Wet extreme mass ratio inspirals may be more common for spaceborne gravitational wave detection

Zhen Pan,[1,*] Zhenwei Lyu,[1,2] and Huan Yang[1,2,†]

[1]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[2]*University of Guelph, Guelph, Ontario N1G 2W1, Canada*

(Received 7 April 2021; accepted 13 August 2021; published 3 September 2021)

Extreme mass ratio inspirals (EMRIs) can be classified as dry EMRIs and wet EMRIs based on their formation mechanisms. Dry (or the "loss-cone") EMRIs, previously considered as the main EMRI sources for the Laser Interferometer Space Antenna, are primarily produced by multibody scattering in the nuclear star cluster and gravitational capture. In this work, we highlight an alternative EMRI formation channel: (wet) EMRI formation assisted by the accretion flow around accreting galactic-center massive black holes (MBHs). In this channel, the accretion disk captures stellar-mass black holes that are initially moving on inclined orbits and subsequently drives them to migrate towards the MBH—this process boosts the formation rate of EMRIs in such galaxies by orders of magnitude. Taking into account the fraction ($\mathcal{O}(10^{-2}$–$10^{-1})$) of active galactic nuclei where the MBHs are expected to be rapidly accreting, we forecast that wet EMRIs will contribute an important or even dominant fraction of all detectable EMRIs by spaceborne gravitational wave detectors.

DOI: 10.1103/PhysRevD.104.063007

I. INTRODUCTION

The primary astrophysical sources for space-based gravitational wave detectors, such as Laser Interferometer Space Antenna (LISA) [1] and TianQin [2], include massive black hole (MBH) bianaries, extreme mass ratio inspirals (EMRIs) [3], galactic binaries and stellar-mass black hole (SBH) binaries. Other systems, e.g., intermediate mass ratio inspirals [4,5], extremely large mass ratio inspirals [6,7] and cosmic strings [8], may also be detectable, albeit with larger uncertainties. Among these sources, EMRIs provide unique opportunities in testing the Kerr spacetime [9,10], probing the galactic-center cluster distribution [11–13], understanding the astrophysical environmental effects [14–16], and inferring the growth history of MBHs [17–19]. Loud EMRIs can serve as dark standard sirens for measuring the Hubble constant H_0 and the dark energy equation of state [20].

EMRI formation mechanism can be classified into two main channels. In the "dry EMRI" channel, an EMRI may be produced after a SBH is gravitationally captured by a MBH, following the multibody scatterings within the nuclear cluster [21–23] (other processes involving tidal disruption or tidal capture of binary SBHs, or tidal stripping of giant stars [24–27] may also contribute a fraction of dry EMRIs). There are two characteristic timescales [28,29] in this process: the GW emission timescale $t_{\rm gw}$ on which the SBH orbit shrinks and the relaxation timescale t_J on which the orbital angular momentum of the SBH changes, due to scatterings by stars and other SBHs. If $t_{\rm gw} > t_J$, then the SBH will likely be randomly scattered either into or away from the MBH (SBHs scattered into the MBH are known as prompt infalls). If $t_{\rm gw} < t_J$, then the SBH orbit gradually spirals into the MBH to form an EMRI while random scatterings are negligible. The generic rate can be obtained by solving the Fokker-Planck equation or by N-body simulations [11–13], subject to assumptions on the initial distributions of stars and SBHs in the nuclear cluster. In addition to the generic rate per MBH, the EMRI rate density in the Universe also depends on the mass function of MBHs, the fraction of MBHs living in stellar cusps and the relative abundance of SBHs in stellar clusters. Taking into account these astrophysical uncertainties, Babak *et al.* [22,23] and Fan *et al.* [30] forecasted that there will be a few to thousands of detectable (dry) EMRIs per year by LISA and TianQin, respectively. In a recent paper [31], Zwick *et al.* reanalyzed the GW emission timescales of inspiraling eccentric binaries and realized post-Newtonian (PN) corrections to the commonly used Peters's formula [32] are necessary. With PN corrections implemented, the dry EMRI rate decreases by approximately at least one order of magnitude [33].

Wet EMRIs come from MBHs in gas-rich environments, where the distributions of nearby stars and SBHs are significantly affected by the accretion flow. About 1% low-redshift ($z \lesssim 1$) galaxies and 1%–10% high-redshift ($1 \lesssim z \lesssim 3$) galaxies are active [34,35] and known as active galactic nuclei (AGNs), in which galactic MBHs are

*zpan@perimeterinstitute.ca
†hyang@perimeterinstitute.ca

believed to be rapidly accreting gas in a disk configuration. In the presence of an accretion disk, the periodic motion of a SBH generally generates density waves which in turn affect the SBH's motion by damping both the orbital inclination with respect to the disk plane and the orbital eccentricity, and driving the SBH's migration in the radial direction [36–39]. As long as the SBH is captured onto the disk, the density waves together with other disk-SBH interactions, e.g., head wind [40,41], accelerate its inward migration until the vicinity of the MBH where GW emissions become prevalent. In addition to SBHs captured onto the disk, star formation and subsequent birth of SBHs in the AGN disk may also contribute to wet EMRI formation [42–44]. In this paper, we show that an accretion disk usually boosts the EMRI intrinsic rate per individual MBH by orders of magnitude compared with the loss-cone channel [45]. In particular, we suggest that wet EMRI formation is an important or even dominant channel for all observable EMRIs by spaceborne GW detectors.

The remaining part of this paper is organized as follows. In Sec. II, we summarize the interactions of AGN disks with SBHs and stars. In Sec. III, we introduce the Fokker-Planck equation that governs the evolution of SBHs and stars in a cluster with or without the presence of an AGN disk. In Sec. IV, we present the generic dry EMRI rate per MBH and the wet EMRI rate per AGN. In Sec. V, we calculate the LISA detectable EMRI rate from both channels, and we discuss the applications of wet EMRIs in Sec. VI.

Throughout this paper, we will use geometrical units $G = c = 1$ and assume a flat ΛCDM cosmology with $\Omega_{\rm m} = 0.307$, $\Omega_\Lambda = 1 - \Omega_{\rm m}$ and $H_0 = 67.7$ km/s/Mpc.

II. DISK-SBH AND DISK-STAR INTERACTIONS

In addition to the gravitational forces from the MBH and the stars/SBHs in the cluster, the orbital motion of a SBH around an accreting MBH is influenced by disk-SBH interactions: density waves, head wind [36–39,41], and other subdominant interactions including dynamic friction [46,47] and heating torque [48,49].

