



The Competing Effect of Gas and Stars in the Evolution of Massive Black Hole Binaries

Elisa Bortolas^{1,2} , Alessia Franchini^{1,2} , Matteo Bonetti^{1,2} , and Alberto Sesana¹

¹ Dipartimento di Fisica “G. Occhialini,” Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy; elisa.bortolas@unimib.it

² INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

Received 2021 July 6; revised 2021 August 6; accepted 2021 August 10; published 2021 August 27

Abstract

Massive black hole binaries are predicted to form during the hierarchical assembly of cosmic structures and will represent the loudest sources of low-frequency gravitational waves (GWs) detectable by present and forthcoming GW experiments. Before entering the GW-driven regime, their evolution is driven by the interaction with the surrounding stars and gas. While stellar interactions are found to always shrink the binary, recent studies predict the possibility of binary outspiral mediated by the presence of a gaseous disk, which could endlessly delay the coalescence and impact the merger rates of massive binaries. Here we implement a semianalytical treatment that follows the binary evolution under the combined effect of stars and gas. We find that binaries may outspiral only if they accrete near or above their Eddington limit and only until their separation reaches the gaseous disk self-gravitating radius. Even in case of an outspiral, the binary eventually reaches a large enough mass for GW to take over and drive it to coalescence. The combined action of stellar hardening, mass growth, and GW-driven inspiral brings binaries to coalescence in a few hundreds of megayears at most, implying that gas-driven expansion will not severely affect the detection prospects of upcoming GW facilities.

Unified Astronomy Thesaurus concepts: [Supermassive black holes \(1663\)](#); [Stellar dynamics \(1596\)](#); [Gravitational wave sources \(677\)](#); [Galaxy accretion disks \(562\)](#); [Computational astronomy \(293\)](#)

1. Introduction

In the past two decades it has been established that massive black holes (MBHs) inhabit the center of massive galaxies (Kormendy & Ho 2013). As galaxies grow by merging with other galaxies along the cosmic history, MBH binaries (MBHBs) are considered to be a natural outcome of the hierarchical structure formation scenario. In fact, following a galaxy merger, the two MBHs hosted by the parent galaxies find themselves inside the same host galaxy (Begelman et al. 1980).

The evolution of such MBH pairs critically depends on the characteristics of the surrounding environment: after a successful dynamical friction-driven inspiral (Chandrasekhar 1943; Capelo & Dotti 2017; Pfister et al. 2019; Bortolas et al. 2020; Bonetti et al. 2020, 2021), bound binaries form in the central parsecs of galaxies. MBHBs can then further shrink down to the gravitational-wave (GW) emission stage through different mechanisms. The most explored ones involve the interaction with either the stellar background (e.g., Khan et al. 2011; Vasiliev et al. 2014; Bortolas et al. 2016, 2018a, 2018b; Gualandris et al. 2017; Varisco et al. 2021) or a gaseous circumbinary disk (e.g., Armitage & Natarajan 2002; Escala et al. 2004; Dotti et al. 2006; Mayer et al. 2007; Cuadra et al. 2009; Lodato et al. 2009). Generally, the evolution is studied either in purely stellar or purely gaseous environments (with the notable exception of Kelley et al. 2017, 2019), but in reality, often both components are present at the same time. While the evolution driven by stars is always found to shrink the binary via three-body stellar encounters (e.g., Quinlan 1996), the effect of the interaction with a gaseous disk is still under debate. The picture outlined in early simulations (e.g., Artymowicz & Lubow 1994), that coherently points toward binary shrinking in response to the interaction with a gaseous disk, has been recently challenged by several studies (Moody et al. 2019; Muñoz et al. 2019; Duffell et al. 2020).

These recent hydrodynamical simulations find possible binary expansion within the range of binary and disk parameters explored.

More recently, it has been pointed out by Tiede et al. (2020) and then confirmed with 3D hydrodynamical simulations of a live binary by Heath & Nixon (2020) that the evolution of the binary semimajor axis depends on the temperature of the disk. Binaries tend to expand only for relatively thick disks, with aspect ratio $H/R \gtrsim 0.05 - 0.2$, the exact threshold depending on the details of the numerical approach employed.

Irrespective of the specific disk configurations, a possible “outspiral” phase is critically relevant in view of the present and future low-frequency GW observatories that are expected to unveil coalescing binaries across the universe, as pulsar timing arrays (PTAs; Desvignes et al. 2016; Reardon et al. 2016; Perera et al. 2019; Alam et al. 2021), and the Laser Interferometer Space Antenna (LISA; Amaro-Seoane et al. 2017; Schödel et al. 2017; Barack et al. 2019). Stalled or expanding binaries could dramatically impact the expected coalescence rates for the forthcoming facilities, therefore affecting the detection prospects and ultimately their mission science cases. On the other hand, long-lived outspiral binaries might be fairly frequent in the universe, increasing the chance of detection in electromagnetic surveys.

In this Letter, we combine the gas-induced expansion found in the aforementioned studies (Muñoz et al. 2019, 2020; Duffell et al. 2020) with the stellar and GW hardening, to verify in which regime the binary is found to possibly stall or expand its separation under the concurrent effect of different evolutionary processes. We consider circular, equal-mass binaries and an infinite supply of gas (see also Muñoz et al. 2020, who compared the orbital decay of MBHBs interacting with finite or infinite supplies of gas), which are idealized assumptions aimed at maximizing the effect of gas outspiral. Nevertheless, we anticipate that we found the bound binary orbit to always decay in a relatively short timescale ($\sim 10^8$ yr),

indicating that the gas-induced binary expansion can only mildly delay the final coalescence, and that gas-driven expansion necessarily has to revert into a shrinking as the binary mass gets large enough due to gas accretion.

The Letter is organized as follows: in Section 2 we describe the semianalytical approach used to study the binary semimajor axis evolution; in Section 3 we outline and comment on the obtained results; in Section 4 we discuss the possible implications of our findings together with the possible caveats and we draw our conclusions.

2. Theoretical Framework

In this section, we present the different prescriptions describing the binary semimajor axis evolution, \dot{a} , as a result of interactions with stars, gas, and of GW emission.

