A&A, 685, A51 (2024)
https://doi.org/10.1051/0004-6361/202348509
© The Authors 2024

Astronomy & Astrophysics

Impact of gas hardening on the population properties of hierarchical black hole mergers in active galactic nucleus disks⋆

Maria Paola Vaccaro[1,2], Michela Mapelli[1,2,3,4], Carole Périgois[2,3], Dario Barone[2], Maria Celeste Artale[5], Marco Dall'Amico[2,3], Giuliano Iorio[2,3,4], and Stefano Torniamenti[1,2,3,4]

[1] Institut für Theoretische Astrophysik, ZAH, Universität Heidelberg, Albert-Ueberle-Straße 2, 69120 Heidelberg, Germany
e-mail: mariapaolavaccaro@gmail.com; mapelli@uni-heidelberg.de
[2] Physics and Astronomy Department Galileo Galilei, University of Padova, Vicolo dell' Osservatorio 3, 35122 Padova, Italy
[3] INFN – Padova, Via Marzolo 8, 35131 Padova, Italy
[4] INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, 35122 Padova, Italy
[5] Departamento de Ciencias Fisicas, Universidad Andres Bello, Fernandez Concha 700, Las Condes, Santiago, Chile

Received 6 November 2023 / Accepted 31 January 2024

ABSTRACT

Hierarchical black hole (BH) mergers in active galactic nuclei (AGNs) are unique among formation channels of binary black holes (BBHs) because they are likely associated with electromagnetic counterparts and can efficiently lead to the mass growth of BHs. Here, we explore the impact of gas accretion and migration traps on the evolution of BBHs in AGNs. We have developed a new fast semi-analytic model, that allows us to explore the parameter space while capturing the main physical processes involved. We find that an effective exchange of energy and angular momentum between the BBH and the surrounding gas (i.e., gas hardening) during inspiral greatly enhances the efficiency of hierarchical mergers, leading to the formation of intermediate-mass BHs (up to $10^4\ M_\odot$) and triggering spin alignment. Moreover, our models with efficient gas hardening show both an anticorrelation between the BBH mass ratio and the effective spin and a correlation between the primary BH mass and the effective spin. In contrast, if gas hardening is inefficient, the hierarchical merger chain is already truncated after the first two or three generations. We compare the BBH population in AGNs with other dynamical channels as well as isolated binary evolution.

Key words. black hole physics – gravitation – gravitational waves – stars: black holes – stars: kinematics and dynamics – galaxies: active – galaxies: nuclei

1. Introduction

The first direct detection of gravitational waves (GWs) in 2015 (Abbott et al. 2016) paved the ground for the study of binary black holes (BBHs). More than 90 GW event candidates have been detected to date, most of them associated with BBHs (Abbott et al. 2023a,b). A few BBH candidates, such as GW190521 (Abbott et al. 2020) and possibly GW190403_051519 and GW190426_190642 (Abbott et al. 2021, 2023a,b), stand out because they involve black holes (BHs) in the pair-instability mass gap, challenging traditional models of stellar evolution (Woosley et al. 2002; Woosley & Heger 2014, 2021; Belczynski et al. 2016; Spera & Mapelli 2017; Stevenson et al. 2019; O'Brien 2021; Siegel et al. 2022; Sabhahit et al. 2023; Umeda & Nagele 2024) and raising questions about their formation (Farmer et al. 2019, 2020; Mapelli et al. 2020; Belczynski 2020; Marchant & Moriya 2020; Costa et al. 2021; Farrell et al. 2021; Vink et al. 2021; Tanikawa et al. 2021, 2022; Dall'Amico et al. 2021; Banerjee 2022; Méndez et al. 2023).

Stellar dynamics provides some of the most straightforward channels to explain the birth of such oversized BHs, via star–star collisions (Di Carlo et al. 2019, 2020; Kremer et al. 2020; Renzo et al. 2020; Torniamenti et al. 2022; Costa et al. 2022; Ballone et al. 2023) or repeated mergers of stellar-origin BHs (Miller & Hamilton 2002; Fishbach et al. 2017; Gerosa & Berti 2017; Rodriguez et al. 2019; Doctor et al. 2020; Kimball et al. 2020; Flitter et al. 2021; see, e.g., Mapelli 2021 for a review). The latter process, often called hierarchical mergers, takes place only in dense star clusters, where merger remnants can be retained inside the system (e.g., Antonini et al. 2019; Fragione & Silk 2020) and pair up again with other single BHs via dynamical encounters (e.g., Heggie 1975; Portegies Zwart & McMillan 2000). In order to constrain the origin of the observed BBH mergers, it is important to characterize the hierarchical merger process in different environments, such as young star clusters (YSCs, e.g., Ziosi et al. 2014; Mapelli 2016; Banerjee 2017a,b, 2020; Di Carlo et al. 2020; Kumamoto et al. 2019, 2020), globular clusters (GCs, e.g., Downing et al. 2010; Rodriguez et al. 2015, 2016, 2018; Askar et al. 2017; Fragione & Kocsis 2018; Zevin et al. 2019; Antonini & Gieles 2020; Antonini et al. 2023), nuclear star clusters (NSCs, e.g., O'Leary et al. 2009; Miller & Lauburg 2009; Antonini & Rasio 2016; Petrovich & Antonini 2017; Leigh et al. 2017; Arca Sedda & Gualandris 2018; Arca Sedda 2020; Atallah et al. 2023; Chattopadhyay et al. 2023), and active galactic nuclei (AGNs, e.g., McKernan et al. 2012, 2020, 2022; Bartos et al. 2017; Stone et al. 2017; Yang et al. 2019a; Secunda et al. 2020; Tagawa et al. 2020b,a, 2022, 2023; Ford & McKernan 2022).

In AGNs, a central supermassive black hole (SMBH) is surrounded by a dense gaseous accretion disk. AGNs can appear as

⋆ The data underlying this article will be shared upon reasonable request to the corresponding authors.

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

extremely luminous objects called quasars, but also as Seyfert galaxies, radio galaxies, or blazars, depending on their luminosity and our viewing angle (see Netzer 2015 for a review of the unification scheme and its controversy). Stars and stellar-sized BHs orbiting the SMBH are subject to gas torques, which can bend their orbits, aligning them with the disk. This is expected to lead to a large overdensity of BHs with similar orbits in small areas of the disk called migration traps (McKernan et al. 2012; Bellovary et al. 2016), where BHs can pair up efficiently via gas capture (DeLaurentiis et al. 2023; Li et al. 2023; Rowan et al. 2023, 2024; Whitehead et al. 2023). Therefore, AGNs are potential factories of BBH mergers with masses in the pair-instability mass gap and above (Yang et al. 2019a; Tagawa et al. 2020a,b).

This channel has attracted great interest because GW signals from mergers could be accompanied by some electromagnetic emission (Bartos et al. 2017; Tagawa et al. 2023), possibly anticipated by a neutrino detection (Zhu 2024; Zhou & Wang 2023). To explore this possibility, electromagnetic follow-up observations have been carried out for several GW events, but so far all associations with BBH mergers remain controversial (Greiner et al. 2016; Coughlin et al. 2020; Calderón Bustillo et al. 2021).

Several studies have investigated the possibility that the high-mass BBH merger event GW190521 is associated with an AGN disk (Tagawa et al. 2021; Samsing et al. 2022) because of its large mass, high spin (Abbott et al. 2020), and claimed electromagnetic counterpart (Graham et al. 2020; Calderón Bustillo et al. 2021; Morton et al. 2023). Yang et al. (2019a) explored the AGN scenario for another high-mass BBH, GW170729, for which there is evidence of nonzero effective and precessing spins. Some authors, such as Gayathri et al. (2021) and Ford & McKernan (2022), have predicted that a sizeable fraction of the observed BBH mergers (~20% up to 80%) may originate in AGNs. However, according to an analysis based on the sky localization of GW signals, the fraction of detected BBH mergers that originated in bright AGNs ($L \geq 10^{46}$ erg s^{-1}) cannot be higher than 17% (Veronesi et al. 2023).

Binary BH pair-up in AGNs can happen either in the migration traps or at other locations in the disk (Wang et al. 2021) collectively referred to as 'the bulk'. McKernan et al. (2020) find that, although more than 50% of mergers happen in the bulk, hierarchical mergers are only efficient in migration traps. Moreover, Tagawa et al. (2020b) find that BBHs assembled in the bulk via gas capture migrate toward the migration trap while hardening.

The complex physics of BBHs in AGN disks has been extensively explored in the literature, with studies accounting for the effects of binary–single interactions (Stone et al. 2017; Leigh et al. 2017) and gas torques (using models borrowed from protoplanetary physics; e.g., McKernan et al. 2012; Bartos et al. 2017; Yang et al. 2019a,b; Secunda et al. 2020; or hydrodynamical simulations; e.g., Li et al. 2023; Kaaz et al. 2023; Rowan et al. 2023, 2024; Whitehead et al. 2023). For example, Ishibashi & Gröbner (2020) explore the effects of energy and angular momentum exchange between the BBH and the surrounding gas. In particular, they assume that the BBH evolves surrounded by a circumbinary disk, which induces orbital decay and increases its orbital eccentricity (i.e., gas hardening).

We studied hierarchical BBH mergers in the migration traps of AGN disks and compared our results to mergers in YSCs, GCs, and NSCs (Mapelli et al. 2021, 2022). Specifically, we tested the impact of gas hardening (GH) on the BBH merger population in AGNs (Ishibashi & Gröbner 2020). We find that the hierarchical merger process is pronouncedly more efficient when accounting for gas hardening.

We have developed a new semi-analytic code for the simulation of hierarchical mergers in AGNs. It is effective at exploring the parameter space, and it models the relevant physical processes much faster than an N-body or hydrodynamical code. The new code is publicly available within the FASTCLUSTER software environment[1] (Mapelli et al. 2021, 2022).

2. Methods

2.1. AGN disk model

We assumed the AGN disk to be as described by Sirko & Goodman (2003, hereafter, SG). These authors introduce a hydro-dynamical model for a geometrically thin and optically thick disk with steady-state accretion of 0.1 $\dot{M}_{\rm Edd}$ onto the central SMBH. Gas turbulence is assumed to be the cause of disk viscosity, characterized by the viscosity coefficient $\alpha \in [0, 1]$. This model neglects any effects due to magnetic fields and general relativity. The physical parameters of the disk are functions of the mass of the SMBH $M_{\rm SMBH}$, the distance R from the SMBH, and the viscosity parameter α. We slightly simplified the radial dependence of the SG model for numerical ease and re-scaled their functions to allow for different $M_{\rm SMBH}$ and α parameters, as shown in Fig. 1. The resulting expressions for the gas surface density, $\Sigma_{\rm gas}$, the disk aspect ratio, $h = H/R$ (defined as the ratio between the height of the disk, $H(R)$, and its radius, R), and the sound speed, $c_{\rm s}$, are the following:

$$\Sigma_{\rm gas}(R, M_{\rm SMBH}, \alpha) = 7.94 \times 10^5 \ {\rm g\ cm} \left(\frac{M_{\rm SMBH}}{10^8\ M_\odot}\right)^{4/5} \left(\frac{\alpha}{0.01}\right) \begin{cases} \left(\frac{R}{10^{3.2}\ R_{\rm g}}\right)^{0.8} & R \leq 10^{3.2}\ R_{\rm g} \\ \left(\frac{R}{10^{3.2}\ R_{\rm g}}\right)^{-1.49} & R > 10^{3.2}\ R_{\rm g} \end{cases} \quad (1)$$

$$h(R, M_{\rm SMBH}, \alpha) = 7.59 \times 10^{-3} \left(\frac{M_{\rm SMBH}}{10^8\ M_\odot}\right)^{-3/20} \left(\frac{\alpha}{0.01}\right)^{-1} \begin{cases} \left(\frac{R}{10^{3.2}\ R_{\rm g}}\right)^{-0.6} & R \leq 10^{3.2}\ R_{\rm g} \\ \left(\frac{R}{10^{3.2}\ R_{\rm g}}\right)^{0.5} & R > 10^{3.2}\ R_{\rm g} \end{cases} \quad (2)$$

$$c_{\rm s}(R, M_{\rm SMBH}, \alpha) = 5.37 \times 10^6 \ {\rm km\ s^{-1}} \left(\frac{M_{\rm SMBH}}{10^8\ M_\odot}\right)^{3/2} \left(\frac{\alpha}{0.01}\right)^{-1} \begin{cases} \left(\frac{R}{10^{3.2}\ R_{\rm g}}\right)^{-1.1} & R \leq 10^{3.2}\ R_{\rm g} \\ 1 & R > 10^{3.2}\ R_{\rm g} \end{cases} . \quad (3)$$

In the above equations, we define the gravitational radius as $R_{\rm g} = GM_{\rm SMBH}/c^2$, where G is the gravity constant and c the speed of light.

The viscosity coefficient α is a free parameter. We assumed it to be constant over the whole extension of the disk and to have a constant value independent of the other physical properties of the disk. The assumed value is $\alpha = 0.1$ which is consistent with observations (King et al. 2007).

Migration traps are locations in the disk where migration stalls and BHs pile up. Bellovary et al. (2016) find that migration traps are located where the slope of the gas surface density profile changes sign from positive to negative (i.e. at local maxima).

[1] FASTCLUSTER is an open-source code available at https://gitlab.com/micmap/fastcluster_open

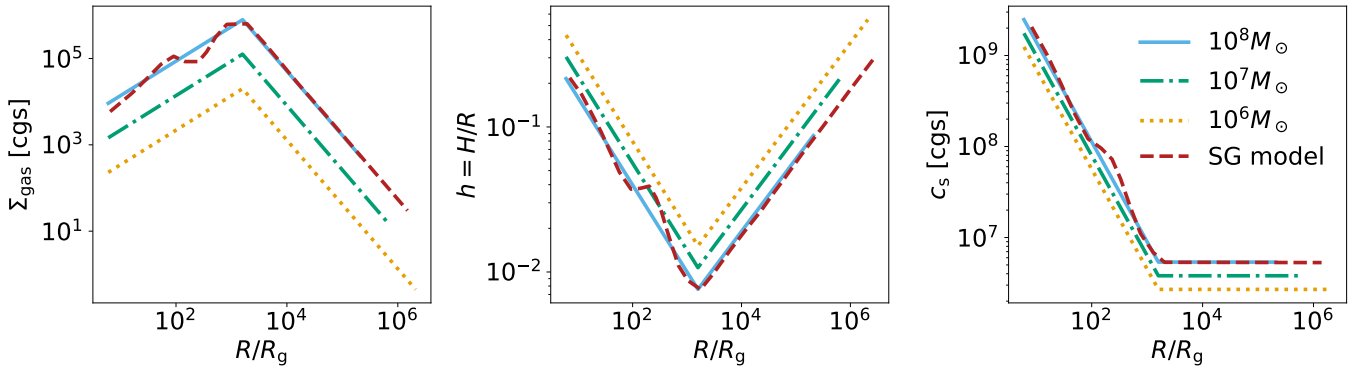

Fig. 1. Surface density, Σ_{gas}, aspect ratio, h and sound speed, c_s, profiles as a function of the scale distance, R/R_g. Dashed maroon lines represent the SG model for a disk with viscosity $\alpha = 0.01$ around a $10^8\,M_\odot$ SMBH. Solid blue lines show our broken power-law best fit to the SG model for $\alpha = 0.01$ and $M_{SMBH} = 10^8\,M_\odot$. The dash-dotted green and dotted orange lines show our fits for a $10^7\,M_\odot$ and a $10^6\,M_\odot$ SMBH, respectively (both with $\alpha = 0.01$).

In a SG disk there are two local maxima in the gas surface density and therefore two migration traps: an inner trap at $\approx 100\,R_g$ and an outer one at $\approx 1300\,R_g$.