As a SBH orbits around the MBH, its periodic motion excites density waves consisting of three components [38,39]: regular density waves arising from the circular motion, eccentricity waves arising from the noncircular motion and bending waves arising from the motion normal to the disk. The density waves in turn affect the motion of the SBH: the regular density waves exert a (type-I) migration torque on the SBH and drives its migration in the radial direction on the timescale $t_{\rm mig,I}$; the eccentricity and bending density waves damp the orbit eccentricity and the inclination with respect to the disk plane on the timescale $t_{\rm wav}$. Previous analytic studies [38,39] calibrated with numerical simulations [50] show that the type-I migration torque can be formulated as

$$\dot{J}_{\rm mig,I} = C_{\rm I}\frac{m_{\rm bh}}{M}\frac{\Sigma}{M}\frac{r^4\Omega^2}{h^2}, \tag{1}$$

where $m_{\rm bh}$ is a SBH mass, and $M = M(<r)$ is the total mass of the MBH, stars, SBHs and the disk within radius r; the prefactor $C_{\rm I} = -0.85 + d\log\Sigma/d\log r + 0.9d\log T_{\rm mid}/d\log r$ depends on the disk profile; $\Sigma(r), T_{\rm mid}(r), h(r), \Omega(r)$ are the disk surface density, the disk middle plane temperature, the disk aspect ratio, and the SBH angular velocity, respectively. The corresponding migration timescale and damping timescale are

$$t_{\rm mig,I} = \frac{J}{|\dot{J}_{\rm mig,I}|} \sim \frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^2}{\Omega}, \qquad t_{\rm wav} = \frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^4}{\Omega}, \tag{2}$$

where $J = r^2\Omega$ is the specific angular momentum of the SBH, and $t_{\rm wav} \approx t_{\rm mig,I}h^2$, i.e., the eccentricity/inclination damping is faster than the migration by a factor h^2. Therefore the orbit should become circular long before the SBH migrate into the LISA band. A gap in the disk opens up if the SBH is so massive that its tidal torque removes surrounding gas faster than the gas replenishment via viscous diffusion. After a gap is opened, the type-I migration turns off and the SBH is subject to type-II migration driven by a type-II migration torque $\dot{J}_{\rm mig,II}$ [51].

For a SBH embedded in the gas disk, surrounding gas in its gravitational influence sphere flows towards it. Considering the differential rotation of the disk, the inflow gas generally carries nonzero angular momentum relative to the SBH, so that the inflow tends to circularize and form certain local disk or bulge profile around the SBH. Depending on the radiation feedback and magnetic fields, a major part of captured materials may escape in the form of outflow and only the remaining part is accreted by the SBH [52,53]. Because of the circularization process, it is reasonable to expect that the outflow carries minimal net momentum with respect to the SBH. As a result, the head wind in the influence sphere of the SBH is captured, and the momentum carried by the wind eventually transfers to the SBH. Therefore the specific torque exerted on the SBH from the head wind is

$$\dot{j}^{\rm id}_{\rm wind} = -\frac{r\delta v_\phi \dot{m}_{\rm gas}}{m_{\rm bh}}, \tag{3}$$

where the upper index "id" denotes in-disk objects, $\delta v_\phi := v_{\phi,\rm gas} - v_{\phi,\rm bh}$ is the head wind speed, and $\dot{m}_{\rm gas}$ is the amount of gas captured per unit time (see [13] for detailed calculation).

In summary, the migration timescales of in-disk (id) SBHs and those outside (od) are

$$t_{\rm mig}^{\rm bh,id}=\frac{J}{|\dot{J}_{\rm mig,I,II}+\dot{J}_{\rm gw}+\dot{J}_{\rm wind}|},\qquad t_{\rm mig}^{\rm bh,od}=\frac{J}{|\dot{J}_{\rm mig,I}+\dot{J}_{\rm gw}|},\tag{4}$$

where $\dot{J}_{\rm mig,I,II}=\dot{J}_{\rm mig,I}$ or $\dot{J}_{\rm mig,II}$ and $\dot{J}_{\rm wind}=\dot{J}_{\rm wind}^{\rm id}$ [Eq. (3)] or 0, depending on whether a gap is open. The specific torque arising from GW emissions is [54]

$$\dot{J}_{\rm gw}=-\frac{32}{5}\frac{m_{\rm bh}}{M}\left(\frac{M}{r}\right)^{7/2}.\tag{5}$$

The damping timescale of SBH orbital inclination and eccentricity is given by Eq. (2)

$$t_{\rm wav}^{\rm bh,od}=\frac{M}{m_{\rm bh}}\frac{M}{\Sigma r^2}\frac{h^4}{\Omega}.\tag{6}$$

The above discussion of disk-SBH interactions also equally applies to stars in the cluster, except stars are usually lighter ($m_{\rm star}<m_{\rm bh}$), and the head wind impact on stars is weak ($\dot{J}_{\rm wind}^{\rm star}\approx 0$) considering that the wind could be largely suppressed in the presence of star radiation feedback and solar wind [55,56]. Because the structure of AGN disks has not been fully understood, we consider three commonly used AGN disk models: α disk, β disk [57], and TQM disk (an AGN disk model proposed by Thompson, Quataert, and Murray [58]) in this work.

III. FOKKER-PLANCK EQUATION

Statistical properties of stars and SBHs in the stellar cluster are encoded in their distribution functions $f_i(t,E,R)$ ($i=$ star/bh) in the phase space, where

$$E\coloneqq\phi(r)-v^2(r)/2,\qquad R\coloneqq J^2/J_c^2(E)\tag{7}$$

are the specific orbital (binding) energy and the normalized orbital angular momentum, respectively. Here $\phi(r)$ is the (positive) gravitational potential, v is the orbital speed, and $J_c(E)$ is the specific angular momentum of a circular orbiter with energy E. Given initial distributions $f_i(t=0,E,R)$, the subsequent evolution is governed by the orbit-averaged Fokker-Planck equation. In the case of no gas disk, the Fokker-Planck equation (for both stars and SBHs) is formulated as [59–61]

$$\mathcal{C}\frac{\partial f}{\partial t}=-\frac{\partial}{\partial E}F_E-\frac{\partial}{\partial R}F_R,\tag{8}$$