2.1. Binary Evolution in a Stellar Environment

The stellar-driven binary hardening rate can be written as (Quinlan 1996; Sesana & Khan 2015)

$$\dot{a}_* = -\frac{HG\rho}{\sigma} a^2, \quad (1)$$

where ρ and σ are the stellar density and velocity dispersion at the binary influence radius, respectively, and H is a dimensionless quantity that mildly depends on the ratio a/a_h and on the binary mass ratio $q = m_2/m_1 \leq 1$, where m_1 and m_2 are the masses of the two binary components; we also set $m = m_1 + m_2$. The specific value of H ranges between 12 and 20 for binary separations $a < a_h$, where $a_h = Gm_2/(4\sigma^2) \propto m^{0.543}$ (see below) is the so-called hard binary separation, i.e., the separation below which the hardening in purely stellar environments is dominated by three-body interactions, and is well described by Equation (1). This constant is generally computed using scattering experiments, in particular we here use the fit obtained by Sesana et al. (2006).

In order to obtain a general description of the stellar hardening, we need to express ρ and σ in terms of the binary mass m . We make the conservative assumption that no nuclear stellar cluster is present in the system, so that stellar hardening is solely due to interactions of the binary with the bulge structure of the galaxy. We thus express the σ as (Kormendy & Ho 2013)

$$\sigma = 200 \left(\frac{m}{3.09 \times 10^8 M_\odot} \right)^{\frac{1}{4.38}} \text{ km s}^{-1}, \quad (2)$$

while we write the central density as $\rho = 2m/(4/3\pi r_{\text{infl}}^3) \propto m^{-0.68}$, where r_{infl} is the influence radius containing twice the binary mass, which can be obtained from scaling relations and is equal to $r_{\text{infl}} = 35 \left(\frac{m}{10^8 M_\odot} \right)^{0.56} \text{ pc}$ (Merritt et al. 2009). This implementation implies $\dot{a}_* \propto m^{-0.91}$, so that the hardening is less efficient for more massive binaries.

2.2. Binary Evolution in a Gaseous Environment

According to the 2D hydrodynamical simulations performed by Muñoz et al. (2019, 2020), the binary semimajor axis

evolution can be written as

$$\dot{a}_{\text{gas}} = 2.68 \frac{\dot{m}}{m} a, \quad (3)$$

where \dot{m} is the binary mass accretion rate.³ This is essentially Equation (9) in Muñoz et al. (2020)⁴ and is formally valid only for $q \gtrsim 0.1$ (Duffell et al. 2020 explore gas-driven evolution for $q = 0.01 - 1$ and also find positive, nearly constant torques for $q \gtrsim 0.1$) for a prograde binary located in the disk plane.

In our controlled experiment, we are free to choose the mass accretion rate \dot{m} . We explore two distinct possibilities:

1. A fixed physical value \dot{m} throughout the binary evolution, so that the mass flux onto the binary is the same even if the binary mass increases.
2. A fixed fraction f_{Edd} (Eddington ratio) of the Eddington accretion rate, $\dot{m} = f_{\text{Edd}} \dot{m}_{\text{Edd}}$, where

$$\dot{m}_{\text{Edd}} = \frac{4\pi Gm m_p}{\eta \sigma_T c} = 2.26 \times 10^{-2} \left(\frac{\eta}{0.1} \right)^{-1} \times \left(\frac{m}{10^6 M_\odot} \right) M_\odot \text{ yr}^{-1}. \quad (4)$$

Here m_p is the proton mass, c is the speed of light, σ_T is the Thomson scattering cross section, and η is the accretion efficiency, which we assume to be 10%. Note that in this case the accretion rate onto the binary does increase with the binary mass.

If the binary components do not have the same mass, the lighter component is typically found, in numerical simulations, to accrete mass more efficiently compared to the heavier component. The consequence is that all the binaries tend to evolve toward equal mass. In this Letter we thus only show cases for which $q = 1$ from the start. We tested the evolution accounting for the change in the binary mass ratio \dot{q} for $q > 0.1$, using the prescriptions inferred by Farris et al. (2014) and Duffell et al. (2020), and found no significant difference from the results presented here, as the binary becomes equal mass very quickly, in the gas-dominated regime, before entering the GW-driven inspiral, which is the only process whose timescale is sensitive to q .

2.3. Disks in the Self-gravitating Regime

Circumbinary disks surrounding very massive black hole binaries, i.e., $m \gtrsim 10^7 M_\odot$, are likely to be self-gravitating (Franchini et al. 2021). When the disk self-gravity contribution is nonnegligible, the circumbinary disk can either self-regulate on a quasi-stable state through the formation of spirals or can fragment into clumps if the disk cooling mechanism is very efficient compared to the heating generated by the shocks induced by the spirals. Since the evolution in Equation (3) has been inferred for non-self-gravitating disks, we assume the binary gas-driven expansion (Equation (3)) to be suppressed as the binary semimajor axis reaches the radius beyond which the disk self-gravity cannot be neglected. This is motivated by the results of Franchini et al. (2021), who find that, regardless of initial temperature, self-gravitating disks regulate themselves to

³ We note that Equation (3) is valid in a regime where the change in accretion rate occurs over a longer timescale compared to the viscous time. It is easy to verify that this condition is generally fulfilled in cold accretion disks.

⁴ See the upper panel of their Figure 7.

$H/R < 0.1$, promoting binary shrinking. The self-gravitating radius can be expressed as (Perego et al. 2009)

$$R_{\text{sg}} = 1.158 \times 10^{-2} f_{\text{Edd}}^{-22/45} \left(\frac{\alpha}{0.1} \right)^{28/45} \left(\frac{\eta}{0.1} \right)^{22/45} \times \left(\frac{m}{10^6 M_{\odot}} \right)^{-7/45} \text{ pc} \quad (5)$$

for a fixed the Eddington ratio, or as

$$R_{\text{sg}} = 1.817 \times 10^{-3} \left(\frac{\dot{m}}{1 M_{\odot} \text{ yr}^{-1}} \right)^{-22/45} \left(\frac{\alpha}{0.1} \right)^{28/45} \times \left(\frac{m}{10^6 M_{\odot}} \right)^{15/45} \text{ pc} \quad (6)$$

for a fixed accretion rate, where α is the disk viscosity parameter, which encapsulates the angular momentum transport mechanism within the disk (Shakura & Sunyaev 1973).

We assume the expansion to be suppressed when the binary semimajor axis is above $C R_{\text{sg}}$; here we set $C = 0.5$ since the radius of the disk cavity extends to $2a$, and therefore we neglect, as a first approximation, gaseous interaction once the inner disk cavity radius equals the disk self-gravitating radius. We also tried larger values of C finding no significant differences in the global evolution.

The prescription for binary expansion may also break down at the radius beyond which the disk temperature drops below 10^4 K, possibly triggering the so-called ionization instability (e.g., Haiman et al. 2009). This can occur at scales smaller than the self-gravitating radius. We however do not consider the ionization instability radius as a limit for the expansion as there is so far no evidence that gas-induced expansion would be hampered beyond this scale. This also means our treatment is rather conservative, as the binary expansion could in principle be prevented in a wider region of the parameter space.