We see in Fig. 1 that the inner migration trap coincides with a local maximum in the density profile of the original SG model, while the outer migration trap corresponds to a global maximum. For simplicity, we ignored the local overdensity in the disk at $10^2\,R_g$ and we only assumed the existence of the outer migration trap. We thus defined the trap radius as $R_{trap} = 10^{3.1}\,R_g$, although its location and existence may be affected by other effects (e.g. Grishin et al. 2024; Pan & Yang 2021 more details in Sect. 4.4).

The disk is assumed to have radial extension between the innermost stable circular orbit (ISCO) radius for a nonrotating BH, which we call $R_{min} = 6\,R_g$ in this context, and an outer radius $R_{max} = 0.1\,\mathrm{pc}(M_{SMBH}/10^6\,M_\odot)^{1/2}$, beyond which the disk's self-gravity becomes important (Goodman 2003; Yang et al. 2019b). For $R > R_{max}$, the disk is expected to fragment and experience star formation. Energetic feedback from the newly formed stars may keep the disk vertically supported (Sirko & Goodman 2003), but the outcome of viscous interactions with the BHs in this region of the disk is highly uncertain, so we conservatively neglected it. Moreover, we neglected the contribution of stars formed in AGN disks, which are expected to be massive and whose compact remnants might also become GW sources (Cantiello et al. 2021; Wang et al. 2023).

We assumed a non-spinning SMBH. We randomly generated the SMBH mass from the observational distribution derived by Greene & Ho (2007) in the Local Universe, which is a Gaussian distribution with mean $\langle \log M_{SMBH}/M_\odot \rangle = 6.576 \pm 0.591$.

A mass-accretion episode onto a SMBH lasts for a finite amount of time, which we refer to as the AGN disk lifetime. The lifetime of AGN disks is subject to large uncertainty: different estimates span several orders of magnitude in the range of 10^{-2}–10^3 Myr (Khrykin et al. 2021). Also, it is not clear whether accretion onto SMBHs happens continuously over a given time span, or episodically through many cycles of efficient accretion. Here, we used the estimate by Khrykin et al. (2021), based on observations of quasars' proximity effect. They find that the quasar lifetime τ is distributed according to a Gaussian distribution with mean $\langle \log \tau/\mathrm{Myr} \rangle = 0.22 \pm 0.80$. We randomly extracted the lifetime for AGN disks in our model from this distribution.

There are other models for stable AGN accretion disks, with notably different features compared to the one we adopt here. For example, the model by Thompson et al. (2005) differs from the SG model in both density and aspect ratio (McKernan et al. 2022). We will explore the impact of different disk models in a follow-up study.

2.2. Nuclear star cluster (NSC)

Supermassive black holes and NSCs commonly inhabit galactic spheroids with stellar masses ranging from $10^8\,M_\odot$ to $10^{11}\,M_\odot$ (Graham & Spitler 2009). For $M_{SMBH} \leq 5\times10^7\,M_\odot$, the mass of the SMBH and that of the NSC scale as

$$\log\left(\frac{M_{SMBH}}{M_{SMBH}+M_{NSC}}\right) = \frac{2}{3}\log\left(\frac{M_{SMBH}}{5\times10^7\,M_\odot}\right). \quad (4)$$

For $M_{SMBH} > 5\times10^7\,M_\odot$, the SMBH generally does not coexist with a NSC and we set $M_{NSC} = 0$.

The effective radius of a NSC mildly correlates to its mass (Neumayer et al. 2020). The best-fit relation between the mass of the NSC and its effective radius is

$$\log\left(\frac{r_h}{\mathrm{pc}}\right) = \begin{cases} 0.538 & M_{NSC} < 10^6\,M_\odot \\ 0.228\,\log\left(\frac{M_{NSC}}{M_\odot}\right) - 1.19 & M_{NSC} \geq 10^6\,M_\odot. \end{cases} \quad (5)$$

To account for the spread in the data, we sampled $\log(r_h/\mathrm{pc})$ from a Gaussian distribution with the mean centered on the corresponding value determined by Eq. (5) and width $\sigma = 0.1$ (0.2) for $M_{NSC} < 10^6\,M_\odot$ ($\geq 10^6\,M_\odot$) so that most of the data fell under the $\pm 2\sigma$ dispersion.

We approximated the spatial distribution of stars in the NSC with a Plummer (1911) model, with mass

$$M_{NSC}(R) = M_{NSC}\,\frac{R^3}{\left(R^2 + a_{PL}^2\right)^{3/2}} \quad (6)$$

where R is the distance from the SMBH, and $a_{PL} = r_h/(1.3\,\mathrm{pc})$ is the scale parameter for the Plummer model.

The mass fraction of stellar-origin BHs in the AGN disk f_{BH} is sampled from a Gaussian with mean 0.04 and standard deviation 0.01 to account for mass segregation in the NSC (Bartos et al. 2017). The number of BHs in the AGN disk and their cumulative mass are thus given by the following equations:

$$N_{BH} = \frac{f_{BH}\,f_{trap}}{2}\,\frac{M_{NSC}(R_{max})}{\langle m_* \rangle}, \quad M_{BH}^{max} = N_{BH}\,\langle m_{BH} \rangle, \quad (7)$$

where $M_{\rm NSC}(R_{\rm max})$ is defined in Eq. (6), the mean stellar-origin BH mass $\langle m_{\rm BH}\rangle$ is computed from data obtained from the population synthesis simulation code SEVN[2] (Iorio et al. 2023) at solar metallicity (Appendix A), the distance $R_{\rm max}$ is the outer radius of the disk (Sect. 2.1), the factor 1/2 accounts for prograde orbiters[3] only, and the mean stellar mass $\langle m_*\rangle \simeq 1\,M_\odot$ is computed using a Kroupa (2001) initial mass function.

The parameter $f_{\rm trap}$ is the ratio between the number of BHs that are able to reach the migration trap on a timescale shorter than the disk lifetime τ and the total number of BHs that interact with the disk: this study focuses on BBH pair-ups in the migration trap, and therefore we are not interested in any BHs situated outside of that location. An operational definition of $f_{\rm trap}$ is given in Sect. 2.3.3.

We assumed the velocity dispersion of stars to scale with the SMBH mass as (Merritt & Ferrarese 2001)

$$\sigma = 200\,{\rm km\,s^{-1}}\left(\frac{M_{\rm SMBH}}{10^8\,M_\odot}\right)^{1/5}. \tag{8}$$

However, it has been shown (e.g. Scott & Graham 2013; Sahu et al. 2019; Graham 2022) that this result depends on the morphology of galaxies included in the sample. For example, Sahu et al. (2019) find an exponent of ∼1/6 and show that it is caused by the combined contribution of Sérsic (i.e., following the Sérsic 1963 brightness profile) and core-Sérsic (i.e., centrally depleted) galaxies following two different $M_{\rm SMBH}-\sigma$ relations with exponents $(5.75\pm0.34)^{-1}$ and $(8.64\pm1.10)^{-1}$, respectively.

2.3. First-generation (1g) BHs

We randomly drew first-generation (1g, i.e., stellar-origin) BH masses $m_{\rm BH}$ from a catalog obtained with the population synthesis code SEVN (Spera & Mapelli 2017; Spera et al. 2019; Mapelli et al. 2020; Iorio et al. 2023). SEVN relies on up-to-date stellar tracks (Bressan et al. 2012; Costa et al. 2019; Nguyen et al. 2022) and models the formation of compact objects by taking into account electron-capture (Giacobbo & Mapelli 2019), core-collapse (Iorio et al. 2023) and pair-instability supernovae (Mapelli et al. 2020). In particular, here we used the rapid core-collapse supernova model from Fryer et al. (2012), which enforces the existence of a mass gap between the maximum neutron star mass and the minimum BH mass (Özel et al. 2010). We used the fiducial model from Iorio et al. (2023) and considered single stellar evolution only, as described in Appendix A. We assumed metallicity $Z = 0.02$ (i.e., approximately solar), which matches the typical metallicity at the center of massive galaxies in the Local Universe (Gallazzi et al. 2008).

We randomly drew an initial radial position for each BH. The radial extension of the accretion disk is small compared to the typical dimension of a NSC, so we neglected mass segregation on this scale and considered the numerical density of objects at radii $R < R_{\rm max}$ to be uniform in radius: $R \sim \mathcal{U}(R_{\rm min}, R_{\rm max})$.

We drew the dimensionless spin magnitude, χ, from a Maxwellian distribution with one-dimensional root-mean-square $\sigma_\chi = 0.1$, truncated at $\chi = 1$. We chose $\sigma_\chi = 0.1$ because it is reminiscent of the spins inferred from the third GW transient catalog (GWTC-3, Abbott et al. 2023b). This assumption does not take into account that the BBH population in GWTC-3 likely comes from multiple formation channels, including the AGN disk scenario. Current data are not sufficiently informative to differentiate between formation channels. We set the primary spin tilt as in Appendix B.

2.3.1. Gas capture

After setting up the properties of 1g BHs, we followed their evolution in the disk. When NSC objects orbit around the central SMBH, their orbits can cross the disk and gather some of the disk gas, subjecting them to strong gas drag. This is expected to dampen both the inclination i and the eccentricity e of their orbit (Cresswell et al. 2007). Therefore, after a sufficient number of laps, these objects will have circular orbits embedded in the disk. This process is called gas capture or orbital damping.

The gas accretion and subsequent gas drag are significant only for prograde orbiters. Hence, we neglected any variation in the orbits of retrograde orbiters.

We defined the inclination damping timescale, $t_{\rm damp}$, for a BH of mass, $m_{\rm BH}$, on an initial orbit of semimajor axis[4], A, around a SMBH of mass, $M_{\rm SMBH}$, as (Wang et al. 2024)

$$t_{\rm damp} \simeq \frac{c^4}{18\,G\,\Sigma_{\rm gas}\,m_{\rm BH}}\sqrt{\frac{A^3}{G\,(M_{\rm SMBH}+m_{\rm BH})}}, \tag{9}$$

where $\Sigma_{\rm gas}\,(A)$ is the surface density of the gas. In this model, we neglect the mass increase due to gas accretion. Hence, during the orbital evolution, $m_{\rm BH}$ is a constant quantity.

2.3.2. Migration

Once a BH is embedded in the disk, it exchanges angular momentum with the surrounding gas and is subject to gas torques. Torques can be both positive or negative, leading to outward or inward migration, respectively. Similar to what happens to planet seeds in protoplanetary disks, migration can happen in two different ways called Type I and Type II.

Small to medium-mass objects are subject to Type I migration, meaning that they change their radial position in the disk without significantly perturbing the density distribution of the disk itself. For a BH of mass $m_{\rm BH}$ on a circular orbit with radius R, this happens on the timescale (Lyra et al. 2010; McKernan et al. 2012; Baruteau et al. 2014)

$$t_{\rm migr,I} = \frac{M_{\rm SMBH}^2\,h^2}{m_{\rm BH}\,\Sigma_{\rm gas}\,R^2\,\Omega}, \tag{10}$$

where h is the aspect ratio of the disk and Ω is the Keplerian angular velocity around the SMBH. Differently from Eq. (9), here we are considering a radius, R, rather than a semimajor axis, A, because gas capture happens necessarily before migration[5], so the orbits have already been circularized when migration sets in.

In our disk model (Fig. 1), torques are positive in the inner region of the disk, where the slope of the surface density is

[2] The latest public version of SEVN can be downloaded from this repository: https://gitlab.com/sevncodes/sevn

[3] Prograde orbiters are objects that orbit in the same direction as the disk. Here, we assume that half of the total BH population in the NSC are prograde orbiters.

[4] We often use coordinates such as semimajor axes and radii. We use capital letters A and R to refer to orbits around the central SMBH while we use lower-case letters a and r for orbits inside a binary system.

[5] Migration can only set-in when $i = 0$ and the orbit is embedded in the disk. Because of gas drag, $i \to 0$ and $e \to 0$ on roughly the same timescale.

positive, and they are negative in the outer region, where the slope of the surface density is negative (Bellovary et al. 2016). Therefore, Type I migration is directed outward in the inner disk and inward in the outer disk. At the location where the torques change sign, called a migration trap, migration will stall leading to a large accumulation of objects. Hence, after a timescale $t_{\rm migr,\,I}$ (Eq. (10)), the migrating object will be in the migration trap.

Larger objects, on the other hand, can open gaps in the disk. This happens because the motion of a massive object exerts an intense tidal perturbation on the disk, which effectively pushes material away from the orbit's trail (Bryden et al. 1999). This is called Type II migration. An object of mass $m_{\rm BH}$ can open a gap in the disk if (McKernan et al. 2014)

$$q > \sqrt{\frac{\alpha}{0.09}\,h^5}, \tag{11}$$

where $q = m_{\rm BH}/M_{\rm SMBH}$ is the mass ratio with respect to the central SMBH, α is the viscosity parameter, and $h = h(R)$ is the aspect ratio of the disk at radius R.

If an object opens a gap in the disk, assuming that no gas can cross the gap, its migration follows the viscous evolution of the disk's gas. Hence, the timescale for Type II migration is the timescale for the viscous evolution of the disk (McKernan et al. 2012):

$$t_{\rm migr,\,II} = t_{\rm visc} = (\alpha\, h^2\, \Omega)^{-1}. \tag{12}$$

However, pressure forces in the disk push to close the gap. So, even if an object is massive enough to open a gap, such a gap can stay open against pressure forces only if (Bryden et al. 1999; McKernan et al. 2014)

$$q \gtrsim \alpha\,(40\,h)^2\,. \tag{13}$$

Type II migrators typically have high mass: taking as fiducial values $\alpha = 0.1$ and $h = 0.02$, Eqs. (11) and (13) entail $m_{\rm BH} \gtrsim 0.1\,M_{\rm SMBH}$. They are bound to their radial location in the disk and can only move with the disk on its viscous timescale. Hence, they will never reach the migration trap. Moreover, the gaps they create will prevent some Type I migrators from reaching the trap: they will intercept inward-moving migrators if they are located at a radius greater than the trap's, or they will intercept outward moving migrators if the opposite is true. These intercepted BHs can potentially pair up and merge with the Type II migrator, although their merger would not be assisted by GH. However, we expect such a massive BH in an AGN disk only in two cases: a high-generation hierarchical BH or the central BH of a dwarf galaxy dragged into the AGN disk after a galaxy–galaxy merger (e.g., Di Matteo et al. 2008). Describing galaxy mergers is out of the scope of this paper, while a high-generation hierarchical BH can form only in the late stages of the disk's lifetime. We can assume that by that point most of the BHs have already migrated in the migration trap, and we can neglect any further pair-up event with a Type II migrator.

Therefore, in our model we consider the onset of Type II migration to be one of the processes that can halt hierarchical mergers. Moreover, the presence of gaps has noteworthy consequences on disk structure. Nevertheless, we bluntly neglect any evolution of the disk density profile.

2.3.3. Pair-up

Depending on the physical properties of the AGN (such as viscosity, gas density, aspect ratio, and SMBH mass), gas capture and migration can happen on short timescales. When these processes are efficient, they can lead to a large abundance of BHs in the narrow region of the migration trap. Also, all BHs in the migration traps are on similar orbits (prograde and quasi-Keplerian), so their relative velocities of encounter are small. Under such conditions, it is easy for two BHs to become gravitationally bound in a binary. Therefore, efficient damping and migration lead to efficient binary pair-up.

In this work, we assume that the pair-up of a primary and secondary BH is immediate as soon as the primary reaches the migration trap. This assumption is fully justified by the efficiency of dynamical friction in the migration trap (see Qian et al. 2024 and Sect. 4). The pairing timescale of a BBH is therefore

$$t_{\rm pair} = t_{\rm damp} + t_{\rm migr,\,I} + t_{\rm in}, \tag{14}$$

where $t_{\rm in}$ is the formation time of the primary BH since the time of formation of the disk.

We considered each BBH to be in a circular orbit in the migration trap at a radius $R_{\rm trap}$ (Sect. 2.1). We computed the fraction, $f_{\rm trap}$, of BHs that reach the migration trap by counting the number of $1g$ BHs for which $t_{\rm pair} < \tau$, where τ is the disk lifetime, and dividing it by the total number, N, of simulated first-generation BHs. We used this parameter for the computation of the maximum mass that can be accreted by a single BH, as in Eq. (7).