where $f=f_i(t,E,R)$, $\mathcal{C}=\mathcal{C}(E,R)$ is a normalization coefficient, and $F_{E,R}$ is the flux in the E/R direction:

$$\begin{aligned}-F_E&=\mathcal{D}_{EE}\frac{\partial f}{\partial E}+\mathcal{D}_{ER}\frac{\partial f}{\partial R}+\mathcal{D}_E f,\\ -F_R&=\mathcal{D}_{RR}\frac{\partial f}{\partial R}+\mathcal{D}_{ER}\frac{\partial f}{\partial E}+\mathcal{D}_R f,\end{aligned}\tag{9}$$

where the diffusion coefficients $\{\mathcal{D}_{EE},\mathcal{D}_{ER},\mathcal{D}_{RR}\}_i$ and the advection coefficients $\{\mathcal{D}_E,\mathcal{D}_R\}_i$ are functions of $f_i(t,E,R)$ [59–61]. From flux $\{F_E,F_R\}_{\rm bh}$, we can compute the EMRI rate via the lose cone mechanism as

$$\Gamma_{\rm lc}(t)=\int_{E>E_{\rm gw}}\vec{F}\cdot d\vec{l},\tag{10}$$

where $\vec{F}=(F_E,F_R)$, $d\vec{l}=(dE,dR)$ is the line element along the boundary of the loss cone, and $E_{\rm gw}$ is a characteristic energy scale above which the SBH GW emission is dominant with $t_{\rm gw}<t_J$ [11–13,28,31,33].

In the presence of an AGN disk, stars, and SBHs settle as two components: a cluster component and a disk component. We expect the distribution functions of cluster-component stars and SBHs acquire some dependence on the orbital inclination as interacting with the disk. For convenience, we choose to integrate out the inclination and work with the inclination-integrated distribution functions $f_i(t,E,R)$ of the cluster-component stars and SBHs. Considering the density waves excited on the disk to damp the orbital inclinations and eccentricities of orbiters, and to drive the orbiters' inward migration together with head winds and GW emissions, we rewrite the Fokker-Planck equation as

$$\mathcal{C}\frac{\partial f}{\partial t}=-\frac{\partial}{\partial E}F_E-\frac{\partial}{\partial R}F_R+S,\tag{11}$$

where flux $F_{E,R}$ are defined in Eq. (9), with the advection coefficients modified by disk-star/SBH interactions as

$$\begin{aligned}&\mathcal{D}_{E,\rm bh}\to\mathcal{D}_{E,\rm bh}-\mathcal{C}\frac{E}{t_{\rm mig}^{\rm bh,od}},\quad \mathcal{D}_{R,\rm bh}\to\mathcal{D}_{R,\rm bh}-\mathcal{C}\frac{1-R}{t_{\rm wav}^{\rm bh,od}},\\ &\mathcal{D}_{E,\rm star}\to\mathcal{D}_{E,\rm star}-\mathcal{C}\frac{E}{t_{\rm mig}^{\rm star,od}},\quad \mathcal{D}_{R,\rm star}\to\mathcal{D}_{R,\rm star}-\mathcal{C}\frac{1-R}{t_{\rm wav}^{\rm star,od}},\end{aligned}$$

and the negative source term $S=S_i(t,E,R)$ arising from spherical-component stars/SBHs captured onto the disk is parametrized as

$$S_{\rm bh}=-\mu_{\rm cap}\mathcal{C}\frac{f_{\rm bh}}{t_{\rm mig}^{\rm star,id}},\quad S_{\rm star}=-\mu_{\rm cap}\frac{m_{\rm star}}{m_{\rm bh}}\mathcal{C}\frac{f_{\rm star}}{t_{\rm mig}^{\rm star,id}},\tag{12}$$

with $\mu_{\rm cap}\in[h,1]\frac{m_{\rm bh}}{m_{\rm star}}$ as a phenomenological parameter quantifying the disk capture efficiency (see [13] for more details). A new EMRI forms if a SBH is captured onto the disk and migrate to the vicinity of the MBH within the disk lifetime $T_{\rm disk}$, therefore the EMRI rate assisted by the AGN disk is given by

$$\Gamma_{\rm disk}(t;T_{\rm disk})=\int\int_{t_{\rm mig}^{\rm bh,id}<T_{\rm disk}}-S_{\rm bh}(t,E,R)dEdR.\tag{13}$$

IV. EMRI RATE PER MBH/AGN: DRY AND WET

A. Dry EMRIs

Given initial distributions of stars and SBHs in the stellar cluster, one can evolve the system according to the Fokker-Planck equation (8) and calculate the EMRI rate in the loss cone channel using Eq. (10). As shown in Refs. [11–13], the EMRI rate mainly depends on the total number of stars within the MBH influence radius, which determines the relaxation timescale and the relative abundance of SBHs in the stellar cluster. Following Ref. [22], the time-averaged EMRI rate per MBH can be parametrized as

$$\Gamma_{\rm dry}(M_\bullet;N_p)=C_{\rm dep}(M_\bullet;N_p)C_{\rm grow}(M_\bullet;N_p)\Gamma_{\rm lc}(M_\bullet), \quad (14)$$

with

$$\Gamma_{\rm lc}(M_\bullet)=30\left(\frac{M_\bullet}{10^6\ M_\odot}\right)^{-0.19}\ {\rm Gyr}^{-1}, \quad (15)$$

where N_p is the average number of prompt infalls per EMRI; $C_{\rm dep}$ and $C_{\rm grow}$ are correction factors accounting for possible depletion of SBHs in the cusp as SBHs accreted by the MBH and capping the maximum MBH growth via accreting SBHs, respectively, and the loss-cone EMRI rate in Eq. (15) is lower than previous calculations [11–13,22] by one order of magnitude because these previous results were based on Peters's formula [32], which underestimate the GW emission timescales of eccentric binaries and the true EMRI rate should be lower by approximately at least one order of magnitude [31,33].