2.4. Gravitational-wave-driven Orbital Decay

If the binary semimajor axis drops below a certain threshold, then the subsequent evolution of the binary is driven by GW emission. Considering a zero orbital eccentricity, the binary semimajor axis shrinks as (Peters 1964)

$$\dot{a}_{\text{GW}} = -\frac{64}{5} \frac{G^3}{c^5} \frac{q}{(1+q)^2} \frac{m^3}{a^3} \quad (7)$$

until the binary reaches coalescence.

Two examples of the hardening (or expansion) rate as a function of a are presented in Figure 1 for an equal-mass, $10^6 M_{\odot}$ binary with different accretion rates. The plots are produced by keeping the total binary mass m fixed (although the employed prescription for \dot{a}_{gas} in Equation (3) also accounts for the angular momentum change induced by mass accretion into the binary; Muñoz et al. 2020), so that the binary with $f_{\text{Edd}} = 1$ appears to stall once gas starts to dominate its evolution.

2.5. Transition Radii

If we neglect for a moment the fact that the mass of the binary is not fixed in time, the different power-law dependencies of the binary hardening rates with a imply that at large scales stellar hardening ($\dot{a}_{\star} \propto -a^2$) is always the dominant

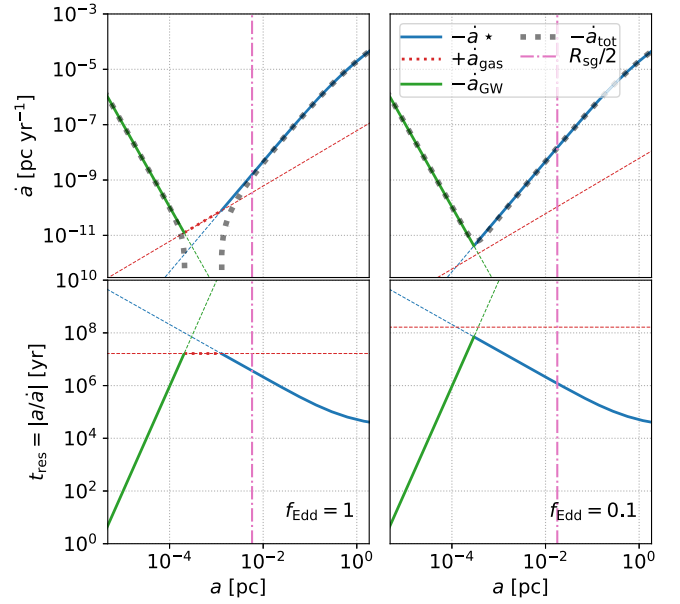


Figure 1. Hardening of a $10^6 M_{\odot}$, equal-mass binary modeled via the prescriptions detailed in Section 2. The top panels show the binary shrinking (or expanding, in the case of gas) rate due to gas, stars, and GWs, and the cumulative hardening \dot{a}_{tot} (only shown when < 0). The bottom panels show the associated residence timescale for all the components. Note that for the gas, the residence time has little meaning as gas tends to expand instead of shrinking the binary in the presented picture. We also show the self-gravitating radius of the disk (Equation (5)) above which the gaseous hardening presented here does not hold anymore. The left panels show the evolution for a binary accreting at \dot{m}_{Edd} (whose shrinking is thus hindered by gas), the right panels refer to a binary with $\dot{m} = 0.1 \dot{m}_{\text{Edd}}$, for which gas plays virtually no role. Very importantly, the evolution shown here does not take into account the time variation of the binary mass, which significantly changes the picture.

evolutionary mechanism. Conversely, at the smallest scales the binary is driven by GWs as $\dot{a}_{\text{GW}} \propto -a^{-3}$. Gas-driven evolution ($\dot{a}_{\text{gas}} \propto a$), which may hinder the shrinking, can thus only be dominant between the stellar- and GW-driven evolution; depending on the involved parameters, it could also be always subdominant.

If gas is ever dominant, the transition between the star- and gas-dominated regimes occurs at

$$a_{\star \rightarrow \text{gas}} = 2.68 \frac{\dot{m}}{m} \frac{\sigma}{H G \rho}. \quad (8)$$

If instead gas is always subdominant, the GW-driven hardening rate is equal to the stellar hardening rate at

$$a_{\star \rightarrow \text{GW}} = \left(\frac{64}{5} \frac{G^2}{c^5} \frac{\sigma}{H \rho} \frac{q}{(1+q)^2} m^3 \right)^{1/5}. \quad (9)$$

Analogously, we can write the transition between the gas-driven to the GW-driven hardening to occur at

$$a_{\text{gas} \rightarrow \text{GW}} = 1.48 \left(\frac{G^3}{c^5} \frac{q}{(1+q)^2} \frac{m^4}{\dot{m}} \right)^{1/4}. \quad (10)$$

2.6. Threshold for Gas-driven Evolution

In the presented theoretical framework, ignoring for now the mass change of the binary in time, we can define the region of the parameter space in which the gas-driven evolution is important. In particular, gas can play a nonnegligible role if, at

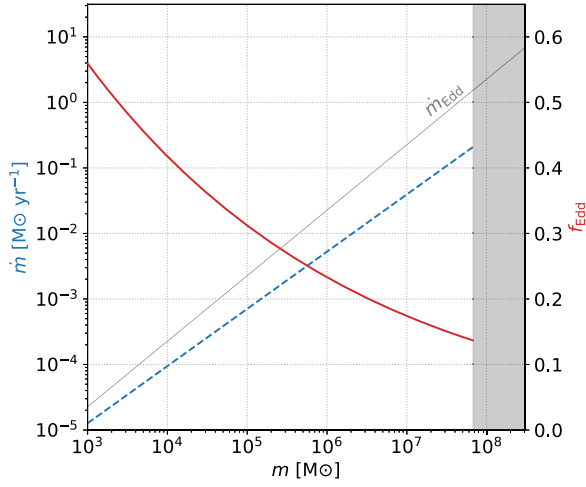


Figure 2. Minimum mass accretion rate above which gas-driven evolution becomes nonnegligible and the binary expansion can occur as a function of the total binary mass m . The blue dashed curve refers to the left-hand vertical axis and considers a fixed binary mass accretion rate (Equation (12)). For reference, we also show the Eddington accretion rate \dot{m}_{Edd} with a solid gray line. The red line refers to the right-hand vertical axis and expresses the accretion in terms of the Eddington ratio (Equation (13)). The curves are limited by a gray vertical shaded region (Equation (14)), which defines the masses above which the binary evolution can no longer be dominated by a non-self-gravitating disk. All curves assume $q = 1$.