2.3.4. BBH properties

We determined the secondary BH mass as in Appendix C, and we set the secondary BH spin in the same way as for the primary BH. We set the secondary BH spin tilt as in Appendix B. We assigned the initial semimajor axis, a, of the binary sampling from a distribution $p(a) \propto a^{9/2}$ (Binney & Tremaine 2008; Tagawa et al. 2020b) for $a \in [a_{\rm min}, a_{\rm max}]$, where $a_{\rm min} = 1\,R_\odot$ and $a_{\rm max}$ is equal to the Hill radius, computed as

$$R_{\rm Hill} = R_{\rm trap}\left(\frac{m_1 + m_2}{3\,M_{\rm SMBH}}\right)^{1/3}, \tag{15}$$

where m_1 and m_2 are the primary and secondary BH mass, respectively. We set the initial eccentricity e following a thermal distribution $p(e) \propto 2e$ for e between 0 and 1 (Jeans 1919).

According to the Heggie (1975) law, a binary can survive in a star cluster only if it is hard, meaning that its binding energy is greater than the average kinetic energy of a field star

$$E_b = \frac{G\,m_1\,m_2}{2\,a} \geq \langle E_k\rangle = \frac{1}{2}\,\langle m_*\rangle\,\sigma^2, \tag{16}$$

where a is the semimajor axis of the binary, $\langle m_*\rangle$ is the average mass of a star in the NSC, and σ is the three-dimensional velocity dispersion. Hence, we dynamically evolveed hard binaries only.

2.4. Orbital evolution and merger

Once the binary is in the migration trap, we set $R = R_{\rm trap}$ and update the quantities in Eqs. (1)–(3) accordingly. The semimajor axis, a, and eccentricity, e, of a binary with component masses m_1 and m_2, embedded in the disk, evolve due to GH, irrespective of whether the binary is prograde or retrograde, as

(Ishibashi & Gröbner 2020)

$$\dot{a}_{\rm gas} = -\frac{24\pi\,\alpha\,c_{\rm s}^2\,\Sigma_{\rm gas}\,(1+e)^2\,a}{\mu\,\Omega_{\rm b}}, \tag{17}$$

$$\dot{e}_{\rm gas} = \frac{12\pi\,\alpha\,c_{\rm s}^2\,\Sigma_{\rm gas}\,(1-e^2)^{1/2}\,(1+e)^2\,[1-(1-e^2)^{1/2}]}{\mu\,\Omega_{\rm b}\,e}, \tag{18}$$

where $\mu = m_1 m_2/(m_1+m_2)$ is the reduced mass and $\Omega_{\rm b} = \sqrt{G(m_1+m_2)/a^3}$ is the Keplerian orbital frequency.

The binary also hardens due to the effect of GW emission, which will govern the evolution at small semimajor axes. The evolution of the semimajor axis, $\dot{a}_{\rm GW}$, and eccentricity, $\dot{e}_{\rm GW}$, due to GW hardening proceeds as in Peters (1964).

The interaction with other objects also contributes to the hardening of hard binaries according to Heggie (1975). We neglected the hardening effect due to three-body interactions because they typically occur on a timescale longer than that of GH (Leigh et al. 2017, see discussion in Sect. 4.3). The overall evolution of the binary is thus described by

$$\dot{a} = \dot{a}_{\rm gas} + \dot{a}_{\rm GW}\,, \qquad \dot{e} = \dot{e}_{\rm gas} + \dot{e}_{\rm GW}. \tag{19}$$

The equations for GH (Eqs. (17) and (18)) are valid under the assumption that gravitational torques from the binary act axisymmetrically upon the disk so that they clear a cavity in the surrounding gas distribution, which remains circular throughout its inspiral. This is an idealized model since cavities in AGN disks can become eccentric (MacFadyen & Milosavljević 2008; Cimerman & Rafikov 2024) and can lead the orbital separation to grow in time (Miranda et al. 2016). Hence, we considered an "optimistic" model in which the evolution is given by Eq. (19) (hereafter, the GH model), and a "pessimistic" model in which we neglect GH and only consider the effect of GW emission (hereafter, the no-GH model). In both the optimistic and pessimistic cases, we integrated the hardening equations using the Euler method and an adaptive timestep (Mapelli et al. 2021).

We refer to the delay time between the pair-up and the merger as $t_{\rm del}$. If $t_{\rm pair} + t_{\rm del}$ is longer than the lifetime of the disk, the disk has evaporated before the binary could merge. In this case, the BBH keeps hardening due to GW emission only.

We assumed that the BBH merges when its members cross the ISCO radius of a non-spinning BH with mass equal to the total mass of the binary system, $r_{\rm ISCO} = 6\,G\,(m_1+m_2)/c^2$, with a tolerance of $0.1\,r_{\rm ISCO}$. This happens on a merger timescale

$$t_{\rm merg} = t_{\rm pair} + t_{\rm del} = t_{\rm in} + t_{\rm damp} + t_{\rm migr,I} + t_{\rm del}. \tag{20}$$

We modeled the mass and spin of the merger remnant using fitting formulas from numerical relativity, as described by Jiménez-Forteza et al. (2017). At birth, merger remnants receive a relativistic kick $v_{\rm kick}$ because of the transfer of linear momentum caused by asymmetries in GW emission. We used the model of Maggiore (2018, Eq. (14.202)) for the magnitude and direction of the kick velocity, $\boldsymbol{v}_{\rm kick}$.

The relativistic kick $\boldsymbol{v}_{\rm kick}$ pushes the merger remnant out of the migration trap. We computed the new velocity as $\boldsymbol{v}_{\rm fin} = \boldsymbol{v}_{\rm Kepl}(R_{\rm trap}) + \boldsymbol{v}_{\rm kick}$, where the Keplerian velocity in the migration trap is computed accounting for the mass of the SMBH $M_{\rm SMBH}$ and the inner part of the NSC $M_{\rm NSC}$, while neglecting the mass of the gas disk:

$$v_{\rm Kepl}\,(R) = \sqrt{\frac{G\,(M_{\rm SMBH} + M_{\rm NSC}(R)\,)}{R}} \equiv \sqrt{\frac{GM_{\rm TOT}\,(R)}{R}}. \tag{21}$$

The final semimajor axis $A_{\rm fin}$ of the remnant orbit, computed by means of simple orbital transfer calculations (Hohmann 1960), is

$$A_{\rm fin} = \frac{GM_{\rm TOT}(R_{\rm trap})}{v_{\rm fin}^2}\left[\frac{R_{\rm trap}\,v_{\rm fin}^2/GM_{\rm TOT}(R_{\rm trap})}{2 - R_{\rm trap}\,v_{\rm fin}^2/GM_{\rm TOT}(R_{\rm trap})}\right]. \tag{22}$$

The quantities in Eqs. (1)–(3) were updated accordingly.

As a safety check, we ensured that the new orbital semimajor axis, $A_{\rm fin}$, is smaller than the maximum radius of the disk, $R_{\rm max}$, meaning that the remnant can experience damping and become embedded in the disk. Otherwise, we did not consider the remnant for future generations.

2.5. Nth-generation (Ng) BHs

A seed BH can only go through a finite number of hierarchical mergers before it comes across one of the following scenarios:

1. The disk has evaporated. Therefore, damping, migration, and any other effects due to gas torques stop. We evaluated this by checking if, at generation N, the merger timescale, $t_{\rm merg}^{(N)}$, is shorter than the disk lifetime, τ. In our formalism, we defined the merger timescale of a Ng BH as $t_{\rm merg}^{(N)} = t_{\rm in}^{(N)} + \left[t_{\rm damp} + t_{\rm migr,I} + t_{\rm del}\right]^{(N)}$, where $t_{\rm in}^{(N)}$ is the evolutionary time of the previous generations.
2. The relativistic kick received at merger is so strong that the remnant is ejected from the AGN. We computed the escape velocity considering only the gravitational potential of the SMBH and the inner NSC, while neglecting the mass of the gaseous disk, as

$$v_{\rm esc}\,(R) = \sqrt{\frac{2\,G\,M_{\rm TOT}\,(R)}{R}}. \tag{23}$$

   Then we computed the final velocity after kick $\boldsymbol{v}_{\rm fin} = \boldsymbol{v}_{\rm in} + \boldsymbol{v}_{\rm kick}$ and ensured that it is smaller than the escape velocity, $v_{\rm esc}$. We also ensured that the merger remnant is on a disk-crossing orbit by comparing its semimajor axis, $A_{\rm fin}$, with the outermost disk radius, $R_{\rm max}$.
3. The number of BHs in the NSC is finite. Therefore, the mass that can be accreted is limited and the BH may not find any more companions to pair up with. We kept track of the mass accreted by a single BH, which is the sum of its initial mass, m_1, and of the masses of all the secondaries, $m_2^{(N)}$, it pairs up and merges with

$$M_{\rm acc} = m_1 + \sum_N m_2^{(N)}, \tag{24}$$

   where the index N represents the generation number.
   We ensure that the BH does not accrete more mass than what is available in the inner NSC in the form of other BHs, namely $M_{\rm acc} \leq M_{\rm BH}^{\rm max}$, where $M_{\rm BH}^{\rm max}$ is obtained as in Eq. (7).
4. The BH is so massive that it opens a gap in the disk and can only move from its radial location due to Type II migration. We checked whether both conditions in Eqs. (11) and (13) are respected. For typical values of viscosity and aspect ratio, these conditions entail $m_{\rm BH} \gtrsim 10^{-1} M_{\rm SMBH}$.

If any of conditions (1)–(4) was met, we did not consider the merger remnant for future generations. We followed the evolution of Nth generation BHs with a procedure similar to the one outlined for first-generation BHs. In hierarchical mergers,

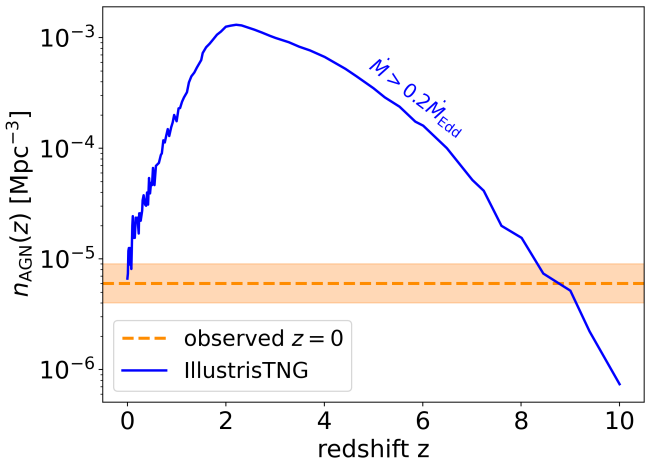

Fig. 2. Number density of active SMBHs per unit comoving volume, $n_{AGN}(z)$, in the ILLUSTRISTNG100 simulation (solid blue lines) vs. the observational values at redshift zero for unobscured (hydrogen column density $N_H \lesssim 10^{23}$ cm^{-2}) high-luminosity ($L \geq 10^{43.2}$ erg s^{-1}) AGNs.

the remnant of an $(N-1)$th-generation merger acts as the primary BH for the Nth-generation. So, the primary mass and spin are simply set as the remnant mass and spin of the previous generation, computed according to Jiménez-Forteza et al. (2017). Similarly, the initial position of the primary component of the Nth-generation BBH is set as the position of the merger remnant of the previous generation, set in Eq. (22). This value is used to compute the pairing time $t_{pair} = t_{damp} + t_{migr,I} + t_{in}$, where for Nth generations $t_{in}^{(N)} = t_{merg}^{(N-1)}$. The migration and pair-up physics are the same as described for the first generation.

We modeled the secondary component of each Nth-generation BBH as described in Appendix C. We define an Nth-generation (Ng) BH to be the result of the repeated merger of N stellar-origin BHs. For instance, an Ng BH can be the result of either an $Mg - 1g$ merger (where $M + 1 = N$) or an $Mg - Lg$ merger (where $M + L = N$, $L > 1$). We kept iterating the same procedure until the considered BH met at least one of requirements (1)–(4).

2.6. Cosmic evolution

We modeled the cosmic evolution of AGNs using data from the ILLUSTRISTNG100, a magneto-hydrodynamical cosmological simulation adopting a cubic box of size $(110.7\,\mathrm{Mpc})^3$ with a resolution of roughly $(6\,\mathrm{Kpc})^3$ (see Pillepich et al. 2018 and Springel et al. 2018 for further details).

We calibrated the AGN density distribution $n_{AGN}(z)$ in the ILLUSTRISTNG data by choosing a threshold in the mass accretion rate such that the value at redshift zero, $n_{AGN}(0)$, is compatible with the observational value of $\sim 6 \times 10^{-6}\,\mathrm{Mpc}^{-3}$ obtained from a sample of X-ray-selected AGNs (Buchner et al. 2015). As displayed in Fig. 2, we find that setting this threshold to a fifth of the Eddington mass accretion rate, namely $\dot{M}_{SMBH} \geq 0.2\,\dot{M}_{Edd}$, provides an appropriate normalization. We employed $n_{AGN}(z)$ to compute the BBH merger rate properties in various redshift bins as explained in Appendix D.

The SMBH mass distribution is expected to evolve as a function of redshift (Weinberger et al. 2018). Nevertheless, since the mass evolution cannot be constrained by current data, we used the observational distribution from Greene & Ho (2007) at every redshift for simplicity.

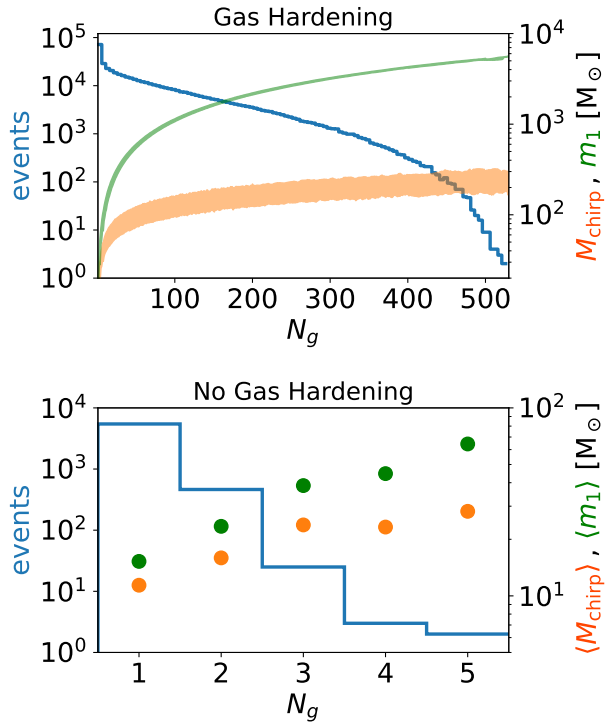

Fig. 3. Number of mergers and characteristic masses at each merger generation. The blue histogram and the left-hand y-axis show the number of BBH merger events for each hierarchical merger generation, N_g. The orange (green) shaded area and the upper right-hand y-axis show the 25%–75% percentile of the chirp mass (primary mass) for merging BBHs of each generation. The orange (green) dots and lower right-hand y-axis show the average chirp mass (primary mass) for merging BBHs of each generation.

2.7. Description of runs

We considered two different models for the AGN channel: with and without GH (hereafter, GH and no-GH). We ran $N = 10^5$ realizations of our two models, each with different SMBH mass and AGN lifetime randomly extracted as described at the end of Sect. 2.1. In each run, we simulated all the BHs that reach the migration trap within the AGN disk lifetime.