Following Ref. [22], we explain the two corrections $C_{\rm dep}(M_\bullet;N_p)$ and $C_{\rm grow}(M_\bullet;N_p)$ to the generic dry EMRI rate. Consider a MBH with mass $M_\bullet$, whose influence sphere ($r<r_c=2M_\bullet/\sigma^2$) encloses a number of SBHs with total mass $\Sigma m_{\rm bh}\simeq 0.06M_\bullet$, and these SBHs will be depleted by the MBH via EMRIs and prompt infalls on a timescale

$$\begin{aligned}T_{\rm dep}(r_c)&=\frac{\sum m_{\rm bh}}{(1+N_p)\Gamma_{\rm lc}(M_\bullet)m_{\rm bh}},\\&=\frac{200}{1+N_p}\left(\frac{m_{\rm bh}}{10\,M_\odot}\right)^{-1}\left(\frac{M_\bullet}{10^6\,M_\odot}\right)^{1.19}{\rm Gyr},\end{aligned} \quad (16)$$

where N_p is the average number of prompt infalls per EMRI. On the influence sphere, the relaxation timescale of the star cluster is approximately [61]

$$T_{\rm rlx}(r_c)\simeq\left(\frac{\sigma}{20\ {\rm km/s}}\right)\left(\frac{r_c}{1\ {\rm pc}}\right)^2\ {\rm Gyr}, \quad (17)$$

where the velocity dispersion is related to the MBH mass by the famous $M_\bullet-\sigma$ relation [62]. The depletion correction $C_{\rm dep}$ is defined as

$$C_{\rm dep}(M_\bullet;N_p):=\min.\left\{\frac{T_{\rm dep}}{T_{\rm rlx}},1\right\}, \quad (18)$$

where

$$\frac{T_{\rm dep}}{T_{\rm rlx}}\simeq\frac{12}{1+N_p}\left(\frac{m_{\rm bh}}{10\ M_\odot}\right)^{-1}\left(\frac{M_\bullet}{10^6\ M_\odot}\right)^{0.06}. \quad (19)$$

The growth correction

$$C_{\rm grow}:=\min.\left\{e^{-1}\frac{M_\bullet}{\Delta M_\bullet},1\right\} \quad (20)$$

arises from requiring the MBH grows no more than e^{-1} via accreting SBHs, where

$$\Delta M_\bullet=m_{\rm bh}(1+N_p)C_{\rm dep}(M_\bullet;N_p)\Gamma_{\rm lc}(M_\bullet)T_{\rm emri}(M_\bullet), \quad (21)$$

is the MBH growth via accreting SBHs, and

$$T_{\rm emri}(M_\bullet)=\int dz\frac{dt}{dz}C_{\rm cusp}(M_\bullet,z) \quad (22)$$

is the effective growth time when the MBH lives in a stellar cusp.

In Fig. 1, we show three sample models of dry EMRIs with $N_p=\{0,10,10^2\}$, where $\Gamma_{\rm dry}(N_p=0)$ is the same as the generic rate [Eq. (15)] in the mass range of interest, $\Gamma_{\rm dry}(N_p=10)$ is capped by the accretion growth limit $C_{\rm grow}$ for light MBHs, and $\Gamma_{\rm dry}(N_p=10^2)$ is further reduced by the SBH depletion $C_{\rm dep}$ across the entire mass range.

B. Wet EMRIs

More technical complications are involved in calculating the wet EMRI rate due to the uncertainties in AGN accretion history, AGN accretion disks and initial conditions of stellar clusters, which we outline as follows in accordance with our previous work [13].

In this paper, we conservatively assume a constant AGN fraction $f_{\rm AGN}=1\%$ throughout the Universe, though it can be 10 times higher [34,35]. Being consistent with the AGN fraction, the total duration of active phases of an AGN is about 10^8 yr [63], therefore MBHs are in quiet phase most of the time. Another complication is that AGN accretion is likely episodic [64,65]; i.e., a MBH may become active for multiple times during its whole life. Without detailed knowledge of the duty cycle of an MBH, we simplify it as a long quiet phase of $T_0=5$ Gyr followed by a short active phase of $T_{\rm disk}=10^7$ or 10^8 yr. For $T_{\rm disk}=10^8$ yr, there is only one active phase. On the other hand, there are on average 10 active phases for models with $T_{\rm disk}=10^7$ yr, and we only consider the following two extremal cases. If a (low-redshift) AGN that has gone through all the 10 active phases and the

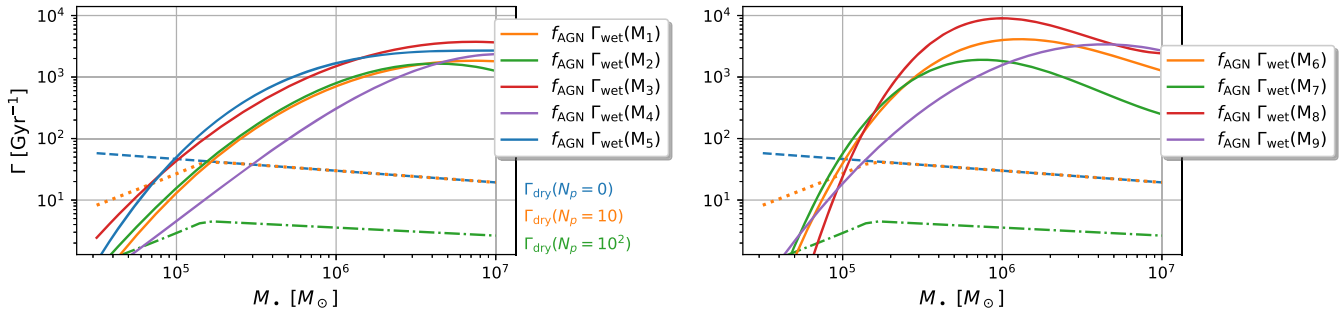

FIG. 1. Average EMRI rates per MBH in the loss cone channel $\Gamma_{\rm dry}(M_\bullet; N_p)$ and per AGN in the disk channel $\Gamma_{\rm wet}(M_\bullet; \mathbb{M})$, where N_p is the number of prompt infalls per EMRI, and $\mathbb{M}$ consists of all model parameters of initial condition of stellar clusters, AGN duty cycles, and AGN disk model, where the AGN fraction is $f_{\rm AGN} = 1\%$.

relaxation between different active phases is not expected to substantially change the SBH distribution, then the average EMRI rate is approximately the same to that in the case of $T_{\rm disk} = 10^8$ yr. If a (high-redshift) AGN has gone through only 1 active phase, then the duty cycle is simply a long quiet phase with duration $T_0 = 5$ Gyr followed by a short active phase with duration $T_{\rm disk} = 10^7$ yr. That is to say, $T_{\rm disk}$ in our model is approximately the total duration of *all* active phases an AGN has gone through.