$a_{\star \rightarrow \text{GW}}, |\dot{a}_{\text{gas}}| > |\dot{a}_{\star}| = |\dot{a}_{\text{GW}}|$. This condition translates into

$$\frac{\dot{m}}{m} > 0.621 \frac{G^{3/5}}{c} \left(\frac{q}{(1+q)^2} \right)^{1/5} \left(\frac{H G \rho}{\sigma} \right)^{4/5} m^{3/5}. \quad (11)$$

If we assume a fixed gas accretion rate \dot{m} , the gas dominates for accretion rates

$$\frac{\dot{m}}{M_{\odot} \text{ yr}^{-1}} > 5.41 \times 10^{-3} \left(\frac{H}{15} \right)^{4/5} \times \left(\frac{4q}{(1+q)^2} \right)^{1/5} \left(\frac{m}{10^6 M_{\odot}} \right)^{0.87}. \quad (12)$$

Instead, if we express the mass accretion rate as a function of the Eddington accretion rate \dot{m}_{Edd} , we obtain a lower limit for the Eddington ratio:

$$f_{\text{Edd}} > 0.24 \left(\frac{H}{15} \right)^{4/5} \left(\frac{4q}{(1+q)^2} \right)^{1/5} \left(\frac{m}{10^6 M_{\odot}} \right)^{-0.13}. \quad (13)$$

Note however that this only applies as long as the disk self-gravity can be neglected. This translates into the additional condition that the self-gravitating radius should be no smaller than the scale of the transition $a_{\star \rightarrow \text{GW}}$, which for the limiting values of accretion in Equations (12) and (13) implies

$$m < 6.72 \times 10^7 M_{\odot} \left(\frac{C}{0.5} \right)^{1.14} \left(\frac{H}{15} \right)^{-0.22} \left(\frac{4q}{(1+q)^2} \right)^{-0.34}. \quad (14)$$

Above this critical mass, the disk self-gravity is important at all the scales where the gaseous disk induces binary evolution, so the modeling presented in Equation (3) no longer holds and in the present model we just assume the gas to stop playing a role. The boundaries of the parameter space relevant for gas-driven evolution are shown in Figure 2.

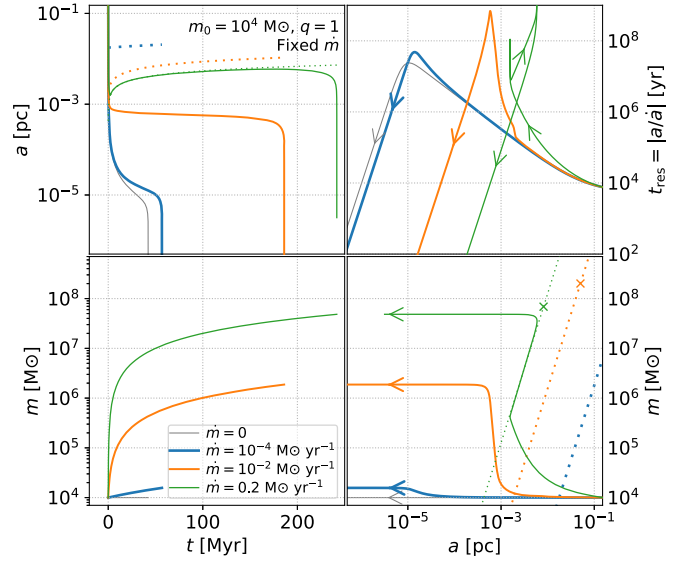


Figure 3. Evolution of a binary resulting from the interaction with gas, stars, and from the emission of GWs. The binary has an initial mass of $10^4 M_{\odot}$, and $q = 1$ throughout the evolution. The plots here show three cases in which the binary is assumed to have a fixed mass accretion rate, whose values are shown in the legend. The left panels show the time evolution of the binary semimajor axis (top) and total mass (bottom). The right panels show, as a function of the binary semimajor axis, the modulus of the residence time $|a/\dot{a}|$ (top) and the binary mass (bottom). In the top-left and bottom-right panels, the dotted lines show the self-gravitating radius as a function of time and mass, respectively, for the assumed mass accretion rates. In the right panels, the arrows show the direction of the evolution; in particular, in the $|a/\dot{a}|$ vs. a plot, while cases with $\dot{m} < 0.2 M_{\odot} \text{ yr}^{-1}$ should be read from right to left as the \dot{a} is always negative and the binary never expands its separation, in the case with $\dot{m} = 0.2 M_{\odot} \text{ yr}^{-1}$ the binary initially shrinks, then it undergoes a phase of expansion, and finally it starts shrinking again. The crosses in the bottom-right panel mark the position of m_{max} (Equation (16)).

2.7. Differential Equations for the Binary Evolution in Time

Below we present the time evolution of the binary semimajor axis induced by stars, gas, and GWs. In particular, we solve a set of differential equations for $\dot{a} = \dot{a}_{\star} + \dot{a}_{\text{gas}} + \dot{a}_{\text{GW}}$ (Equations (1), (3), and (7)) and \dot{m} . The mass accretion rate would be either expressed as a fixed value in time, or by fixing the Eddington ratio f_{Edd} . In order to avoid numerical issues, we smooth out the binary gas-driven evolution at CR_{sg} :

$$\dot{a}_{\text{gas,step}} = \dot{a}_{\text{gas}} (1 + \tanh[A(1 - a/(CR_{\text{sg}}))]), \quad (15)$$

where A is a constant that we set equal to 20. In the present framework, we always assume that the binary can interact with an infinite gas reservoir, so that we maximize the possible effects of gas-driven expansion. We start our integration at $a_0 = 10a_h$ (Section 2.1) and we stop it when the binary is well within the GW-dominated phase. Here we also neglect the binary eccentricity evolution, and we assume an always circular binary.

3. Results

Figures 3 and 4 show different aspects of the binary evolution for an equal-mass binary with initial mass of $10^4 M_{\odot}$ and different mass accretion rates.⁵ In particular,

⁵ The chosen initial mass is relatively small, but a lower initial mass makes the effects of gas-driven expansion more extreme for a given f_{Edd} or \dot{m} , as the binary can grow for a longer time before its mass gets large enough so that the surrounding disk becomes self-gravitating (see, e.g., Figure 5).