Furthermore, we simulated four other channels (isolated, YSC, GC, and NSC) with the same code (FASTCLUSTER; Mapelli et al. 2021, 2022) and using the same underlying initial conditions, such as the stellar evolution model determining the $1g$ BHs mass distribution (Appendix A). This allowed us to filter out any bias that might arise by using different numerical codes for different environments: the differences we see are not due to the numerical approach adopted but to intrinsic differences among channels. We performed a Bayesian analysis to determine the AGN mixing fraction as described in Appendix E.

3. Results

3.1. Impact of gas hardening

We find that the efficiency of the hierarchical merger process is significantly higher in the presence of GH: in our GH model, seed BHs can go through up to roughly 500 merger episodes, whereas in the no-GH case the hierarchical chain usually stops after a few generations (Fig. 3). In the GH (no-GH) model, the hierarchical merger chain stops because of the condition on timescales, maximum mass, Type II migration, or ejection in the

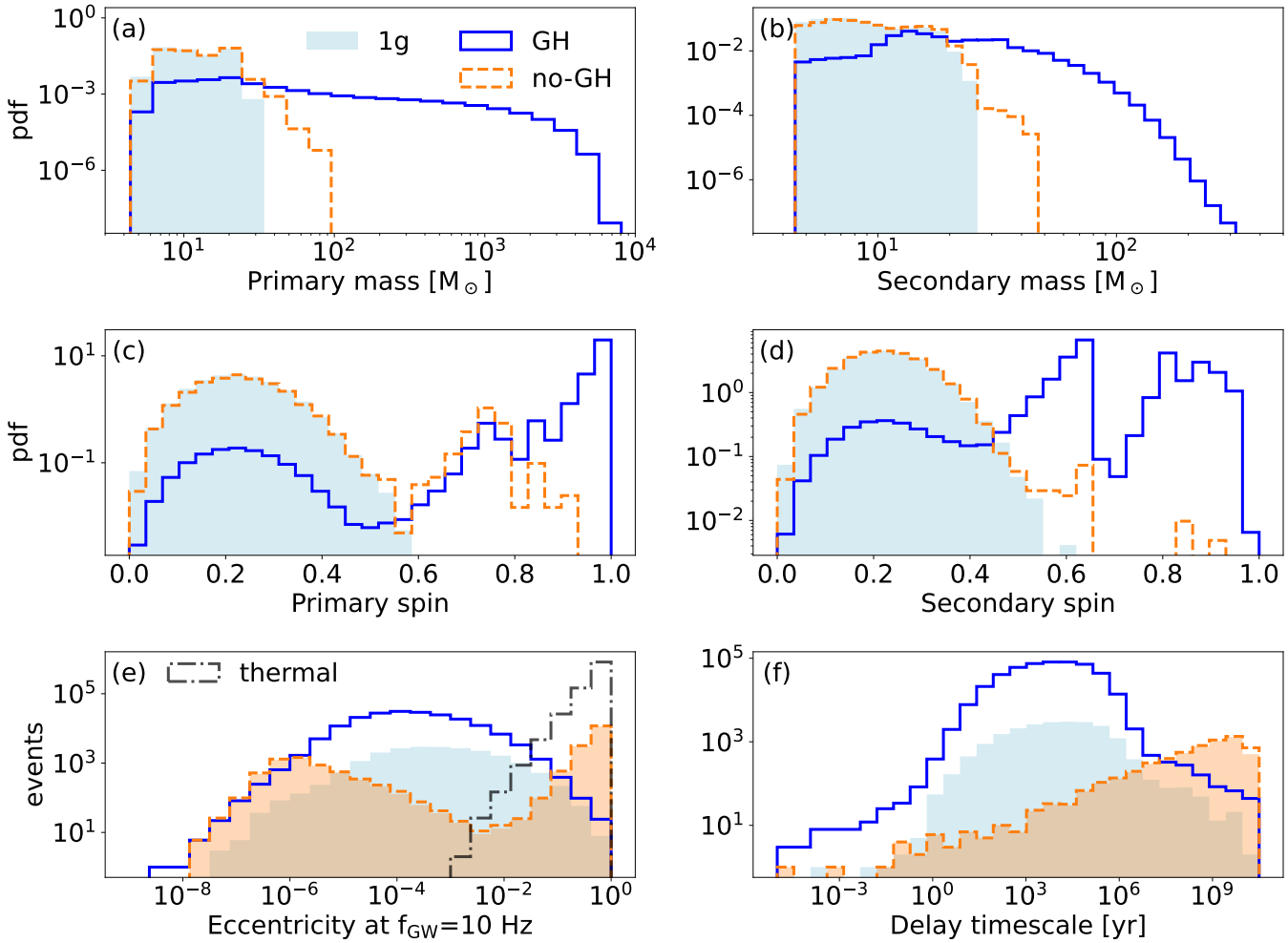

Fig. 4. Main properties of dynamical BBH mergers in our GH and no-GH AGN models. The unfilled solid blue (dashed orange) histogram shows BBHs in the GH (no-GH) AGN scenario. The filled blue (orange) histogram shows the 1g BBHs in the GH (no-GH) AGN scenario. If 1g BHs in the no-GH scenario are identical to the GH scenario, they are not shown. Panels (a) and (b): probability density function (pdf) of the primary and secondary BH mass (m_1 and m_2). Panels (c) and (d): primary and secondary spin magnitudes (χ_1 and χ_2). Panel (e): orbital BBH eccentricity $e(10\,\mathrm{Hz})$ when the GW frequency is $f_{\rm GW} = 10$ Hz. In gray we show the thermal distribution $p(e) = 2e$ for comparison. Panel (f): timescales of delay time, $t_{\rm del}$, between the BBH pair-up and the merger.

58.7% (81.6%), 27.1% (18.0%), 13.6% (0%), and 0.6% (0.4%) of simulations, respectively.

Figure 4 shows the main properties of dynamically assembled BBHs in our AGN simulations. We display only BBHs that merge within a Hubble time. In both models, hierarchical mergers in the migration trap are successful in producing BBHs with primary mass in the pair-instability mass gap and above, but there is a major difference between the GH and no-GH scenarios: the primary BH mass extends only up to 300 $M_\odot$ in the no-GH scenario, while it reaches $\approx 5\times10^3\,M_\odot$ in the GH scenario, because GH dramatically increases the efficiency of hierarchical mergers. Hence, the GH mechanism produces a recognizable feature in the high-mass end ($m_1 \gtrsim 100\,M_\odot$) of the mass spectrum. This is caused by the stark difference in delay timescales (Fig. 4f) due to the different GH prescriptions: in the GH scenario, the relative distribution has a pronounced peak between 1 yr and 1 Myr, which in the no-GH case is absent. Information on other relevant timescales is illustrated in Sect. 3.4. Secondary BHs can also have masses in the pair-instability mass gap but only up to a few times $10^2\,M_\odot$ in the GH scenario, as the AGN channel tends to favor mergers with low mass ratio (Fig. 4b).

The GH model has a sharp peak in the distribution of the primary spin magnitude at $\chi_1 = 1$ as well as a peak in the distribution of the secondary spin magnitude at $\chi_2 \simeq 0.9$ associated with high-generation mergers (Fig. 4c,d). Both models GH and no-GH also have a peak at $\chi_1 \approx 0.75$ (also $\chi_2 \approx 0.75$ for the GH case), corresponding to the second generation.

The distribution of spin-tilt angles is influenced by GH, as shown in Fig. 5 and explained in Appendix B. In the no-GH scenario, the spin tilts $\theta_{1,2}$ are isotropically distributed with respect to the orbital angular momentum $\boldsymbol{L}$. Instead, in the presence of GH, $\boldsymbol{\chi}_1, \boldsymbol{\chi}_2$, and $\boldsymbol{L}$ are all preferentially aligned with each other. The distribution of the secondary spin tilt θ_2 is slightly wider than the primary's spin tilt, because of a cumulative effect: the misalignment between $\boldsymbol{\chi}_2$ and $\boldsymbol{L}$ is set keeping into account the misalignment between $\boldsymbol{\chi}_1$ and $\boldsymbol{L}$ (see Appendix B).

Both scenarios preferentially produce BBHs with low eccentricity when[6] the GW frequency is 10Hz. Nevertheless, they can produce BBHs with eccentricity $e \geq 10^{-2}$ (Fig. 4e), which is potentially detectable by the LIGO-Virgo-KAGRA (LVK)

[6] Here we only refer to BBH mergers with ISCO frequency in the LVK detectability range, $f_{\rm GW}^{\rm ISCO} \geq 10$ Hz. See Sect. 4.1 for more details.

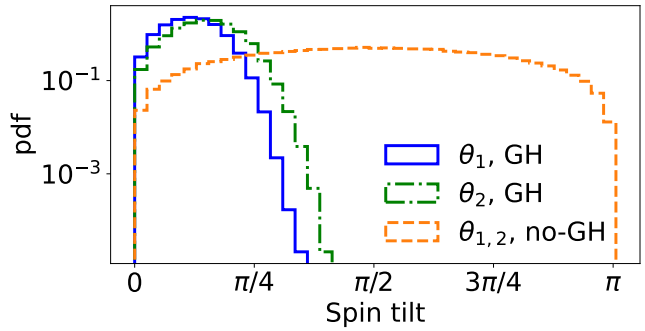

Fig. 5. Probability density function (pdf) of spin tilt angles θ_1 and θ_2. The blue (green) histogram shows the primary (secondary) spin tilt in the GH scenario. The unfilled orange histogram shows the primary and secondary tilts in the no-GH scenario. The first generation is not shown separately in this plot because the distributions are identical at each generation.

interferometers (Romero-Shaw et al. 2021). In our GH model, there are two processes at play: GH pumps the eccentricity (Eq. (18)), whereas GW emission damps it (Peters 1964). Hence, the GH scenario is more likely to produce BBHs with eccentricity in the range $e \in [10^{-5}, 10^{-2}]$ for $f_{\rm GW} = 10$ Hz than the no-GH one. In a more accurate model, we should also include the effect of three-body scatterings (Samsing et al. 2022) and tidal forces exerted by the SMBH (Rom et al. 2024), both of which pump BBH eccentricity.

Figure 6 shows the relation between the BBH mass $m_1 + m_2$ and mass ratio m_2/m_1 in our models, compared to GW data (Abbott et al. 2023a,b). The AGN channel tends to favor mergers with low mass ratio (Appendix C). This effect is particularly important for high primary BH masses $m_1 \gtrsim 10^2\,M_\odot$ when accounting for GH.

The effective spin distribution in the GH model (Fig. 7) peaks at $\chi_{\rm eff} \sim 1$, corresponding to maximum alignment, and shows lower peaks at $\chi_{\rm eff} \simeq 0.2$ and $\chi_{\rm eff} \simeq 0.5$. From Fig. 8, we see that $1g$ mergers populate the peak at $\chi_{\rm eff} \simeq 0.2$, $2g$ mergers create a peak at $\chi_{\rm eff} \simeq 0.5$, while third- and higher-generation BBHs contribute the main peak at $\chi_{\rm eff} \simeq 1$. The precession spin distribution, instead, has no sharp features in this scenario.

In the no-GH scenario, the orbital angular momentum is isotropically oriented with respect to the AGN disk (Appendix B). This produces a bell-shaped effective spin distribution centered on $\chi_{\rm eff} = 0$ with half-width at half-maximum HWHM $\simeq 0.3$ and a precession spin distribution that peaks at $\chi_{\rm p} \simeq 0.2$ and $\chi_{\rm p} \simeq 0.75$ (Fig. 7). As the BBH generation increases, so does the magnitude of its BH spins, making the corresponding effective spin distribution progressively wider as shown in Fig. 8. The peaks at $\chi_{\rm p} \simeq 0.2$ and 0.7 are populated mostly by $1g$ and $2g$ mergers, respectively, whereas the tail at $\chi_{\rm p} > 0.8$ results from higher-generation mergers.

3.2. Anticorrelation between q and $\chi_{\rm eff}$

Figure 9 shows the relationship of the effective spin with primary mass and mass ratio of BBH mergers in our simulations. In the GH scenario, there is a clear correlation between effective spin and BBH mass, as well as a clear anticorrelation between effective spin $\chi_{\rm eff}$ and mass ratio q. This is a possible explanation for the anticorrelation between $\chi_{\rm eff}$ and q found in the LVK data (Callister et al. 2021; Santini et al. 2023). If this anticorrelation stems from hierarchical mergers in the GH regime, our simulations suggest that it should extend to lower mass ratios and higher effective spins than currently observed by LIGO and Virgo.

The anticorrelation is particularly noticeable for $\chi_{\rm eff} \gtrsim 0.4$ because all higher-generation mergers display the following features: high mass, small mass ratio, and high effective spin. In the no-GH model, both the spin alignment and the hierarchical merger process are suppressed; so there is no clear correlation between the aforementioned quantities.

3.3. Comparison with other channels

Figure 10 compares the main properties of BBH mergers in five different formation channels, namely AGN disks, NSCs, GCs, YSCs, and isolated binary evolution (iso). The simulations for NSCs, GCs, YSCs, and iso adopt the same setup as model B of Mapelli et al. (2022), and use as initial conditions catalogs of BH masses derived with SEVN (Iorio et al. 2023), for consistency with the catalogs of the AGN channel. We provide more details about NSC, GC, YSC, and iso simulations in Appendix A.

All dynamical formation channels can produce BBH mergers with masses and spin magnitudes higher than the isolated channel. The no-GH AGN disk scenario resembles most closely the results of the other dynamical channels, whereas the GH AGN disk channel is completely different from the others.

The population of dynamical channels appears as follows: the mass distributions show one or two peaks at a few tens of solar masses and a tail that extends up to a few hundred solar masses; the effective spin distribution is symmetric, peaks at $\chi_{\rm eff} = 0$, and has a half width at half maximum of ∼0.5 (∼0.3 for the YSC channel); and the precession spin distribution has two peaks, at $\chi_{\rm p} \simeq 0.2$ and $\chi_{\rm p} \simeq 0.7$. This contrasts with the isolated channel population, which features a primary (secondary) BH mass distribution that does not extend any higher than $50\,M_\odot$ ($40\,M_\odot$).

The primary BH mass distribution for the GH AGN channel, instead, extends up to $\sim 5 \times 10^3\,M_\odot$. The effective spin distribution has two sharp peaks at $\chi_{\rm eff} \simeq 0.2$ and $\chi_{\rm eff} \simeq 1$, as well as a minor peak at $\chi_{\rm eff} \simeq 0.5$, whereas the precession spin distribution has no evident peaks.

We computed the mixing fractions, $f_{\rm mix}$, for our five channels, as described in Appendix E. To derive $f_{\rm mix}$, we marginalized over the merger rate density, to avoid this extremely uncertain quantity affecting our results. As shown in Fig. 11, the NSC and YSC channels are associated with larger median mixing fractions than the other channels. This happens because, in our models, the NSC channel can account for the peak at low BH mass ($m_1 \approx 10\,M_\odot$) found in the LVK data, while the YSC channel contributes mostly to the high-mass peak at $\sim 35\,M_\odot$ (Abbott et al. 2023b). In fact, in our model, the YSC scenario produces larger median BBH masses than the NSC scenario, because the escape velocity from YSCs is much lower, preventing the low-mass BHs from merging in YSCs (they are ejected by supernova kicks, see the discussion in Mapelli et al. 2021). The isolated channel also produces BBH mergers with a peak of the primary BH mass at $\sim 10\,M_\odot$, but has too little support for zero and negative values of $\chi_{\rm eff}$ with respect to the observed ones. The mass and spin properties of the GC scenario are intermediate between YSCs and NSCs.