The structure of AGN disks has not been fully understood either, partially due to the large range over which an AGN disk extends: from an inner radius of a few gravitational radii of the MBH to the outer radius of parsec scale where the AGN disk connects to the galactic gas disk. Three commonly used AGN disk models, α disk, β disk [57], and TQM disk [58], are different in their prescriptions of disk viscosity and/or disk heating mechanism, which lead to large differences in predicted disk structures. Each disk model is specified by two model parameters, the MBH accretion rate $\dot{M}_\bullet$ and a viscosity parameter [13]: an α parameter which prescribes the ratio between the viscous stress and the local total/gas pressure in the α/β disk and a X parameter that prescribes the ratio between the radial gas velocity and the local sound speed in the TQM disk. For calculating the wet EMRI rate, we consider both an α-disk model [57] with the viscosity parameter $\alpha = 0.1$ and accretion rate $\dot{M}_\bullet = 0.1\dot{M}_\bullet^{\rm Edd}$, and a TQM disk [58] with the viscosity parameter $X = 0.1$ and accretion rate $\dot{M}_\bullet = 0.1\dot{M}_\bullet^{\rm Edd}$ (β disk is different from α disk only in the inner region where radiation pressure dominates over gas pressure, and this difference has little impact on the wet EMRI rate).

For calculating the wet EMRI rate, we also need to specify the initial distributions of stars and SBHs in the stellar cluster, which we assume the commonly used Tremaine's cluster model [66] with a power-law density profile $n_{\rm star}(r) \sim r^{-\gamma}$ deep inside the influence sphere of the MBH and $n_{\rm star}(r) \sim r^{-4}$ far outside, and SBHs are of the same density profile with a relative abundance δ.

Given initial distributions of stars and SBHs in the stellar cluster, we first evolve the system for time T_0 according to the Fokker-Planck equation (8), then turn on an accretion disk and continue the evolution for time $T_{\rm disk}$, according to the modified Fokker-Planck equation (11). In the active phase, the disk assisted EMRI rate is computed using Eq. (13). We show the time-averaged EMRI rate per AGN

$$\Gamma_{\rm wet}(M_\bullet; \mathbb{M}) = \frac{1}{T_{\rm disk}} \int_{T_0}^{T_0+T_{\rm disk}} \Gamma_{\rm disk}(t, M_\bullet; \mathbb{M}) dt, \quad (23)$$

for different models $\mathbb{M}$ in Fig. 1, where $\mathbb{M}$ denotes models parameterizing initial distributions of stars and SBHs in the cluster, duty cycles of MBHs and AGN disk model (see Table I for model parameters for all nine models considered in this work).

Because SBHs are captured onto the disk and migrate inward efficiently, and the SBH loss via prompt infalls is negligible ($N_p \ll 1$), the wet EMRI rate is mainly limited by the number of SBHs available in the stellar cluster. As a result, we find the presence of an AGN disk usually boosts the EMRI formation rate by orders of magnitude [13,67] regardless of the variations of different disk models considered.

V. TOTAL AND LISA DETECTABLE EMRI RATES

For calculating the total EMRI rate, we consider two redshift-independent MBH mass functions in the range of $(10^4, 10^7) M_\odot$,

$$f_{\bullet,-0.3}: \frac{dN_\bullet}{d\log M_\bullet} = 0.01 \left(\frac{M_\bullet}{3\times 10^6\ M_\odot}\right)^{-0.3} {\rm Mpc}^{-3},$$
$$f_{\bullet,+0.3}: \frac{dN_\bullet}{d\log M_\bullet} = 0.002 \left(\frac{M_\bullet}{3\times 10^6\ M_\odot}\right)^{+0.3} {\rm Mpc}^{-3}, \quad (24)$$

where the former one is approximate to the mass function as modeled in Refs. [68–71] assuming MBHs were seeded by Population III stars and accumulated mass via mergers and gas accretion along cosmic history, and the latter one is a phenomenological model [18]. The differential EMRI

TABLE I. Comparison of dry and wet EMRI rates in different models, where $f_\bullet$ is the MBH mass function. The last two columns are the total EMRI rate in the redshift range of $0 < z < 4.5$ and the corresponding LISA detectable (SNR ≥ 20) rate.

Dry EMRIs	$f_\bullet$	N_p					Total rate [yr^{-1}]	LISA detectable rate [yr^{-1}]
	$f_{\bullet,-0.3}$	0					3500	150
		10					1300	120
		10^2					150	14
	$f_{\bullet,+0.3}$	0					160	10
		10					130	10
		10^2					15	1
Wet EMRIs	$f_\bullet$	$\mathbb{M}$:	(γ, δ)	$\mu_{\rm cap}$	$(T_{\rm disk}[{\rm yr}], f_{\rm AGN})$	AGN Disk	Total rate [yr^{-1}]	LISA detectable rate [yr^{-1}]
	$f_{\bullet,-0.3}$	M_1:	(1.5, 0.001)	1	$(10^8, 1\%)$	α disk	11000	600
		M_2:	(1.5, 0.001)	0.1			11000	760
		M_3:	(1.5, 0.002)	1			24000	1500
		M_4:	(1.8, 0.001)	1			8100	240
		M_5:	(1.5, 0.001)	1	$(10^8, 1\%)$	TQM disk	23000	1900
		M_6:	(1.5, 0.001)	1	$(10^7, 1\%)$	α disk	39000	4200
		M_7:	(1.5, 0.001)	0.1			21000	3000
		M_8:	(1.5, 0.002)	1			80000	9800
		M_9:	(1.8, 0.001)	1			22000	1400
	$f_{\bullet,+0.3}$	M_1:	(1.5, 0.001)	1	$(10^8, 1\%)$	α disk	2100	49
		M_2:	(1.5, 0.001)	0.1			2000	57
		M_3:	(1.5, 0.002)	1			4300	100
		M_4:	(1.8, 0.001)	1			1900	18

rates (in observer's frame) in the two formation channels are written as

$$\frac{d^2\mathcal{R}_{\rm dry}}{dM_\bullet dz} = \frac{1}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz} C_{\rm cusp}(M_\bullet, z)\Gamma_{\rm dry}(M_\bullet; N_p),$$
$$\frac{d^2\mathcal{R}_{\rm wet}}{dM_\bullet dz} = \frac{f_{\rm AGN}}{1+z}\frac{dN_\bullet}{dM_\bullet}\frac{dV_{\rm c}(z)}{dz} C_{\rm cusp}(M_\bullet, z)\Gamma_{\rm wet}(M_\bullet; \mathbb{M}), \tag{25}$$

where the factor $1/(1+z)$ arises from the cosmological redshift, $V_{\rm c}(z)$ is the comoving volume of the Universe up to redshift z, $C_{\rm cusp}(M_\bullet, z)$ is the fraction of MBHs living in stellar cusps which are supposed to be evacuated during mergers of binary MBHs and regrow afterwards [68–71]. For cases with mass function $f_{\bullet,+0.3}$, we use the same $C_{\rm cusp}$ function as in [22] and we simply take $C_{\rm cusp} = 1$ for cases with phenomenological mass function $f_{\bullet,+0.3}$ [72].