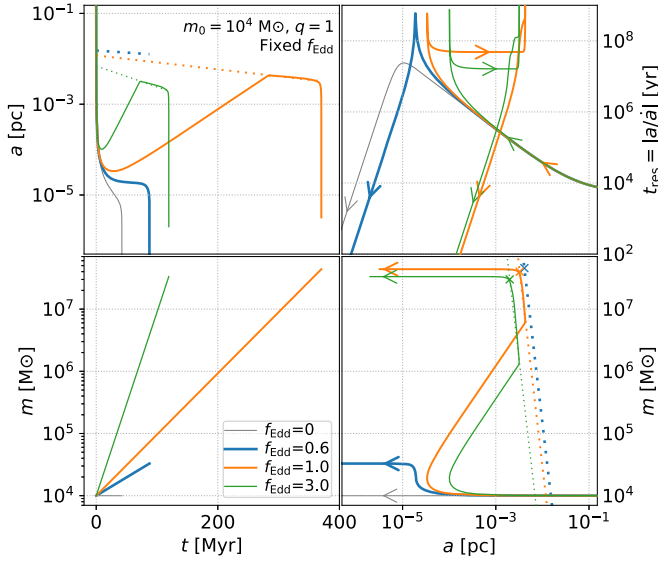


Figure 4. Same as Figure 3, but for a binary whose Eddington ratio is fixed along the evolution, as detailed in the legend. The crosses in the bottom-right panel refer to the mass limit in Equation (17).

Figure 3 assumes the binary to have a fixed \dot{m} along the evolution, while Figure 4 assumes a fixed Eddington ratio. The plots show that, if the mass accretion rate is small enough, the binary semimajor axis decreases as a result of stellar and then GW hardening, without the gas playing a significant role.

The impact of gas becomes stronger as the binary accretes mass more efficiently (i.e., \dot{m} or f_{Edd} get larger). While the initial evolution is always driven by stellar interactions that shrink the binary ($\dot{a} < 0$), the gas-driven expansion may counterbalance the stellar hardening so that the binary shrinking turns into an expansion, delaying GW-driven hardening (green line in Figure 3; green and orange lines in Figure 4).

Note that the reason for the expansion (rather than a stalling) is different in the two accretion scenarios. When the binary accretes at a fixed f_{Edd} , \dot{a}_{gas} only depends on a , while $\dot{a}_* \propto m^{-0.91}$. Therefore, the stellar hardening rate becomes less efficient as the binary increases its total mass and, overall, $\dot{a} > 0$. The binary expansion continues until the binary separation reaches the threshold $a = CR_{\text{sg}}$ where the gas-driven expansion is hampered. The binary semimajor axis then moves along CR_{sg} , which decreases as the binary increases its mass ($R_{\text{sg}} \propto m^{-7/45}$ for fixed f_{Edd}). The accretion continues until the binary mass grows large enough so that $|\dot{a}_* + \dot{a}_{\text{GW}}| > |\dot{a}_{\text{gas}}|$ and the binary starts shrinking very efficiently down to its coalescence.

In the scenario where the binary accretes mass at a fixed \dot{m} , the binary expands only if \dot{a}_{gas} is large enough so that, once a reaches CR_{sg} from larger separations and gas-driven evolution switches on, expansion and shrinking find an equilibrium point at $a = CR_{\text{sg}} \propto m^{38/45}$; this dependence on the mass implies that the binary expands owing to its mass growth, since it is bound to move along CR_{sg} . The binary expansion proceeds until the binary mass becomes large enough for GWs to take over, ensuring coalescence. It is worth stressing that, in the fixed \dot{m} scenario, expansion and significant mass growth only occur for highly super-Eddington initial accretion rates (e.g., in Figure 3, the green curve showing expansion initially assumes $\dot{m} \sim 10^3 \dot{m}_{\text{Edd}}$) for an initial binary mass of $10^4 M_{\odot}$.

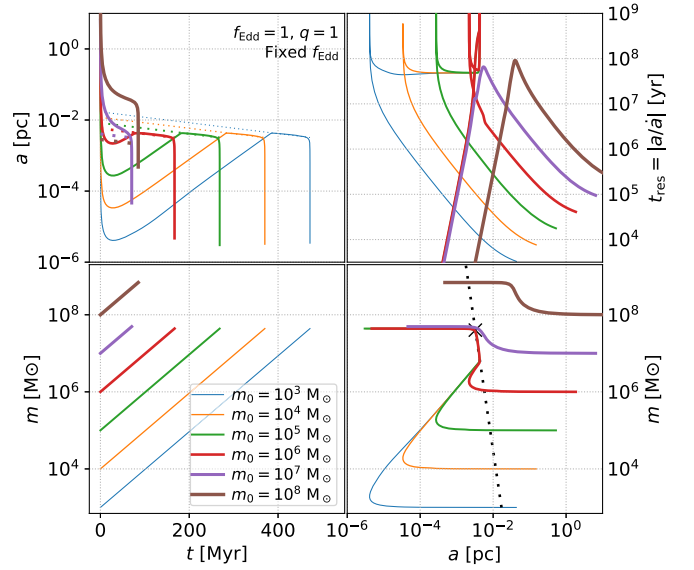


Figure 5. Same as Figure 3, but for a binary accreting always at the Eddington limit (fixed $f_{\text{Edd}} = 1$), varying its initial mass, as detailed in the legend. The cross in the bottom-right panel refers to the mass limit in Equation (17).

The picture outlined above clearly implies that, *although the binary may undergo a phase of expansion, this necessarily reverts into shrinking when m becomes large enough*, so that the binary will always enter the GW emission phase and reach its final coalescence.

Given the picture outlined above, we can compute the maximum mass that can be reached by the binary before GWs counteract the effect of gas and ensure a prompt shrinking; note that this mass is an upper limit that may be attained only if there exists a moment in time at which $\dot{a} > 0$.⁶ We compare R_{sg} with $a_{\text{gas}} \rightarrow \text{GW}$ and we obtain, for a fixed \dot{m}

$$\frac{m_{\text{max}}}{3.58 \times 10^8 M_{\odot}} = \left(\frac{C}{0.5}\right)^{1.5} \left(\frac{\dot{m}}{M_{\odot} \text{yr}^{-1}}\right)^{-0.36} \times \left(\frac{\alpha}{0.1}\right)^{0.93} \left(\frac{4q}{(1+q)^2}\right)^{-0.375}; \quad (16)$$

if instead we fix f_{Edd} , we find

$$\frac{m_{\text{max}}}{4.00 \times 10^7 M_{\odot}} = \left(\frac{C}{0.5}\right)^{1.1} \left(\frac{f_{\text{Edd}}}{1}\right)^{-0.26} \times \left(\frac{\alpha}{0.1}\right)^{0.69} \left(\frac{\eta}{0.1}\right)^{0.54} \left(\frac{4q}{(1+q)^2}\right)^{-0.28}. \quad (17)$$

This maximum mass decreases with f_{Edd} or \dot{m} , implying that the total mass accreted along the inspiral as a function of f_{Edd} or \dot{m} peaks near the smallest f_{Edd} or \dot{m} for which $\dot{a} > 0$ is attained.⁷

⁶ This threshold mass, once reached, is very close to the mass at the merger, as the binary shrinks very quickly in the GW regime and is not able to accrete a significant amount of gas.