There is a small difference between the mixing-fraction distribution of the GH and no-GH scenarios because this analysis only considers detectable events: the most massive BHs of the GH population (which are the main difference with respect to the no-GH model) have no impact on $f_{\rm mix}$. Overall, our results

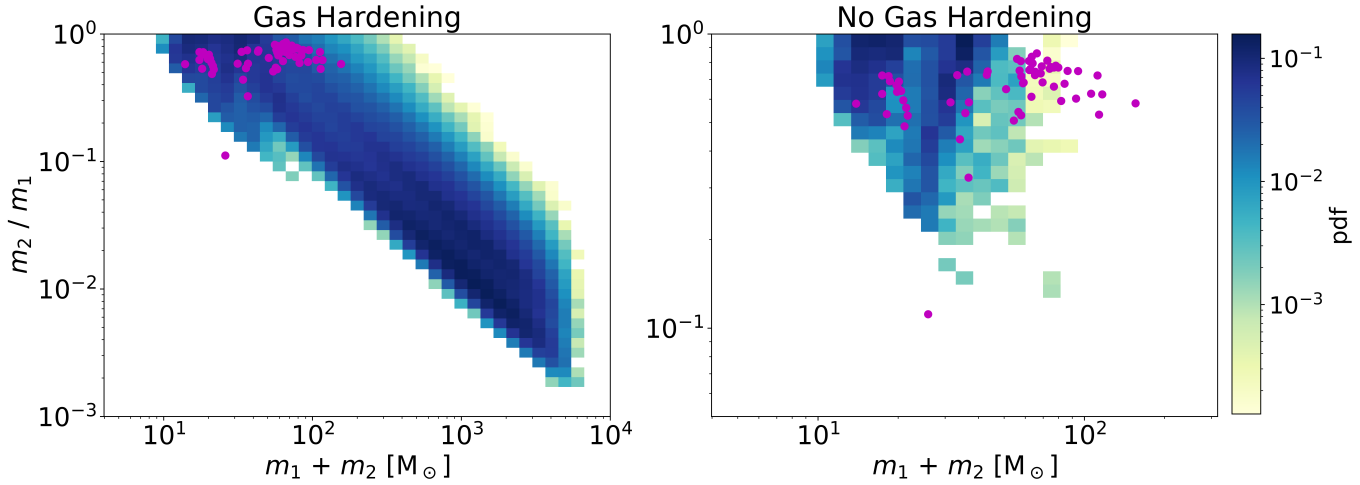

Fig. 6. Probability density function (pdf) of the binary mass and mass ratio for all BBH mergers in our simulations (blue color palette). We do not apply any selection effects to our simulations. Left-hand (right-hand) panels: GH (non-GH) scenario. The magenta dots show the median values of the parameters for LVK BBH merger event candidates with $p_{\rm astro} > 0.9$ from GWTC-3 (Abbott et al. 2023a,b). We do not include error bars for readability purposes.

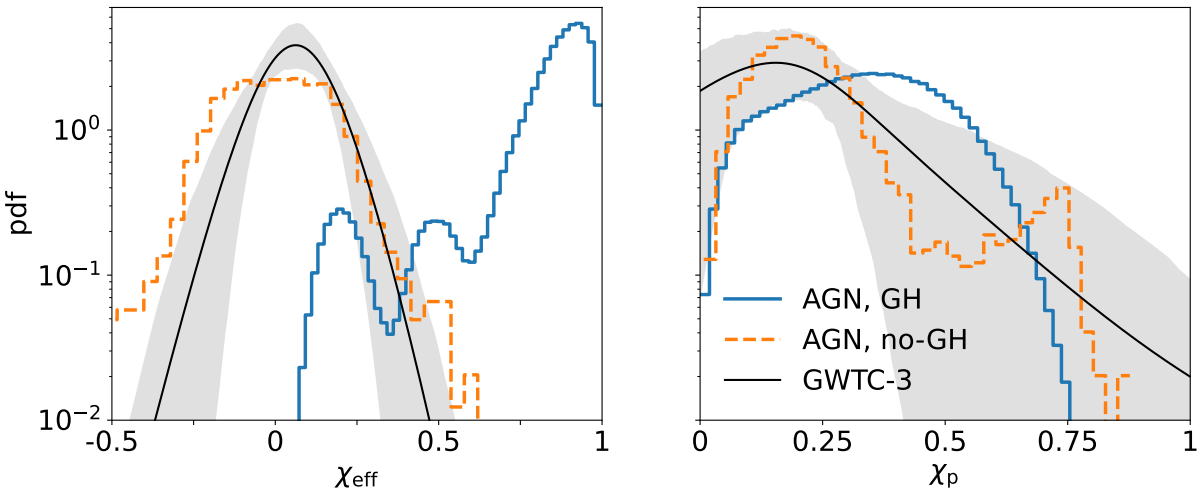

Fig. 7. Effective (left) and precession spin (right) probability density functions (pdf) GH (solid blue lines) and no-GH (dashed orange lines) model. We do not apply any selection effects to our simulations. Solid black lines show the posterior distribution inferred from LVK data, after applying a parametric-model description of the intrinsic BBH population with hierarchical Bayesian inference (Abbott et al. 2023b).

confirm that the current LVK sample of BBH mergers is too small and the uncertainties on theoretical models are too large to draw an informative mixing-fraction analysis with five channels (see Sect. 3.4 of Mapelli et al. 2022).

3.4. Timescales

Figure 12 shows the distribution of all relevant timescales with and without GH. Both the damping time $t_{\rm damp}$ and the migration time $t_{\rm migr,I}$ span a large range, which is representative of the large range in SMBH mass and initial position R. The damping timescale for the first hierarchical merger generation has a flat distribution in the range 10^{-4}–10^{2} Myr, while the migration timescale ranges between 10^{-2} and 10^{3} Myr. In the GH scenario, $t_{\rm damp}$ and $t_{\rm migr,I}$ display additional peaks for subsequent generations at 10^{-5} Myr and 10^{-2} Myr, respectively. The delay timescale $t_{\rm del}$ spans a large range of values, because of the large range encompassed in BH masses and initial BBH semimajor axes. In the GH scenario, it has a pronounced peak at ∼1 yr, whereas in the no-GH case it is typically as large as the Hubble timescale (∼14 Gyr). Indeed, the longest timescale in the no-GH scenario is the delay timescale. In the GH scenario, instead, the evolution is overall governed by the migration timescale.

We also display the timescale for gas dynamical friction τ_0, responsible for BBH pair-up, which we discuss in Sect. 4.2. It is typically on the order of 10^{-13} Myr and can be as short as 10^{-20} Myr for the disk parameters in our model; hence, it is negligible compared to other timescales at play.

4. Discussion

4.1. Outlook for ET and LISA

We studied the evolution of hierarchical mergers in the migration traps of AGN disks and explored the role of GH. Our models, especially the GH scenario, predict the formation of merger events with higher mass than is detectable by current ground-based detectors. Indeed, the frequency emitted by a BBH

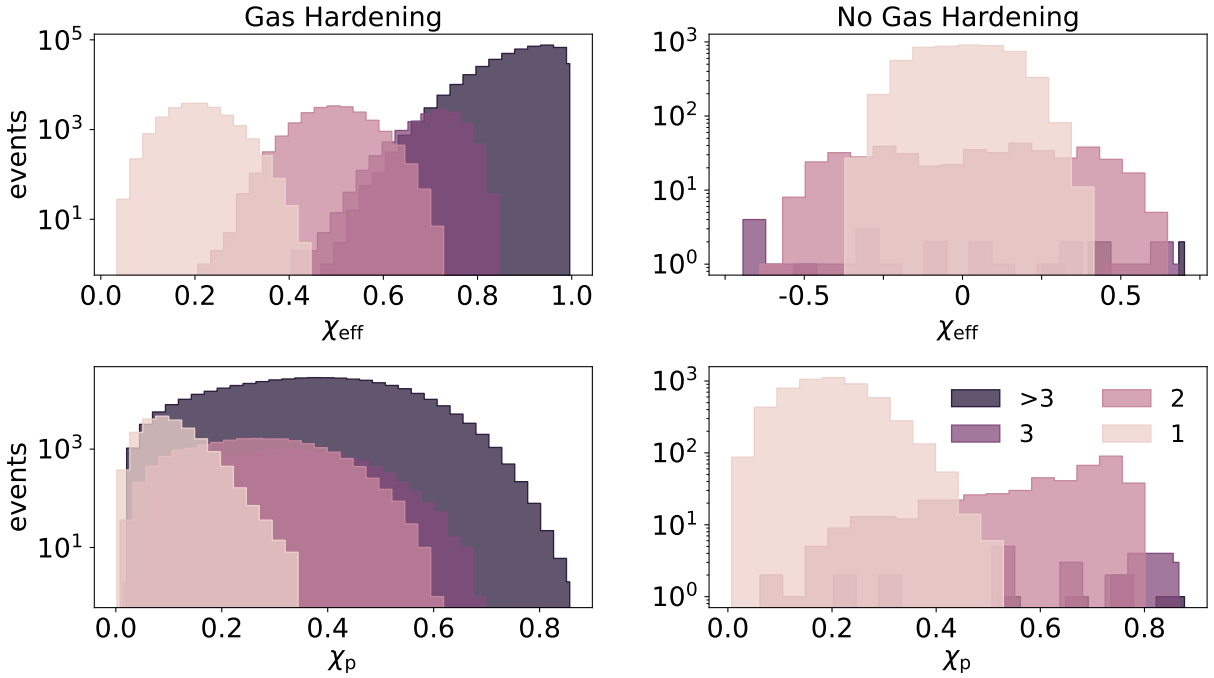

Fig. 8. Effective and precession spin distributions for the GH (left-hand panels) and no-GH (right-hand panels) models, where the shade represents different hierarchical merger generations. The lightest hue refers to the first generation (stellar-origin BHs), while the darkest hue refers to generations higher than the third.

depends on its mass and steadily increases during its inspiral. The frequency of the ISCO is equal to (Maggiore 2008)

$$f_{\rm GW}^{\rm ISCO} = \frac{1}{\pi 6\sqrt{6}} \frac{c^3}{G\,(m_1+m_2)} \simeq 4.4\,{\rm KHz}\left(\frac{M_\odot}{m_1+m_2}\right). \tag{25}$$

Figure 13 shows the probability distribution function of the maximum emitted frequency by AGN BBHs. The LVK interferometers are only sensitive in the frequency range from a few tens of hertz up to ∼1 KHz (Abbott et al. 2023a). Hence, only a fraction of our predicted merger events in the GH scenario are observable with existing detectors.

The next generation ground-based interferometers, the *Einstein* Telescope (ET) and Cosmic Explorer (Punturo et al. 2010; Maggiore et al. 2020; Evans et al. 2021; Branchesi et al. 2023), and space-borne interferometers Laser Interferometer Space Antenna (LISA), Deci-hertz Interferometer GW Observatory (DECIGO) and TianQin (Amaro Seoane et al. 2013, 2017; Luo et al. 2016; Kawamura et al. 2019) will be able to detect GW signals with lower frequency than currently possible. For example, the frequency range of the ET and Cosmic Explorer ($\approx 1-10^4$ Hz) has substantial overlap with the GW frequency of BBH mergers from the GH AGN channel, with the exception of the very high-mass tail. LISA will instead be sensitive to frequencies lower than ∼1 Hz (Robson et al. 2019), so it would be able to detect the highest-mass end of the synthetic mergers predicted by our GH AGN model. We will explore detectability by ET and LISA in detail in a follow-up work.

4.2. BBH pair-up

In our model, we assume that the pairing of a primary and a secondary BH is immediate as soon as the primary reaches the migration trap. A realistic model for BBH pair-up in AGN disks should keep into account GW two-body capture, three-body encounters, and gas dissipation. Whitehead et al. (2023) run hydro-dynamical simulations of close encounters between pairs of BHs embedded in the gaseous disk and find that dissipation by gas gravitation is not always efficient for the formation of bound BBHs. Specifically, they find that it is usually an effective mechanism for BBH formation for gas densities ρ_g in the range $-4.5 \lesssim \log\left[\rho_{\rm g} R_{\rm H}^3/(m_1+m_2)\right] \lesssim -2.5$, where $R_{\rm H}$ is the Hill radius of the BBH. In our simulations, we typically have $\log\left[\rho_{\rm g} R_{\rm H}^3/(m_1+m_2)\right] \simeq -6$ for our BBHs because the high gas density in the migration trap is compensated for by small Hill radii due to the proximity to the SMBH. Hence, gas dissipation is expected to be inefficient.

On the other hand, in their recent N-body simulations, Qian et al. (2024) find that BBH formation from two BHs on similar orbits embedded in the disk is efficient for $\Omega\,\tau_0 \lesssim 10$, where Ω is the Keplerian angular velocity of the BHs and

$$\tau_0 = \frac{c_{\rm s}^3}{4\pi G \rho_{\rm g} m_1} \tag{26}$$

is the dynamical friction timescale (Ostriker 1999), which, as shown in Qian et al. (2024), is linearly related to the timescale of BBH pair-up. In our model typically $\Omega\,\tau_0 \sim 10^{-4}-10^{-3}$ due to the high gas density $\rho_{\rm g}$ in the migration trap, which justifies our assumption.

Moreover, we expect a high number density $n_{\rm BH}$ of BHs in the migration trap due to efficient migration. This may further aid BBH pair-up because the timescales for two-body and three-body captures scale respectively as $n_{\rm BH}^{-1}$ (Quinlan & Shapiro 1990) and $n_{\rm BH}^{-2}$ (Fragione & Silk 2020). In future work, we will further explore the process that leads to BBH pair-up, accounting for these additional effects.

In this work we neglect BBH formation in the bulk (i.e., outside of the migration trap), which is expected to be efficient due

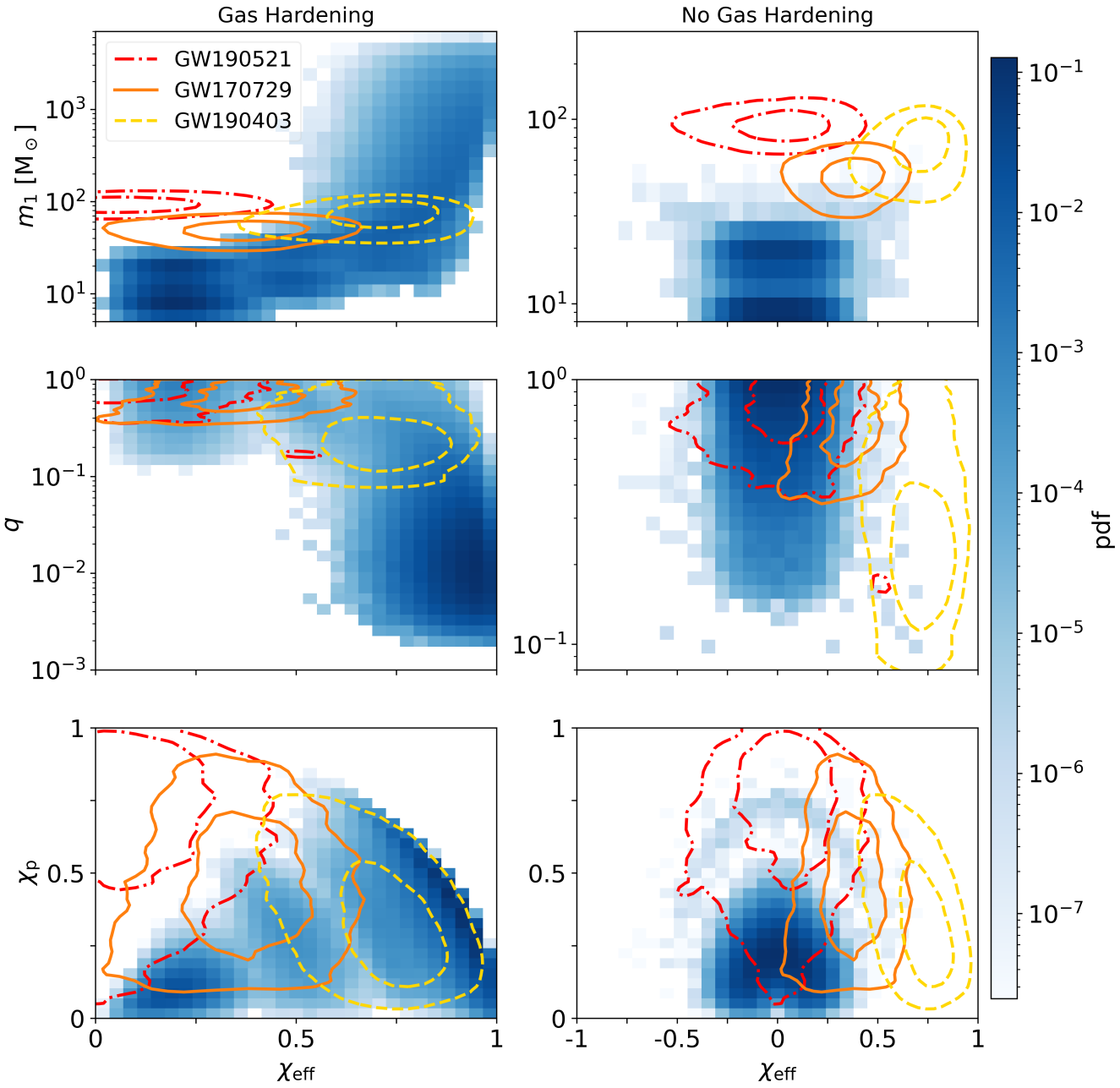

Fig. 9. Effective spin correlations with relevant BBH parameters. The top row shows the probability density distribution for the primary BH mass, m_1, and the effective spin, $\chi_{\rm eff}$, compared with posterior contour plots at credibility levels 50% and 90% of LVK BBH merger events GW190521 (dash-dotted red line, Abbott et al. 2020) and GW170729 (solid orange line, Abbott et al. 2023b), and LVK transient event GW190403_051519 (dashed yellow line, Abbott et al. 2021). The central (bottom) row has the same legend as the first row, but shows the mass ratio, $q = m_2/m_1$, (precession spin, $\chi_{\rm p}$) and the effective spin, $\chi_{\rm eff}$.

to the high number of Type I migrators embedded in the disk (e.g. Tagawa et al. 2020a,b). Interestingly, gas-assisted bulk-assembled BBHs migrate toward the migration trap while hardening (Tagawa et al. 2020b, Fig. 8). Hence, some binaries may merge in the migration trap even if they were assembled in the bulk.