In order to calculate the LISA detectable EMRI rate in each channel, we construct a population of EMRIs with SBH mass $m_{\rm bh} = 10\ M_\odot$, MBH spin $a = 0.98$, and MBH masses and redshifts randomly sampled according to the differential EMRI rates [Eq. (25)]. For each individual EMRI, we need 10 more parameters to uniquely specify its binary configuration at coalescence and its gravitational waveform [73–75]: sky localization $\hat{\boldsymbol{n}}$, MBH spin direction $\hat{\boldsymbol{a}}$, three phase angles, coalescence time t_0, inclination angle ι_0, and eccentricity e_0 at coalescence. For both dry and wet EMRIs, we assume that the sky locations and the MBH spin directions are isotropically distributed on the sphere, three phase angles are uniformly distributed in $[0, 2\pi]$, coalescence times are randomly sampled from [0, 2] yr, and cosines of inclination angles are randomly sampled from $[-1, 1]$. Distributions of eccentricity e_0 are different: uniform distribution of e_0 in [0, 0.2] for dry EMRIs vs $e_0 = 0$ for wet EMRIs.

For each EMRI, we compute its time-domain waveform $h_{+,\times}(t)$ using the augment analytic kludge [73–75] with the conservative Schwarzschild plunge condition, because the PN corrections used for constructing the augment analytic kludge waveform model are increasingly inaccurate as the orbital separation decreases. Extending the waveform to the Kerr last stable orbit likely leads to an overestimate of the signal-to-noise ratio (SNR) [22,75]. The SNR is calculated as a noise weighted inner product in the frequency domain [76]

$$\mathrm{SNR} = \sqrt{4\int_0^\infty \frac{h_+(f)h_+^*(f) + h_\times(f)h_\times^*(f)}{S_{\rm n}(f)} df}, \tag{26}$$

where $S_{\rm n}(f)$ is the sky-averaged detector sensitivity of LISA [22,77]. The expected LISA detectable EMRI rates (SNR ≥ 20) of different models in each mass bin are shown in Fig. 2, and the total event rates and the LISA detectable rates are collected in Table I. From Fig. 2 and Table I, wet EMRI formation is evidently an important or even dominant channel for all the models we have considered.

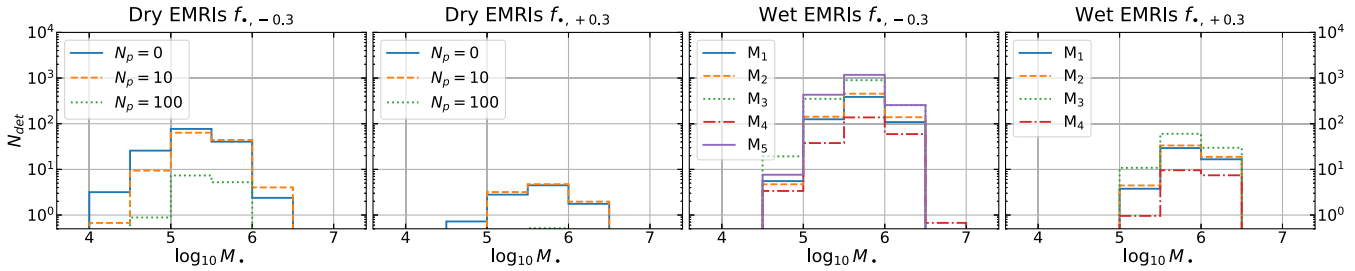

FIG. 2. Forecasted LISA detectable dry and wet EMRI rates $N_{\rm det}$ per mass bin ($M_\bullet[M_\odot]$) per year for different models, where $f_{\bullet,\pm 0.3}$ are the two different MBH mass functions [Eq. (24)], N_p is the number of prompt infalls per EMRI in the loss cone channel, and the wet EMRI model parameters of $\mathrm{M}_{1,\ldots,5}$ are detailed in Table I.

VI. APPLICATIONS OF WET EMRIs

Due to the high LISA sensitivity to the EMRI eccentricity whose value at coalescence can be measured with typical uncertainty as low as 10^{-5} [22], wet EMRIs can be distinguished from dry ones via eccentricity measurements, as wet EMRIs are expected to be circular in the LISA band as a result of the efficient eccentricity damping by the density waves ($t_{\rm wav} \ll t_{\rm mig}$), while dry EMRIs from the loss-cone channel are highly eccentric as entering the LISA band and remain mildly eccentric at coalescence [22,28]. Another subdominant dry EMRI channel involving tidal stripping of giant stars seems unlikely to produce such circular EMRIs either [26], while the prediction of the channel involving tidal disruption of binary SBHs is more uncertain [24,27]. The disk-environmental effects may produce measurable phase shift in the EMRI waveform [40,41,78].

EMRIs have unprecedented potential to probe fundamental laws of gravity and the nature of dark matter [9,10,79,80]. In previous studies, such tests using EMRIs have been implicitly assumed in vacuum without any environmental contamination. However, as we have shown here, wet EMRIs are possibly more common in the Universe, for which the environmental effects on the EMRI waveform are inevitable. The possible degeneracy calls for a systematic framework for searching new fundamental physics with EMRIs, with astrophysical environmental effects taken into account.