⁷ The only caveat to this consideration is that the maximum mass has been derived equating CR_{sg} to $a_{\text{gas}} \rightarrow \text{GW}$, in fact assuming that gas-driven expansion is counterbalanced by GW emission only. Stars however may also give a nonnegligible contribution to counterbalance the gaseous evolution, so that the maximum mass can in fact be slightly smaller than the one reported here.

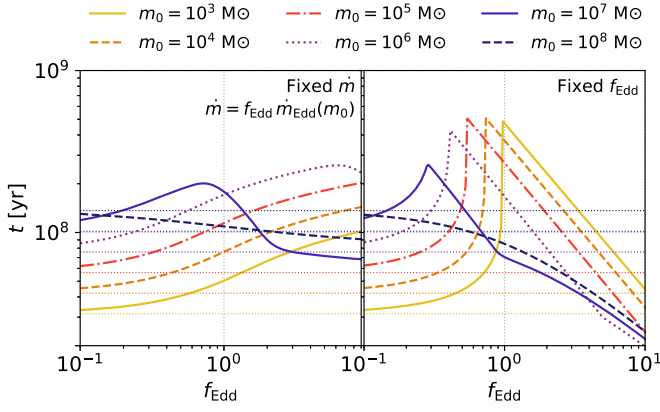


Figure 6. The plot shows the timescale needed by a binary to shrink from $10a_h$ to the final coalescence, as a function of the Eddington factor, for different binary masses (as coded in the legend). In the left panel, the binary mass accretion \dot{m} is assumed to be fixed in time, and equal to f_{Edd} times the initial Eddington accretion rate (\dot{m}_{Edd} evaluated at m_0). In the right panel we instead assume a fixed Eddington ratio throughout the evolution. The thin dotted horizontal lines mark the inspiral time for nonaccreting binaries ($\dot{m} = 0$); they refer to different m_0 , as in the legend. In the right panel, beyond the peak and before the slope change, the inspiral time is well described via Equation (18), as better detailed in the text.

Figure 5 shows the evolution of binaries with different initial mass, assuming a fixed $f_{\text{Edd}} = 1$. For this choice of the Eddington ratio, all binaries with mass $\lesssim m_{\text{max}}$ (evaluated when they cross CR_{sg} the first time) undergo an expansion phase; all their tracks in the (a, m) space overlap during the evolution, so that the final phase is the same. Only the most massive binaries (see the brown solid line in Figure 5) do not undergo expansion and may end up with a final mass $\gg m_{\text{max}}$.

It is worth stressing that, in our model, we assumed the binary to continue growing its mass even for $a > CR_{\text{sg}}$. We also performed a series of tests damping mass accretion together with binary expansion at CR_{sg} ; this can mimic inefficient accretion beyond this radius, or a gaseous disk that is spatially limited to CR_{sg} (or exhausted once this separation is reached). In this situation, we found the overall evolutionary picture depicted here to remain similar and the time needed to reach coalescence to be $\lesssim 10^8$ yr, but the specific evolutionary tracks may vary.⁸

Figure 6 shows the timescale needed by binaries with different initial mass to reach their final coalescence as a function of the Eddington ratio, for a fixed mass accretion rate (left) or Eddington ratio (right). Remarkably, *the inspiral timescale does not get dramatically longer as a result of the gas-driven expansion*: if we fix \dot{m} and consider reasonable values for the mass accretion rate, the inspiral timescale increases by a factor of a few (< 10) at most. The impact of gas is more relevant when adopting a fixed f_{Edd} . In that case, the inspiral time peaks near $f_{\text{Edd}} = 1$, and specifically very close to the smallest f_{Edd} allowing for $\dot{a} > 0$ to be attained along the evolution. The inspiral time declines for large f_{Edd} as the binary efficient mass growth implies that all processes occur faster. Also note that, in proportion, the inspiral time is less impacted by gas-driven expansion as the binary initial mass is larger:

⁸ If mass accretion is damped together with expansion, the solution of the differential equations proceeds so that, in case of expansion, the binary stalls at a radius in which expansion is nearly hampered, but still some mass accretion can occur. Only if accretion is stopped before the binary reaches CR_{sg} may the binary endlessly stall owing to the lack of mass growth; however, we believe this latter configuration to be rather unphysical.

when the initial mass is closer to the limiting mass in Equations (16) or (17), the binary needs a shorter time to get there and enter the GW-driven inspiral.

Clearly, the specific evolutionary timescales obtained here depend both on the choice of C , i.e., the multiple of R_{sg} at which gas expansion is damped, and on how sharp the suppression is at such a point. Nevertheless, we checked that the global scenario depicted here is robust against variations of such quantities.

Finally, we can also estimate the inspiral time of binaries if expansion ever occurs. A fixed f_{Edd} implies that $dm/m = f_{\text{Edd}} [4\pi G m_p / (\eta \sigma_T c)] dt$, so that the mass grows exponentially, and the time needed to reach m_{max} from m_0 (computed via Equation (17)) naturally follows from the Salpeter timescale:

$$t_{\text{exp}} = \log\left(\frac{m_{\text{max}}}{m_0}\right) \frac{1}{f_{\text{Edd}}} \frac{\eta \sigma_T c}{4\pi G m_p} = 4.415 \times 10^7 \left(\frac{f_{\text{Edd}}}{1}\right)^{-1} \log\left(\frac{m_{\text{max}}}{m_0}\right) \text{yr}. \quad (18)$$

The value of t_{exp} is in very good agreement with our results, and nearly perfectly overlaps with the curves in the right panel of Figure 6 beyond the peak (except for values of f_{Edd} so large that $m > m_{\text{max}}$ as CR_{sg} is first reached). In principle, one could make an analogous estimate for the cases with fixed \dot{m} , but this would be of little value as expansion is almost never attained if \dot{m} is fixed.