4.3. Three-body encounters

We neglect three-body interactions and their effects on BH scattering and BBH hardening. We make this approximation because of the dearth of semi-analytical models for these interactions: in the literature, there are some works (e.g., Coleman Miller & Lauburg 2009; Fragione & Silk 2020) that assume an isotropic distribution of velocities and are appropriate for spherical star clusters. In a Keplerian disk geometry, the distribution of velocities is much different and these models are not appropriate (McKernan et al. 2022). Nevertheless, a hard binary hardens via binary–single encounters (Heggie 1975), and this has a twofold effect: the inspiral of BBHs is accelerated via three-body effects, and the third intruding body receives a recoil kick that may prevent it from reaching the migration trap. Also, binary-single interactions are expected to increase the eccentricity of all BBHs to $e\,(10\,{\rm Hz}) \geq 10^{-4}$ and potentially flip the

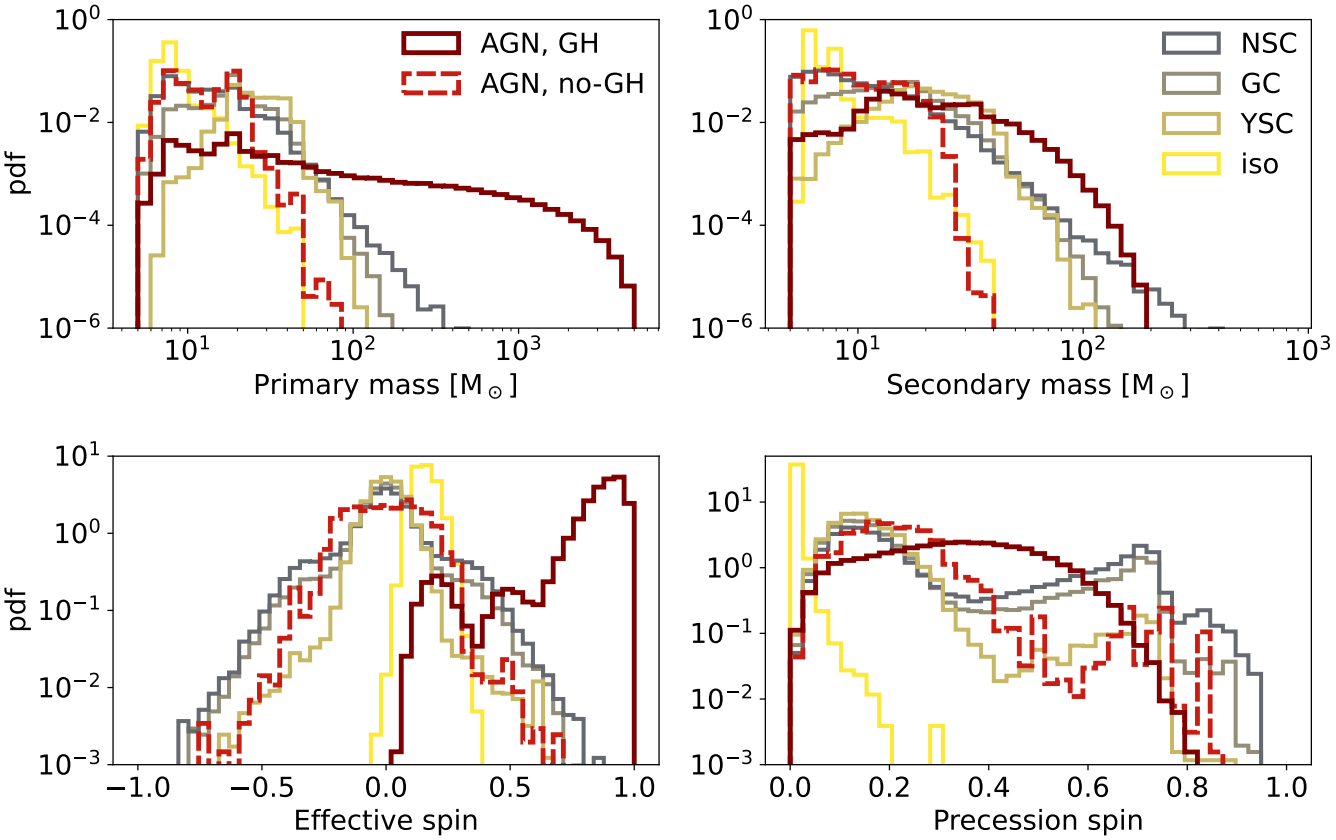

Fig. 10. Probability density distribution (pdf) of the primary and secondary BH masses, $m_{1,2}$, the effective spin, $\chi_{\rm eff}$, and the precession spin, $\chi_{\rm p}$, in the five channels modeled with FASTCLUSTER: active galactic nuclei (AGNs) with GH (solid dark red line) or with no-GH (dashed bright red line), NSCs (dark gray line), GCs (light gray line), YSCs (dark gold line), and isolated binary evolution (light gold line).

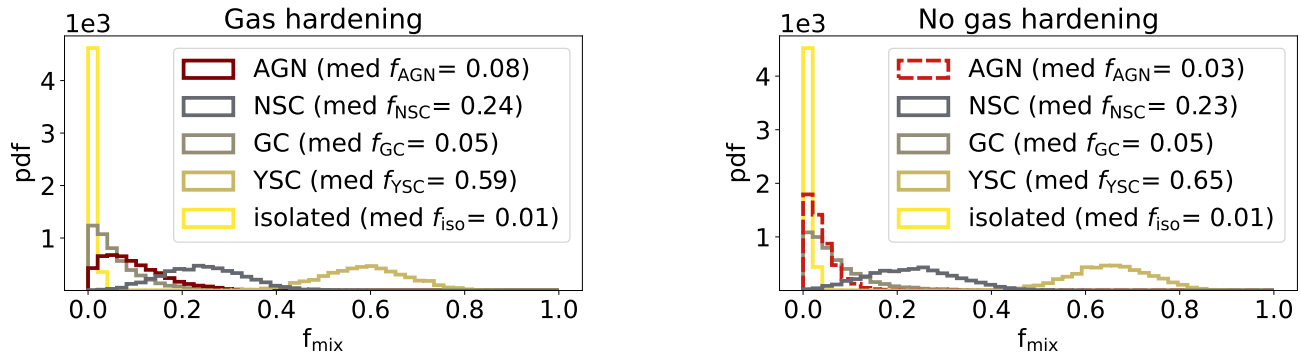

Fig. 11. Mixing fractions obtained with our five-channel analysis (see Appendix E). The left-hand (right-hand) plot shows the GH (no-GH) AGN channel model.

orientation of the BBH orbital angular momentum, leaving key signatures in the effective spin distribution (Samsing et al. 2022). Additionally, Rowan et al. (2023) show that the initial eccentricity of BBHs formed in AGN disks is high, which is not well modeled by our assumption of a thermal distribution (Jeans 1919).

Tagawa et al. (2020b) simulated the evolution of the compact object population in AGN disks using a one-dimensional N-body simulation combined with a semi-analytical model for the formation, disruption, and evolution of binaries. They include the effects of three-body interactions and employ a Thompson et al. (2005) disk model rather than the SG model that we use. Nevertheless, comparing our Fig. 6 with Tagawa et al. (2020b, Fig. 12a), we can point out that the output from our fiducial model is compatible with the output from theirs. This suggests that three-body effects have a relatively minor impact on the population of merging BBHs.

Leigh et al. (2017) find that the approximate encounter timescale between a BH in the migration trap and an intruder is a fraction of the Type I migration timescale (Eq. (10)) of an object of mass m_* starting its migration from the outer skirts of the disk, namely

$$t_{\rm enc} \simeq \frac{t_{\rm migr,\,I}\,(R_{\rm max})}{N_*} = \frac{1}{N_*}\left[\frac{M_{\rm SMBH}^2\,h^2\,(R_{\rm max})}{m_*\,\Sigma_{\rm gas}\,(R_{\rm max})\,R_{\rm max}^2\,\Omega_*}\right], \tag{27}$$

where N_* is the number of Type I migrators in the disk.

Assuming $m_* = 1\,M_\odot$ and $N_* = 100$, we find that the encounter timescale is typically larger than the delay timescale

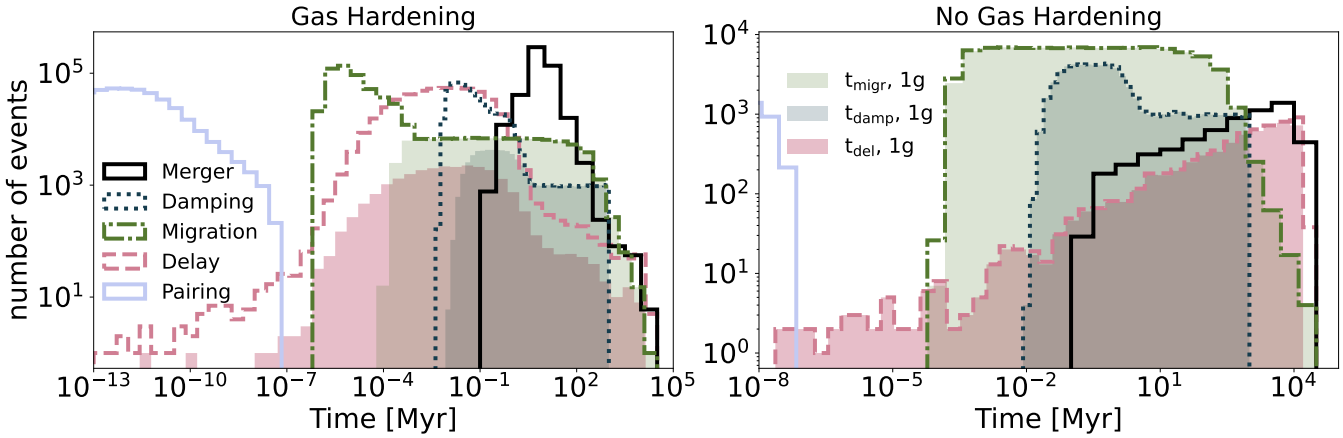

Fig. 12. Relevant timescales in our model. The black line shows the overall merger timescale, $t_{\rm merg}$. The dash-dotted green line shows the gas capture timescale, $t_{\rm damp}$, (the shaded green histogram shows $t_{\rm damp}$ for $1g$ BBHs only). The dotted navy line shoiws the Type I migration timescale, $t_{\rm migr,I}$, (the shaded navy histogram shows $t_{\rm migr,I}$ for $1g$ BBHs only). The dashed pink line shows the delay timescale, $t_{\rm del}$, between the pair-up and the merger (the shaded pink histogram shows $t_{\rm del}$ for $1g$ BBHs only). The light blue line shows the gas dynamical friction timescale, τ_0, related to BBH pairing (Qian et al. 2024).

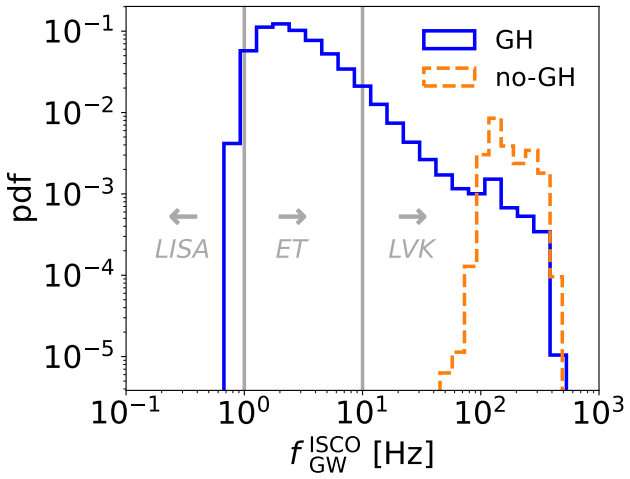

Fig. 13. Predictions for the *Einstein* Telescope (ET) and the Laser Interferometer Space Antenna (LISA). The solid blue (dashed orange) histogram shows the probability density functions (pdf) for ISCO emitted GW frequencies in the GH (no-GH) scenario. The gray lines show the approximate detectable frequencies of LVK detectors, ET, and LISA.

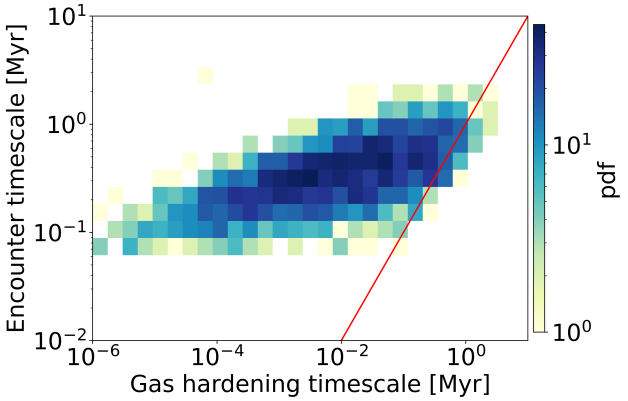

Fig. 14. Delay timescale, $t_{\rm del}$, in the GH model compared against the three-body encounter timescale, $t_{\rm enc}$, estimated as in Leigh et al. (2017). The red line is $t_{\rm del} = t_{\rm enc}$.

in the GH scenario (Fig. 14). This justifies our disregarding of three-body effects in this scenario. Nevertheless, they should be included in the no-GH model.

This implies that Heggie's law (employed in Eq. (16)) may not be valid in the GH case, as even soft binaries may be hardened considerably on a timescale shorter than the one on three-body interactions. We find that forming soft BBHs in our model is an extremely rare event (fewer than 1 in 10 000 BBHs), so including them in our computation would not have affected the results significantly.

Furthermore, efficient damping, migration, and BBH pair-up in AGN disks are expected to give rise to a high concentration of BBHs in migration traps. Hence, we should take into account the effect of binary-binary interactions. Such encounters are expected to efficiently ionize one of the binaries involved and lead to the formation of a stable triple (Marín Pina & Gieles 2024).