In the context of wet EMRIs, AGN jet physics and accretion physics are promising realms where LISA and next-generation Event Horizon Telescope (NGEHT [81]) may synergize. According to the estimate in [19], a fraction of low-redshift ($z \lesssim 0.3$) EMRIs can be traced back to their host galaxies with LISA observations alone, and host AGNs of ∼50% of low-redshift ($z \lesssim 0.5$) wet EMRIs can be identified considering the much lower density of AGNs. Combining GW observations of wet EMRIs with radio obervations of AGN jets by, e.g., NGEHT, one can simultaneously measure the MBH mass $M_\bullet$, the MBH spin $\hat{\boldsymbol{a}}$, the rotation direction of the accretion disk $\hat{\boldsymbol{L}}$, the jet power $\dot{E}_{\rm jet}$ and the jet direction $\hat{\boldsymbol{n}}_{\rm jet}$. This set of observables provide unprecedented opportunities to probe the AGN jet physics. For example, an ensemble of events with $\{\hat{\boldsymbol{n}}_{\rm jet}\cdot\hat{\boldsymbol{a}}, \hat{\boldsymbol{n}}_{\rm jet}\cdot\hat{\boldsymbol{L}}\}$ data may help us to constrain various jet launching models, i.e., powered by the rotating energy of the MBH [82] or by the accretion disk [83]. In addition, certain disk properties are directly constrained with GW observations via the disk environmental effects on the EMRI waveform [40,41,78], and accretion physics of AGN disks is also one of the primary targets of NGEHT.

Wet EMRIs with host AGNs identified are ideal "bright sirens" for constraining the late time cosmology (e.g., the Hubble constant and the equation of state of dark energy), because the luminosity distance and the redshift can be measured from GW and electromagnetic observations, respectively. It will be interesting to compare the sensitivity of this method to other approaches, with the predicted wet EMRI rate from this study.

Wet EMRIs encode additional information of MBH growth in their orbital inclination angles ι_0 with respect to the MBH spin. If all MBHs grow up via coherent gas accretion where gas feeds are from a fixed direction, then orbital inclination angles of wet EMRIs at coalescence should be $\iota_0 \approx \pi/2$. If MBHs grow up via chaotic gas accretion from a random direction in each active phase, then a fraction of wet EMRIs form before the MBH spin direction $\hat{\boldsymbol{a}}$ is aligned with the disk rotation direction $\hat{\boldsymbol{L}}$ via the Bardeen-Petterson mechanism [84], and their orbital inclinations are approximately $\iota_0 \approx \cos^{-1}(\hat{\boldsymbol{a}}\cdot\hat{\boldsymbol{L}})$. In a similar way, MBH growth via different merger channels also imprints differently on the inclinations of wet EMRIs.

ACKNOWLEDGMENTS

We thank Hui-Min Fan and Alberto Sesana for very helpful discussions. The authors are supported by the Natural Sciences and Engineering Research Council of Canada and in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