4. Discussion and Conclusions

In this study we addressed the evolution of MBHBs under the concurrent effect of different mechanisms, specifically the gas-, stellar-, and GW-driven evolution. In particular, we explored the consequences of an outspiral phase mediated by the MBHB interaction with a gaseous circumbinary disk, based on the results of recent numerical simulations (Duffell et al. 2020; Muñoz et al. 2020; Tiede et al. 2020). We described the MBHB evolution through a coupled set of simplified differential equations expressing the time variation of the semimajor axis and the total mass. Specifically, stars, gas, and GWs all contribute to the semimajor axis evolution, while the mass ratio and total mass growth depend on gas dynamics only.

The key result of the present Letter is that a putative phase of gas-driven expansion does not sensibly impact the MBHB evolution and inspiral time. Binary expansion can occur only if the binary mass accretion rate remains close or exceeds the Eddington limit along the entire evolution. In these cases, the binary may undergo an outspiral phase. However, it cannot expand to scales larger than the self-gravitating radius (i.e., the radius at which disk self-gravity cannot be further ignored, and gas-driven expansion can no longer occur; Franchini et al. 2021). Following this possible expansion phase, the semimajor axis nearly stalls at the self-gravitating radius, while the binary continues accreting, until its mass gets large enough for GWs to induce coalescence. Irrespective of the presence of an outspiral phase, we always find that MBHBs reach the coalescence within at most a few hundreds of megayears. Thus, although binaries interacting with gas can undergo an expansion phase, our findings suggest that such a phase does not hinder MBHBs coalescence. Noticeably, a putative expansion would lengthen the phase at which the binary and the disk are coupled at

subparsec scales (see, e.g., Figure 5). Therefore, according to the presented scenario, the luminous subparsec binaries targeted by present and future time-domain surveys could be more common than previously thought.

Given the simplifications made in the current implementation, our treatment is subjected to a number of caveats that we now discuss. We would like however to stress that all our assumptions were in general conservative.

We assumed, for simplicity, that once the binary enters the mass regime where the disk self-gravity is important, the gas-driven expansion is halted. The interaction of a binary with a self-gravitating accretion disk has been studied through hydrodynamical simulations by, e.g., Cuadra et al. (2009), Roedig et al. (2012), and Franchini et al. (2021). These works focused on self-regulated disks, i.e., for which the cooling is not efficient enough for the disk to fragment into bound clumps, and found that the interaction with the binary leads to its shrinking in the wide region of the parameter space explored. If the disk instead does fragment into clumps, possibly forming stars, the interaction with the binary would occur by means of stellar slingshot ejections, which induce the shrinking (see Section 2.1). Therefore, the binary semimajor axis decreases with time in both scenarios, making our choice to suppress the evolution quite conservative. Furthermore, if the structure of the circumbinary disk surrounding an MBHB is similar to the structure of accretion disks in active galactic nuclei (Collin-Souffrin & Dumont 1990), then the disk aspect ratio is expected to be $H/R \approx 0.01-10^{-3}$. In this regime of disk aspect ratios, the binary is found to shrink due to the interaction with the disk even in the non-self-gravitating regime. Moreover, it has been shown that binaries with initial values for $q < 0.05$ or $e > 0.1$ tend to a solution that guarantees the inspiral, regardless of the exact properties of the circumbinary disk, suggesting that even a mild initial eccentricity could avert the expansion (D’Orazio & Duffell 2021).

Our modelization of the stellar hardening (Section 2.1, Equation (1)), was based on well-established scaling relations linking the density and velocity dispersion with the MBHB total mass (see, e.g., Sesana & Khan 2015). This provides an “average” connection between MBHBs and the characteristics of their hosts. This is of course a simplification, but it is conservative in many ways. First of all, the employed scaling relations do not consider the possible presence of a nuclear star cluster in the galaxy center, which would enhance the hardening by increasing the density at the binary influence radius. Therefore the evolutionary timescale found with our implementation should be considered as a conservative estimate of the stellar-driven shrinking, which in many galaxies may be much more efficient. In addition we updated ρ , σ that enter the stellar hardening prescription (Equation (1)) with the binary total mass at any given time; this implies a symbiotic evolution of the binary and the galaxy, although the galaxy properties are likely to evolve on a much longer timescale compared with the binary evolution. Note that this assumption is also conservative, as $\dot{a}_* \propto \rho/\sigma \propto m^{-0.91}$, implying that we would get a more efficient stellar hardening if we were to keep fixed the galaxy properties or to make them evolve more slowly as the binary mass grows. In particular, this assumption is what drives the expansion phase when f_{Edd} is fixed, as described above, and a less conservative assumption on \dot{a}_* may result in a binary stalling phase rather than an expansion phase (which would anyways end with coalescence due to the relentless mass



accretion). The important caveat here is that the stellar content of galaxies in the very high- z universe may have been relatively low, so that the employed scaling relations may not be suited for the early universe, in which much of the gas content still had to be turned into stars. However, the structural parameters of high- z galaxies near their center, their gas and stellar content, and morphological properties are poorly known, and the investigation of this aspect is beyond the scope of the present work.

Another assumption we made in this work concerns the binary eccentricity, which we force to remain zero throughout the binary evolution. This is again a conservative assumption, as the binary eccentricity may be nonzero at the binary formation time and may grow during the stellar hardening phase, especially if the stellar background is isotropic or counterrotates with the binary. Furthermore, the interaction of an initially circular or slightly eccentric binary with a gaseous circumbinary disk is also found to impact the binary eccentricity in both the non-self-gravitating (Ragusa et al. 2020) and the self-gravitating (Roedig et al. 2011) cases. A nonzero eccentricity can impact both stellar- and GW-driven hardening. The effect on the former is generally a mild enhancement of the hardening efficiency by stars, while for the latter it provides a strong boost in the emitted GW power, which strongly increases as binaries become eccentric, determining that a GW-driven inspiral becomes much more efficient compared with the circular cases.

Finally, the existence of a binary expansion phase implies a significant mass growth prior to merger. Although this might shift a significant number of low-mass MBHBs at the margin of the LISA sensitivity band (LISA will be mostly sensitive to MBHBs of $M < 10^7 M_\odot$), it is important to bear in mind that we assumed an unlimited supply of gas. In reality, most of the cold gas funneled to the center of a merger remnant is converted in stars and a circumbinary disk cannot be sustained indefinitely. A putative stalling or expansion phase may in practice be cut off by the consumption of the available cold gas reservoir. To fully assess the putative effect of binary outspiral on the population of MBHBs observable by LISA and PTA and on the statistics of MBHBs possibly observable as electromagnetic periodic sources, our semianalytic model should be coupled to a framework describing the evolution of MBHBs and their hosts along the cosmic history (Izquierdo-Villalba et al. 2020), in analogy with what was done by, e.g., Kelley et al. (2017, who however did not account for gas-driven expansion) in the context of cosmological simulations. This is deferred to future work.