4.4. Position of migration traps

The position of migration traps is highly uncertain: it strongly depends on the disk model and the migration prescription used. For instance, Pan & Yang (2021) compute Type I migration torques in the locally isothermal approximation as well as torques due to winds and to the gravitational interaction with the SMBH for three different disk models. In their approximation, no disk model develops migration traps. Moreover, they find that in the inner part of the disk ($R \lesssim 10^2 R_{\rm g}$) Type I migration torques are always overpowered by the other effects, which prompt embedded BHs to eventually merge with the SMBH. This justifies our assumption that we can neglect the inner migration trap for the SG model (Sect. 2.1), but it implies that we somewhat overestimate the number of Type I migrators.

In their recent work, Grishin et al. (2024) account for both Type I migration, caused by the gravitational perturbation of embedded BHs in the disk, and thermal migration, caused by the thermal response of the disk to the small and overdense accretion disks surrounding embedded BHs. They find that the resulting migration trap position is at much larger distances than previously identified by considering Type I migration only.

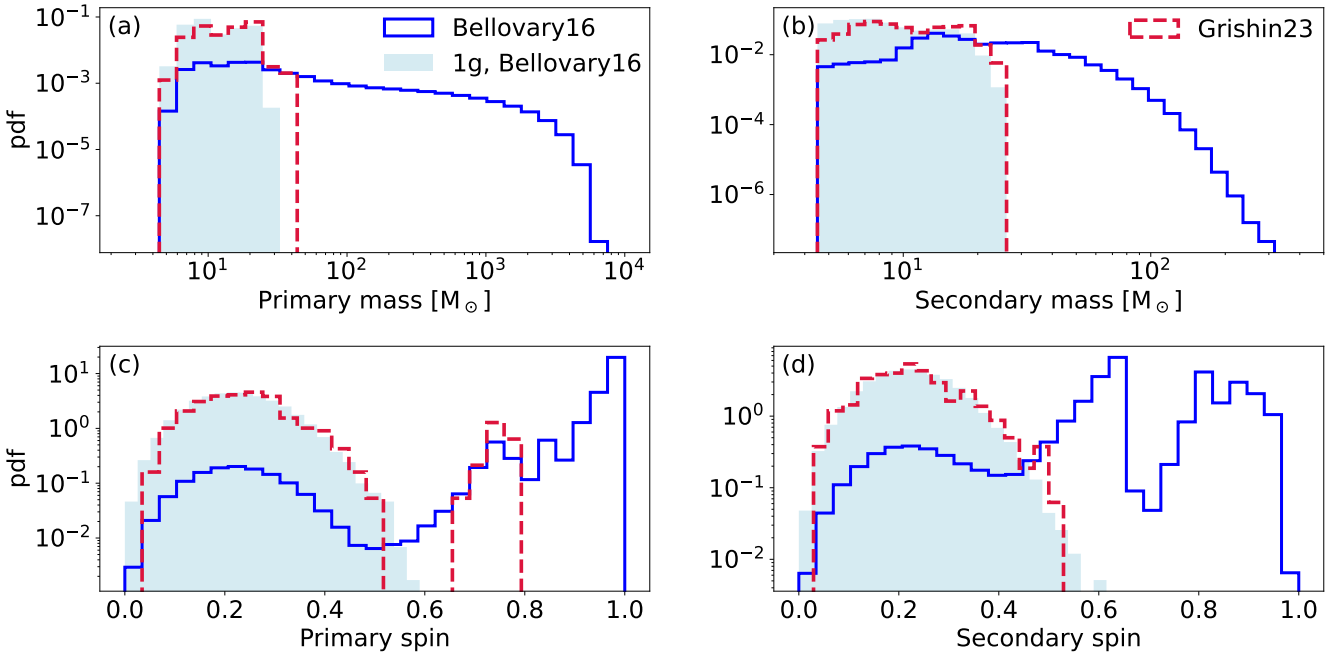

Fig. 15. Main properties of dynamical BBH mergers with two different models for the migration timescale and the location of migration traps. The solid blue line shows the fiducial GH model. The dashed red line shows the results of runs where we used the model by Grishin et al. (2024) for both the migration timescale and the location of the migration trap. Filled light blue histograms refer to $1g$ mergers in the fiducial model. Panels (a) and (b): primary and secondary BH masses. Panels (c) and (d): primary and secondary BH spin magnitudes.

In our fiducial model, we followed the prescription by Bellovary et al. (2016) and identified the position of the migration trap as $R_{\rm trap} = 1324\,R_{\rm g}$, whereas Grishin et al. (2024) find that the updated position of the migration trap is typically larger $R_{\rm trap} = 10^3$–$10^6\,R_{\rm g}$, and has a steep dependence on the SMBH mass as

$$\log\left(R_{\rm trap}/R_{\rm g}\right) \simeq -\log\left(M_{\rm SMBH}/M_{\odot}\right) + 11, \tag{28}$$

for $M_{\rm SMBH} \lesssim 10^8\,M_{\odot}$. Hence, for a typical SMBH of mass $10^6\,M_{\odot}$, the migration trap is at $\sim 10^5\,R_{\rm g}$.

Migration is also affected by the additional thermal torque, which reduces the migration timescale by a factor of

$$t_{\rm migr}^{\rm tot} \simeq \frac{h}{2}\,t_{\rm migr,\,I}, \tag{29}$$

where $t_{\rm migr,\,I}$ is defined as in Eq. (10).

Figure 15 shows a comparison between the results of our fiducial model and the results we obtain assuming the position of migration traps and the migration timescales predicted by Grishin et al. (2024). We find that the hierarchical merger process is suppressed when including a treatment for thermal torques, as only a handful of seed BHs reach the second generation in our simulations. This happens because the gaseous disk is significantly thicker and more diluted in the outer area of the disk where the new migration trap is located (see Sect. 2.1). Therefore, as the migration timescale is proportional to h^3 and $\Sigma_{\rm gas}^{-1}$, the migration process is much slower in this scenario. Moreover, even when BHs successfully reach the migration trap, the delay timescale ends up being too long because of the low gas surface density.

4.5. General caveats

In our model, we disregard dynamical interactions with the SMBH and gravitational perturbations caused by intermediate-mass BHs. For example, Deme et al. (2020) show that the presence of intermediate-mass BHs in the disk may enhance the ionization of BBHs and consequently decrease the merger rate.

Moreover, in Sect. 2.3.1 we assumed that eccentricity and inclination of a BH orbit are damped on similar timescales due to gas torques. Wang et al. (2024, Fig. 4) show that this is not always accurate: if the initial orbit is highly eccentric ($e \gtrsim 0.9$), the eccentricity-damping timescale can be up to five times longer than the inclination-damping one. A BH on a gas-embedded eccentric orbit ($i \sim 0$, $e \gtrsim h$) would be subject to spin-down (McKernan & Ford 2023), which would leave an imprint on the expected spin distribution.

Finally, we ignored the evolution of AGN disks in time. This can happen slowly over the AGN lifetime as BHs are embedded in the disk (Tagawa et al. 2022) and gas is accreted by the SMBH. The efficiency of all dynamical processes strongly depends on disk density and aspect ratio. If these quantities evolve over time, the resulting BBH population will be affected as well.

4.6. GW190521 and other transient events

It has previously been suggested that the transient events GW170729 and GW190521 could have originated in AGN disks (Yang et al. 2019a; Tagawa et al. 2021; Samsing et al. 2022; Graham et al. 2020; Morton et al. 2023). In particular, Yang et al. (2019a) found that it is five times more likely for GW170729 to arise from hierarchical mergers in AGNs than assuming that all events in the GWTC-1 catalog (Abbott et al. 2019) arise from the same channel, whereas Morton et al. (2023) show that the association between GW190521 and the AGN flare ZTF19abanrhr (Graham et al. 2020) is highly preferred over the lack of association, suggesting that the GW transient event was generated in an AGN.

We compare the posterior contours for the primary mass and effective spins of such events with the output of our AGN simulations, as shown in Fig. 9. We also include the posterior contour of the transient event GW190403_051519, although it has a low S/N of $7.6^{+0.6}_{-1.1}$ and a low probability of having an astrophysical origin, $p_{\rm astro} = 0.61$ (Abbott et al. 2021).

We find that the contours of GW170729 have some overlap with both the GH and the no-GH AGN models, and thus might be compatible with an origin in an AGN environment. In contrast, GW190521 has no significant overlap with our models in the $\chi_{\rm eff} - m_1$ space, and negligible overlap with our no-GH model in the $\chi_{\rm eff} - \chi_{\rm p}$ space. Indeed, the posterior probability distribution of GW190521 shows moderate support for high precession spin and low effective spin, suggesting that its BH spin vectors are large but misaligned. In contrast, in our GH model, we predict that they should be aligned for a BBH of primary mass $\sim 100\,M_\odot$. This does not rule out the hypothesis of an AGN origin for GW190521, as we might speculate that the BBH may have been perturbed by three-body encounters, which we do not account for in our GH model (Appendix B). This would alter the orientation of the orbital angular momentum and decrease the effective spin, as well as increase the BBH eccentricity (Samsing et al. 2022).

Finally, the posterior contours of GW190403_051519 significantly overlap with our GH AGN model. However, this event candidate has a high false alarm rate and may not have astrophysical origin (Abbott et al. 2021).

5. Summary

We explored the formation of BBH mergers in AGNs employing a new semi-analytical model. our key findings are as follows:

- The presence of GH significantly increases the efficiency of hierarchical mergers in AGN disks, allowing seed BHs to go through up to 500 merger episodes. This leads to the formation of BBHs with high masses (up to a few thousand solar masses) and low mass ratios ($q \simeq 10^{-2}$).
- In contrast, if GH is not efficient, the hierarchical merger chain is truncated after a few generations, leaving a BBH population with no components more massive than $\sim 10^2\,M_\odot$.
- The distribution of spin tilt angles is influenced by GH. There is a preferential alignment of spins in the GH scenario and isotropic distribution in the no-GH scenario. Hence, GH leads to distinct features in the effective spin distributions: a main peak on $\chi_{\rm eff} \simeq 1$ corresponding to maximum alignment and smaller peaks at $\chi_{\rm eff} \simeq 0.2$ and $\chi_{\rm eff} \simeq 0.5$.
- In the GH scenario, we find an anticorrelation between q and $\chi_{\rm eff}$ that might even extend to lower values of q and higher values of $\chi_{\rm eff}$ than currently observed by LVK.
- Comparison with other formation channels shows that AGN-driven mergers in the GH scenario result in higher primary BH masses and higher effective spins.

In summary, we find that efficient GH in AGN disks enhances the formation of BBH mergers with high primary masses and low mass ratios, and leads to a strong preference for $\chi_{\rm eff} \approx 1$. Given the high masses of BHs, next-generation ground-based detectors such as the ET are an ideal test bed for such unique features.

Acknowledgements. We thank the anonymous referee for their constructive and insightful comments that improved this manuscript. MM, MPV, CP, and ST acknowledge financial support from the European Research Council for the ERC Consolidator grant DEMOBLACK, under contract no. 770017. MPV and MM acknowledge financial support from the German Excellence Strategy via the Heidelberg Cluster of Excellence (EXC 2181 – 390900948) STRUCTURES. MD acknowledges financial support from the Cariparo Foundation under grant 55440. We thank Tamara Bogdanović, Barry McKernan, Dominika Wylezalek, Jenny Greene, Carolin "Lina" Kimmig, Ralf Klessen, Gastón Javier Escobar, Stefano Rinaldi, Giorgio Mentasti, Jacopo Tissino, Imre Bartos, Bence Kocsis, K.E. Saavik Ford, and Hiromichi Tagawa for their useful input. We thank Dylan Nelson for granting us access to the ILLUSTRISTNG JupyterLab workspace.

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102
Abbott, B. P., et al. (LIGO Scientific Collaboration and Virgo Collaboration) 2019, Phys. Rev. X, 9, 031040
Abbott, R., Abbott, T. D., Abraham, S., et al. 2020, Phys. Rev. Lett., 125, 101102
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021, ArXiv e-prints [arXiv:2108.01045]
Abbott, R., Abbott, T. D., Acernese, F., et al. 2023a, Phys. Rev. X, 13, 041039
Abbott, R., Abbott, T. D., Acernese, F., et al. 2023b, Phys. Rev. X, 13, 011048
Amaro Seoane, P., Aoudia, S., Audley, H., et al. 2013, ArXiv e-prints [arXiv:1305.5720]
Amaro Seoane, P., Audley, H., Babak, S., et al. 2017, ArXiv e-prints [arXiv:1702.00786]
Antonini, F., & Gieles, M. 2020, MNRAS, 492, 2936
Antonini, F., & Rasio, F. A. 2016, ApJ, 831, 187
Antonini, F., Gieles, M., & Gualandris, A. 2019, MNRAS, 486, 5008
Antonini, F., Gieles, M., Dosopoulou, F., & Chattopadhyay, D. 2023, MNRAS, 522, 466
Arca Sedda, M. 2020, ApJ, 891, 47
Arca Sedda, M., & Gualandris, A. 2018, MNRAS, 477, 4423
Askar, A., Szkudlarek, M., Gondek-Rosińska, D., Giersz, M., & Bulik, T. 2017, MNRAS, 464, L36
Atallah, D., Trani, A. A., Kremer, K., et al. 2023, MNRAS, 523, 4227
Ballone, A., Costa, G., Mapelli, M., et al. 2023, MNRAS, 519, 5191
Banerjee, S. 2017a, MNRAS, 467, 524
Banerjee, S. 2017b, MNRAS, 473, 909
Banerjee, S. 2020, MNRAS, 500, 3002
Banerjee, S. 2022, A&A, 665, A20
Bartos, I., Kocsis, B., Haiman, Z., & Márka, S. 2017, ApJ, 835, 165
Baruteau, C., Crida, A., Paardekooper, S. J., et al. 2014, Protostars and Planets VI, 667
Belczynski, K. 2020, ApJ, 905, L15
Belczynski, K., Heger, A., Gladysz, W., et al. 2016, A&A, 594, A97
Bellovary, J. M., Low, M. M., McKernan, B., & Saavik Ford, K. E. 2016, ApJ, 819, L17
Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Priceton University Press)
Bogdanović, T., Reynolds, C. S., & Coleman Miller, M. 2007, ApJ, 661, L147
Bouffanais, Y., Mapelli, M., Santoliquido, F., et al. 2021, MNRAS, 507, 5224
Branchesi, M., Maggiore, M., Alonso, D., et al. 2023, JCAP, 2023, 068
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Bryden, G., Chen, X., Lin, D. N. C., Nelson, R. P., & Papaloizou, J. C. B. 1999, ApJ, 514, 344
Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89
Calderón Bustillo, J., Leong, S. H. W., Chandra, K., McKernan, B., & Ford, K. E. S. 2021, ArXiv e-prints [arXiv:2112.12481]
Callister, T. A., Haster, C.-J., Ng, K. K. Y., Vitale, S., & Farr, W. M. 2021, ApJ, 922, L5
Cantiello, M., Jermyn, A. S., & Lin, D. N. C. 2021, ApJ, 910, 94
Chattopadhyay, D., Stegmann, J., Antonini, F., Barber, J., & Romero-Shaw, I. M. 2023, MNRAS, 526, 4908
Chen, Y.-X., & Lin, D. N. C. 2023, MNRAS, 522, 319
Cimerman, N. P., & Rafikov, R. R. 2024, MNRAS, 528, 2358
Coleman Miller, M., & Lauburg, V. M. 2009, ApJ, 692, 917
Costa, G., Girardi, L., Bressan, A., et al. 2019, MNRAS, 485, 4641
Costa, G., Bressan, A., Mapelli, M., et al. 2021, MNRAS, 501, 4514
Costa, G., Ballone, A., Mapelli, M., & Bressan, A. 2022, MNRAS, 516, 1072
Coughlin, M. W., Dietrich, T., Antier, S., et al. 2020, MNRAS, 497, 1181
Cresswell, P., Dirksen, G., Kley, W., & Nelson, R. P. 2007, A&A, 473, 329
Dall'Amico, M., Mapelli, M., Di Carlo, U. N., et al. 2021, MNRAS, 508, 3045
DeLaurentiis, S., Epstein-Martin, M., & Haiman, Z. 2023, MNRAS, 523, 1126
Deme, B., Meiron, Y., & Kocsis, B. 2020, ApJ, 892, 130
Di Carlo, U. N., Giacobbo, N., Mapelli, M., et al. 2019, MNRAS, 487, 2947
Di Carlo, U. N., Mapelli, M., Giacobbo, N., et al. 2020, MNRAS, 498, 495
Di Matteo, T., Colberg, J., Springel, V., Hernquist, L., & Sijacki, D. 2008, ApJ, 676, 33
Doctor, Z., Wysocki, D., O'Shaughnessy, R., Holz, D. E., & Farr, B. 2020, ApJ, 893, 35