[1] J. Baker *et al.*, arXiv:1907.06482.
[2] J. Mei, Y.-Z. Bai, J. Bao, E. Barausse *et al.*, Prog. Theor. Exp. Phys. **2021**, 05A107 (2021).
[3] P. Amaro-Seoane, J. R. Gair, M. Freitag, M. C. Miller, I. Mandel, C. J. Cutler, and S. Babak, Classical Quant. Grav. **24**, R113 (2007).
[4] M. Arca-Sedda and R. Capuzzo-Dolcetta, Mon. Not. R. Astron. Soc. **483**, 152 (2019).
[5] M. Arca-Sedda, P. Amaro-Seoane, and X. Chen, arXiv:2007.13746.
[6] P. Amaro-Seoane, Phys. Rev. D **99**, 123025 (2019).
[7] P. Amaro-Seoane, arXiv:2011.03059.
[8] Z. Khakhaleva-Li and C. J. Hogan, arXiv:2006.00438.
[9] K. Glampedakis and S. Babak, Classical Quant. Grav. **23**, 4167 (2006).
[10] L. Barack and C. Cutler, Phys. Rev. D **75**, 042003 (2007).
[11] M. Preto and P. Amaro-Seoane, Astrophys. J. Lett. **708**, L42 (2010).
[12] P. Amaro-Seoane and M. Preto, Classical Quant. Grav. **28**, 094017 (2011).
[13] Z. Pan and H. Yang, Phys. Rev. D **103**, 103018 (2021).
[14] B. Bonga, H. Yang, and S. A. Hughes, Phys. Rev. Lett. **123**, 101103 (2019).
[15] H. Yang, B. Bonga, Z. Peng, and G. Li, Phys. Rev. D **100**, 124056 (2019).
[16] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014).
[17] E. Berti and M. Volonteri, Astrophys. J. **684**, 822 (2008).
[18] J. R. Gair, C. Tang, and M. Volonteri, Phys. Rev. D **81**, 104014 (2010).
[19] Z. Pan and H. Yang, Astrophys. J. **901**, 163 (2020).
[20] D. Laghi, N. Tamanini, W. Del Pozzo, A. Sesana, J. Gair, and S. Babak, arXiv:2102.01708.
[21] P. Amaro-Seoane, Living Rev. Relativity **21**, 4 (2018).
[22] 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, Phys. Rev. D **95**, 103012 (2017).
[23] J. R. Gair, S. Babak, A. Sesana, P. Amaro-Seoane, E. Barausse, C. P. Berry, E. Berti, and C. Sopuerta, J. Phys. Conf. Ser. **840**, 012021 (2017).
[24] M. C. Miller, M. Freitag, D. P. Hamilton, and V. M. Lauburg, Astrophys. J. **631**, L117 (2005).
[25] X. Chen and W.-B. Han, Commun. Phys. **1**, 53 (2018).
[26] Y. Y. Wang, F. Y. Wang, Y. C. Zou, and Z. G. Dai, Astrophys. J. **886**, L22 (2019).
[27] Y. Raveh and H. B. Perets, Mon. Not. R. Astron. Soc. **501**, 5012 (2021).
[28] C. Hopman and T. Alexander, Astrophys. J. **629**, 362 (2005).
[29] B. Bar-Or and T. Alexander, Astrophys. J. **820**, 129 (2016).
[30] H.-M. Fan, Y.-M. Hu, E. Barausse, A. Sesana, J.-d. Zhang, X. Zhang, T.-G. Zi, and J. Mei, Phys. Rev. D **102**, 063016 (2020).
[31] L. Zwick, P. R. Capelo, E. Bortolas, V. Vazquez-Aceves, L. Mayer, and P. Amaro-Seoane, Mon. Not. R. Astron. Soc. **506**, 1007 (2021).
[32] P. C. Peters and J. Mathews, Phys. Rev. **131**, 435 (1963).
[33] V. Vázquez-Aceves, L. Zwick, E. Bortolas, P. R. Capelo, P. Amaro-Seoane, L. Mayer, and X. Chen, arXiv:2108.00135.
[34] A. Galametz, D. Stern, P. R. M. Eisenhardt, M. Brodwin, M. J. I. Brown, A. Dey, A. H. Gonzalez, B. T. Jannuzi, L. A. Moustakas, and S. A. Stanford, Astrophys. J. **694**, 1309 (2009).
[35] M. Macuga, P. Martini, E. D. Miller, M. Brodwin, M. Hayashi, T. Kodama, Y. Koyama, R. A. Overzier, R. Shimakawa, K.-i. Tadaki, and I. Tanaka, Astrophys. J. **874**, 54 (2019).
[36] P. Goldreich and S. Tremaine, Astrophys. J. **233**, 857 (1979).
[37] P. Goldreich and S. Tremaine, Astrophys. J. **241**, 425 (1980).
[38] H. Tanaka, T. Takeuchi, and W. R. Ward, Astrophys. J. **565**, 1257 (2002).
[39] H. Tanaka and W. R. Ward, Astrophys. J. **602**, 388 (2004).
[40] N. Yunes, B. Kocsis, A. Loeb, and Z. Haiman, Phys. Rev. Lett. **107**, 171103 (2011).
[41] B. Kocsis, N. Yunes, and A. Loeb, Phys. Rev. D **84**, 024032 (2011).
[42] Y. Levin, arXiv:astro-ph/0307084.
[43] G. Sigl, J. Schnittman, and A. Buonanno, Phys. Rev. D **75**, 024034 (2007).
[44] Y. Levin, Mon. Not. R. Astron. Soc. **374**, 515 (2007).
[45] More details can be found in a companion paper [13].
[46] S. Chandrasekhar, Astrophys. J. **97**, 255 (1943).
[47] E. C. Ostriker, Astrophys. J. **513**, 252 (1999).
[48] F. S. Masset, Mon. Not. R. Astron. Soc. **472**, 4204 (2017).
[49] A. M. Hankla, Y.-F. Jiang, and P. J. Armitage, Astrophys. J. **902**, 50 (2020).
[50] S. J. Paardekooper, C. Baruteau, A. Crida, and W. Kley, Mon. Not. R. Astron. Soc. **401**, 1950 (2010).
[51] D. Syer and C. J. Clarke, Mon. Not. R. Astron. Soc. **277**, 758 (1995).
[52] X.-H. Yang, F. Yuan, K. Ohsuga, and D.-F. Bu, Astrophys. J. **780**, 79 (2014).
[53] J. C. McKinney, A. Tchekhovskoy, A. Sadowski, and R. Narayan, Mon. Not. R. Astron. Soc. **441**, 3177 (2014).
[54] GW emission turns out to has little influence on the wet EMRI rate, which is determined by the capture and the migration of SBHs at large separations where GW emission is negligible. We include the GW emission in the calculation because in some special cases the migration torque $\dot{J}_{\rm mig,I}$ changes its sign near the MBH and the GW emission works to overcome the would-be migration trap [13].
[55] A. Gruzinov, Y. Levin, and C. D. Matzner, Mon. Not. R. Astron. Soc. **492**, 2755 (2020).
[56] X. Li, P. Chang, Y. Levin, C. D. Matzner, and P. J. Armitage, Mon. Not. R. Astron. Soc. **494**, 2327 (2020).
[57] E. Sirko and J. Goodman, Mon. Not. R. Astron. Soc. **341**, 501 (2003).
[58] T. A. Thompson, E. Quataert, and N. Murray, Astrophys. J. **630**, 167 (2005).
[59] H. Cohn and R. M. Kulsrud, Astrophys. J. **226**, 1087 (1978).
[60] H. Cohn, Astrophys. J. **234**, 1036 (1979).
[61] J. Binney and S. Tremaine, *Galactic Dynamics* (1987).
[62] K. Gültekin, D. O. Richstone, K. Gebhardt, T. R. Lauer, S. Tremaine, M. C. Aller, R. Bender, A. Dressler, S. M. Faber, A. V. Filippenko, R. Green, L. C. Ho, J. Kormendy, J. Magorrian, J. Pinkney, and C. Siopis, Astrophys. J. **698**, 198 (2009).

[63] A. Soltan, Mon. Not. R. Astron. Soc. **200**, 115 (1982).
[64] A. King and C. Nixon, Mon. Not. R. Astron. Soc. **453**, L46 (2015).
[65] K. Schawinski, M. Koss, S. Berney, and L. F. Sartori, Mon. Not. R. Astron. Soc. **451**, 2517 (2015).
[66] S. Tremaine, D. O. Richstone, Y.-I. Byun, A. Dressler, S. M. Faber, C. Grillmair, J. Kormendy, and T. R. Lauer, Astron. J. **107**, 634 (1994).
[67] H. Tagawa, Z. Haiman, and B. Kocsis, Astrophys. J. **898**, 25 (2020).
[68] E. Barausse, Mon. Not. R. Astron. Soc. **423**, 2533 (2012).
[69] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, Astrophys. J. **794**, 104 (2014).
[70] F. Antonini, E. Barausse, and J. Silk, Astrophys. J. **806**, L8 (2015).
[71] F. Antonini, E. Barausse, and J. Silk, Astrophys. J. **812**, 72 (2015).
[72] There was some misplot of the $C_{\rm cusp}(M_\bullet, z)$ function in [22], and we thank Alberto Sesana for kindly providing the correct one.
[73] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).
[74] A. J. Chua and J. R. Gair, Classical Quant. Grav. **32**, 232002 (2015).
[75] A. J. Chua, C. J. Moore, and J. R. Gair, Phys. Rev. D **96**, 044005 (2017).
[76] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).
[77] P. Amaro-Seoane, H. Audley, S. Babak *et al.*, arXiv:1702.00786.
[78] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014).
[79] O. A. Hannuksela, K. W. K. Wong, R. Brito, E. Berti, and T. G. F. Li, Nat. Astron. **3**, 447 (2019).
[80] J. Zhang and H. Yang, Phys. Rev. D **101**, 043020 (2020).
[81] https://www.ngeht.org/.
[82] R. D. Blandford and R. L. Znajek, Mon. Not. R. Astron. Soc. **179**, 433 (1977).
[83] R. D. Blandford and D. G. Payne, Mon. Not. R. Astron. Soc. **199**, 883 (1982).
[84] J. M. Bardeen and J. A. Petterson, Astrophys. J. **195**, L65 (1975).