We warmly thank the anonymous referee for useful comments and suggestions, and Massimo Dotti for fruitful discussion. E.B., A.F., and A.S. acknowledge financial support provided under the European Unions H2020 ERC Consolidator Grant “Binary Massive black hole Astrophysics” (B Massive, Grant Agreement: 818691). M.B. acknowledges funding from MIUR under the grant PRIN 2017-MB8AEZ.

ORCID iDs

Elisa Bortolas  <https://orcid.org/0000-0001-9458-821X>
 Alessia Franchini  <https://orcid.org/0000-0002-8400-0969>
 Matteo Bonetti  <https://orcid.org/0000-0001-7889-6810>
 Alberto Sesana  <https://orcid.org/0000-0003-4961-1606>

References

- Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021, *ApJS*, **252**, 5
- Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv:1702.00786
- Armitage, P. J., & Natarajan, P. 2002, *ApJL*, **567**, L9
- Artymowicz, P., & Lubow, S. H. 1994, *ApJ*, **421**, 651
- Barack, L., Cardoso, V., Nissanke, S., et al. 2019, *CQGra*, **36**, 143001
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, **287**, 307
- Bonetti, M., Bortolas, E., Lupi, A., & Dotti, M. 2021, *MNRAS*, **502**, 3554
- Bonetti, M., Bortolas, E., Lupi, A., Dotti, M., & Raimundo, S. I. 2020, *MNRAS*, **494**, 3053
- Bortolas, E., Capelo, P. R., Zana, T., et al. 2020, *MNRAS*, **498**, 3601
- Bortolas, E., Gualandris, A., Dotti, M., & Read, J. I. 2018a, *MNRAS*, **477**, 2310
- Bortolas, E., Gualandris, A., Dotti, M., Spera, M., & Mapelli, M. 2016, *MNRAS*, **461**, 1023
- Bortolas, E., Mapelli, M., & Spera, M. 2018b, *MNRAS*, **474**, 1054
- Capelo, P. R., & Dotti, M. 2017, *MNRAS*, **465**, 2643
- Chandrasekhar, S. 1943, *ApJ*, **97**, 255
- Collin-Souffrin, S., & Dumont, A. M. 1990, *A&A*, **229**, 292
- Cuadra, J., Armitage, P. J., Alexander, R. D., & Begelman, M. C. 2009, *MNRAS*, **393**, 1423
- Desvignes, G., Caballero, R. N., Lentati, L., et al. 2016, *MNRAS*, **458**, 3341
- D’Orazio, D. J., & Duffell, P. C. 2021, *ApJL*, **914**, L21
- Dotti, M., Colpi, M., & Haardt, F. 2006, *MNRAS*, **367**, 103
- Duffell, P. C., D’Orazio, D., Derdzinski, A., et al. 2020, *ApJ*, **901**, 25
- Escala, A., Larson, R. B., Coppi, P. S., & Mardones, D. 2004, *ApJ*, **607**, 765
- Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, *ApJ*, **783**, 134
- Franchini, A., Sesana, A., & Dotti, M. 2021, *MNRAS*, in press (doi:10.1093/mnras/stab2234)
- Gualandris, A., Read, J. I., Dehnen, W., & Bortolas, E. 2017, *MNRAS*, **464**, 2301
- Haiman, Z., Kocsis, B., & Menou, K. 2009, *ApJ*, **700**, 1952
- Heath, R. M., & Nixon, C. J. 2020, *A&A*, **641**, A64
- Izquierdo-Villalba, D., Bonoli, S., Dotti, M., et al. 2020, *MNRAS*, **495**, 4681
- Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2017, *MNRAS*, **471**, 4508
- Kelley, L. Z., Haiman, Z., Sesana, A., & Hernquist, L. 2019, *MNRAS*, **485**, 1579
- Khan, F. M., Just, A., & Merritt, D. 2011, *ApJ*, **732**, 89
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Lodato, G., Nayakshin, S., King, A. R., & Pringle, J. E. 2009, *MNRAS*, **398**, 1392
- Mayer, L., Kazantzidis, S., Madau, P., et al. 2007, *Sci*, **316**, 1874
- Merritt, D., Schnittman, J. D., & Komossa, S. 2009, *ApJ*, **699**, 1690
- Moody, M. S. L., Shi, J.-M., & Stone, J. M. 2019, *ApJ*, **875**, 66
- Muñoz, D. J., Lai, D., Kratter, K., & Miranda, R. 2020, *ApJ*, **889**, 114
- Muñoz, D. J., Miranda, R., & Lai, D. 2019, *ApJ*, **871**, 84
- Perego, A., Dotti, M., Colpi, M., & Volonteri, M. 2009, *MNRAS*, **399**, 2249
- Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, *MNRAS*, **490**, 4666
- Peters, P. C. 1964, *PhRv*, **136**, 1224
- Pfister, H., Volonteri, M., Dubois, Y., Dotti, M., & Colpi, M. 2019, *MNRAS*, **486**, 101
- Quinlan, G. D. 1996, *NewA*, **1**, 35
- Ragusa, E., Alexander, R., Calcino, J., Hirsh, K., & Price, D. J. 2020, *MNRAS*, **499**, 3362
- Reardon, D. J., Hobbs, G., Coles, W., et al. 2016, *MNRAS*, **455**, 1751
- Roedig, C., Dotti, M., Sesana, A., Cuadra, J., & Colpi, M. 2011, *MNRAS*, **415**, 3033
- Roedig, C., Sesana, A., Dotti, M., et al. 2012, *A&A*, **545**, A127
- Schödel, R., Gallego-Cano, E., Nogueras Lara, F., Dong, H., & Gallego Calvente, T. 2017, in Proc. XII Scientific Meeting of the Spanish Astronomical Society, Highlights on Spanish Astrophysics IX, ed. S. Arribas et al.
- Sesana, A., Haardt, F., & Madau, P. 2006, *ApJ*, **651**, 392
- Sesana, A., & Khan, F. M. 2015, *MNRAS*, **454**, L66
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, **24**, 337
- Tiede, C., Zrake, J., MacFadyen, A., & Haiman, Z. 2020, *ApJ*, **900**, 43
- Varisco, L., Bortolas, E., Dotti, M., & Sesana, A. 2021, arXiv:2104.14570
- Vasiliev, E., Antonini, F., & Merritt, D. 2014, *ApJ*, **785**, 163