Downing, J. M. B., Benacquista, M. J., Giersz, M., & Spurzem, R. 2010, MNRAS, 407, 1946
Evans, M., Adhikari, R. X., Afle, C., et al. 2021, ArXiv e-prints [arXiv:2109.09882]
Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., & Justham, S. 2019, ApJ, 887, 53
Farmer, R., Renzo, M., de Mink, S. E., Fishbach, M., & Justham, S. 2020, ApJ, 902, L36
Farrell, E., Groh, J. H., Hirschi, R., et al. 2021, MNRAS, 502, L40
Fishbach, M., Holz, D. E., & Farr, B. 2017, ApJ, 840, L24
Fishbach, M., Holz, D. E., & Farr, W. M. 2018, ApJ, 863, L41
Flitter, J., Muñoz, J. B., & Kovetz, E. D. 2021, MNRAS, 507, 743
Ford, K. E. S., & McKernan, B. 2022, MNRAS, 517, 5827
Fragione, G., & Kocsis, B. 2018, Phys. Rev. Lett., 121, 161103a
Fragione, G., & Silk, J. 2020, MNRAS, 498, 4591
Fryer, C. L., Belczynski, K., Wiktorowicz, G., et al. 2012, ApJ, 749, 91
Gallazzi, A., Brinchmann, J., Charlot, S., & White, S. D. M. 2008, MNRAS, 383, 1439
Gayathri, V., Yang, Y., Tagawa, H., Haiman, Z., & Bartos, I. 2021, ApJ, 920, L42
Gerosa, D., & Berti, E. 2017, Phys. Rev. D, 95, 124046
Giacobbo, N., & Mapelli, M. 2019, MNRAS, 482, 2234
Goodman, J. 2003, MNRAS, 339, 937
Graham, A. W. 2022, MNRAS, 518, 6293
Graham, A. W., & Spitler, L. R. 2009, MNRAS, 397, 2148
Graham, M. J., Ford, K. E. S., McKernan, B., et al. 2020, Phys. Rev. Lett., 124, 251102a
Greene, J. E., & Ho, L. C. 2007, ApJ, 667, 131
Greiner, J., Burgess, J. M., Savchenko, V., & Yu, H. F. 2016, ApJ, 827, L38
Grishin, E., Gilbaum, S., & Stone, N. C. 2024, MNRAS, 530, 2114
Heggie, D. C. 1975, MNRAS, 173, 729
Hohmann, W. 1960, The Attainability of Heavenly Bodies, NASA Technical Translation (National Aeronautics and Space Administration)
Iorio, G., Mapelli, M., Costa, G., et al. 2023, MNRAS, 524, 426
Ishibashi, W., & Gröbner, M. 2020, A&A, 639, A108
Jeans, J. H. 1919, MNRAS, 79, 408
Jiménez-Forteza, X., Keitel, D., Husa, S., et al. 2017, Phys. Rev. D, 95, 064024
Kaaz, N., Schrøder, S. L., Andrews, J. J., Antoni, A., & Ramirez-Ruiz, E. 2023, ApJ, 944, 44
Kawamura, S., Nakamura, T., Ando, M., et al. 2019, Int. J. Mod. Phys. D, 28, 1845001
Khrykin, I. S., Hennawi, J. F., Worseck, G., & Davies, F. B. 2021, MNRAS, 505, 649
Kimball, C., Talbot, C., Berry, C. P. L., et al. 2020, ApJ, 900, 177
King, A. R., Lubow, S. H., Ogilvie, G. I., & Pringle, J. E. 2005, MNRAS, 363, 49
King, A. R., Pringle, J. E., & Livio, M. 2007, MNRAS, 376, 1740
Kremer, K., Spera, M., Becker, D., et al. 2020, ApJ, 903, 45
Kroupa, P. 2001, MNRAS, 322, 231
Kumamoto, J., Fujii, M. S., & Tanikawa, A. 2019, MNRAS, 486, 3942
Kumamoto, J., Fujii, M. S., & Tanikawa, A. 2020, MNRAS, 495, 4268
Leigh, N. W. C., Geller, A. M., McKernan, B., et al. 2017, MNRAS, 474, 5672
Li, J., Dempsey, A. M., Li, H., Lai, D., & Li, S. 2023, ApJ, 944, L42
Lubow, S. H., Martin, R. G., & Nixon, C. 2015, ApJ, 800, 96
Luo, J., Chen, L.-S., Duan, H.-Z., et al. 2016, CQG, 33, 035010
Lyra, W., Paardekooper, S.-J., & Low, M.-M. M. 2010, ApJ, 715, L68
MacFadyen, A. I., & Milosavljević, M. 2008, ApJ, 672, 83
Maggiore, M. 2008, GravitationalWaves: Volume 1: Theory and Experiments, Gravitational Waves (Oxford University Press)
Maggiore, M. 2018, Gravitational Waves: Volume 2: Astrophysics and Cosmology, Gravitational Waves (Oxford University Press)
Maggiore, M., Broeck, C. V. D., Bartolo, N., et al. 2020, JCAP, 2020, 050
Mapelli, M. 2016, MNRAS, 459, 3432
Mapelli, M. 2021, Handbook of Gravitational Wave Astronomy (Springer), 16
Mapelli, M., Spera, M., Montanari, E., et al. 2020, ApJ, 888, 76
Mapelli, M., Dall'Amico, M., Bouffanais, Y., et al. 2021, MNRAS, 505, 339
Mapelli, M., Bouffanais, Y., Santoliquido, F., Arca Sedda, M., & Artale, M. C. 2022, MNRAS, 511, 5797
Marchant, P., & Moriya, T. J. 2020, A&A, 640, L18
Marín Pina, D., & Gieles, M. 2024, MNRAS, 527, 8369
McKernan, B., & Ford, K. E. S. 2023, MNRAS, submitted [arXiv:2309.15213]
McKernan, B., Ford, K. E. S., Lyra, W., & Perets, H. B. 2012, MNRAS, 425, 460
McKernan, B., Ford, K. E. S., Kocsis, B., Lyra, W., & Winter, L. M. 2014, MNRAS, 441, 900
McKernan, B., Ford, K. E. S., O'Shaugnessy, R., & Wysocki, D. 2020, MNRAS, 494, 1203
McKernan, B., Ford, K. E. S., Callister, T., et al. 2022, MNRAS, 514, 3886
Méndez, E. M., Colle, F. D., López-Cámara, D., & Vigna-Gómez, A. 2023, MNRAS, 522, 1686
Merritt, D., & Ferrarese, L. 2001, ApJ, 547, 140
Miller, M. C., & Hamilton, D. P. 2002, MNRAS, 330, 232
Miller, M. C., & Lauburg, V. M. 2009, ApJ, 692, 917
Miranda, R., Muñoz, D. J., & Lai, D. 2016, MNRAS, 466, 1170
Morton, S. L., Rinaldi, S., Torres-Orjuela, A., et al. 2023, Phys. Rev. D, 108, 123039
Netzer, H. 2015, ARA&A, 53, 365
Neumayer, N., Seth, A., & Böker, T. 2020, A&A Rev., 28, 4
Nguyen, C. T., Costa, G., Girardi, L., et al. 2022, A&A, 665, A126
O'Brien, B., Szczepańczyk, M., Gayathri, V., et al. 2021, Phys. Rev., D, 104
O'Leary, R. M., Kocsis, B., & Loeb, A. 2009, MNRAS, 395, 2127
O'Leary, R. M., Meiron, Y., & Kocsis, B. 2016, ApJ, 824, L12
Ostriker, E. C. 1999, ApJ, 513, 252
Özel, F., Psaltis, D., Narayan, R., & McClintock, J. E. 2010, ApJ, 725, 1918
Pan, Z., & Yang, H. 2021, Phys. Rev. D, 103, 103018
Peters, P. C. 1964, Phys. Rev., 136, B1224
Petrovich, C., & Antonini, F. 2017, ApJ, 846, 146
Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648
Plummer, H. C. 1911, MNRAS, 71, 460
Portegies Zwart, S. F., & McMillan, S. L. W. 2000, ApJ, 528, L17
Punturo, M., Abernathy, M., Acernese, F., et al. 2010, CQG, 27, 194002
Qian, K., Li, J., & Lai, D. 2024, ApJ, 962, 143
Quinlan, G. D., & Shapiro, S. L. 1990, ApJ, 356, 483
Renzo, M., Cantiello, M., Metzger, B. D., & Jiang, Y. F. 2020, ApJ, 904, L13
Robson, T., Cornish, N. J., & Liu, C. 2019, CQG, 36, 105011
Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, Phys. Rev. Lett., 115, 051101
Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, Phys. Rev. D, 93, 084029
Rodriguez, C. L., Amaro-Seoane, P., Chatterjee, S., et al. 2018, Phys. Rev. D, 98, 123005
Rodriguez, C. L., Zevin, M., Amaro-Seoane, P., et al. 2019, Phys. Rev. D, 100, 043027
Rom, B., Sari, R., & Lai, D. 2024, ApJ, 964, 43
Romero-Shaw, I., Lasky, P. D., & Thrane, E. 2021, ApJ, 921, L31
Rowan, C., Boekholt, T., Kocsis, B., & Haiman, Z. 2023, MNRAS, 524, 2770
Rowan, C., Whitehead, H., Boekholt, T., Kocsis, B., & Haiman, Z. 2024, MNRAS, 527, 10448
Sabhahit, G. N., Vink, J. S., Sander, A. A. C., & Higgins, E. R. 2023, MNRAS, 524, 1529
Sahu, N., Graham, A. W., & Davis, B. L. 2019, ApJ, 887, 10
Samsing, J., Bartos, I., D'Orazio, D. J., et al. 2022, Nature, 603, 237
Santini, A., Gerosa, D., Cotesta, R., & Berti, E. 2023, Phys. Rev. D, 108, 083033
Santoliquido, F., Mapelli, M., Bouffanais, Y., et al. 2020, ApJ, 898, 152
Santoliquido, F., Mapelli, M., Giacobbo, N., Bouffanais, Y., & Artale, M. C. 2021, MNRAS, 502, 4877
Scott, N., & Graham, A. W. 2013, ApJ, 763, 76
Secunda, A., Bellovary, J., Mac Low, M.-M., et al. 2020, ApJ, 903, 133
Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41
Siegel, D. M., Agarwal, A., Barnes, J., et al. 2022, ApJ, 941, 100
Sirko, E., & Goodman, J. 2003, MNRAS, 341, 501
Spera, M., & Mapelli, M. 2017, MNRAS, 470, 4739
Spera, M., Mapelli, M., Giacobbo, N., et al. 2019, MNRAS, 485, 889
Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676
Stevenson, S., Sampson, M., Powell, J., et al. 2019, ApJ, 882, 121
Stone, N. C., Metzger, B. D., & Haiman, Z. 2017, MNRAS, 464, 946
Tagawa, H., Haiman, Z., Bartos, I., & Kocsis, B. 2020a, ApJ, 899, 26
Tagawa, H., Haiman, Z., & Kocsis, B. 2020b, ApJ, 898, 25
Tagawa, H., Kocsis, B., Haiman, Z., et al. 2021, ApJ, 908, 194
Tagawa, H., Kimura, S. S., Haiman, Z., et al. 2022, ApJ, 927, 41
Tagawa, H., Kimura, S. S., Haiman, Z., Perna, R., & Bartos, I. 2023, ApJ, 950, 13
Tanikawa, A., Kinugawa, T., Yoshida, T., Hijikawa, K., & Umeda, H. 2021, MNRAS, 505, 2170
Tanikawa, A., Moriya, T. J., Tominaga, N., & Yoshida, N. 2022, MNRAS, 519, L32
Thompson, T. A., Quataert, E., & Murray, N. 2005, ApJ, 630, 167
Torniamenti, S., Rastello, S., Mapelli, M., et al. 2022, MNRAS, 517, 2953
Umeda, H., & Nagele, C. 2024, ApJ, 961, 146
Vajpeyi, A., Thrane, E., Smith, R., McKernan, B., & Saavik Ford, K. E. 2022, ApJ, 931, 82
Veronesi, N., Rossi, E. M., & van Velzen, S. 2023, MNRAS, 526, 6031
Vink, J. S., Higgins, E. R., Sander, A. A. C., & Sabhahit, G. N. 2021, MNRAS, 504, 146
Wang, Y.-H., McKernan, B., Ford, S., et al. 2021, ApJ, 923, L23

Wang, J.-M., Zhai, S., Li, Y.-R., et al. 2023, ApJ, 954, 84
Wang, Y., Zhu, Z., & Lin, D. N. C. 2024, MNRAS, 528, 4958
Weinberger, R., Springel, V., Pakmor, R., et al. 2018, MNRAS, 479, 4056
Whitehead, H., Rowan, C., Boekholt, T., & Kocsis, B. 2023, MNRAS, submitted [arXiv:2309.11561]
Woosley, S. E., & Heger, A. 2014, Very Massive Stars in theLocal Universe (Springer International Publishing), 199
Woosley, S. E., & Heger, A. 2021, ApJ, 912, L31
Woosley, S. E., Heger, A., & Weaver, T. A. 2002, Rev. Mod. Phys., 74, 1015
Yang, Y., Bartos, I., Gayathri, V., et al. 2019a, Phys. Rev. Lett., 123, 181101
Yang, Y., Bartos, I., Haiman, Z., et al. 2019b, ApJ, 876, 122
Zevin, M., & Holz, D. E. 2022, ApJ, 935, L20
Zevin, M., Samsing, J., Rodriguez, C., Haster, C.-J., & Ramirez-Ruiz, E. 2019, ApJ, 871, 91
Zhou, Z.-H., & Wang, K. 2023, ApJ, 958, L12
Zhu, J.-P. 2024, MNRAS, 528, L88
Ziosi, B. M., Mapelli, M., Branchesi, M., & Tormen, G. 2014, MNRAS, 441, 3703

Appendix A: Catalogs for AGN disks, NSCs, GCs, YSCs, and isolated binaries

We generated catalogs of BH masses for AGNs, NSCs, GCs, YSCs, and isolated binaries with the SEVN binary population synthesis code (Iorio et al. 2023). For the isolated binary systems, we integrated the evolution of 7.5×10^7 binary systems divided into 15 metallicity bins ranging from $Z = 10^{-4}$ to 3×10^{-2}. The initial properties of such binary systems and the treatment of binary evolution processes (stable mass transfer, common envelope, tides, natal kicks, GW decay, etc.) are the same as in the fiducial model by Iorio et al. (2023). Also, we consider the rapid model for core-collapse supernovae (Fryer et al. 2012) and include a treatment for pair instability as described by Mapelli et al. (2020).

The masses of BHs in AGNs, NSCs, GCs, and YSCs are obtained from a single star population run with SEVN. In particular, for NSCs, GCs, and YSCs we simulated the same set of 15 different metallicities ranging from $Z = 10^{-4}$ to 3×10^{-2} as in the isolated binary case but we turn off binary evolution. These catalogs are the same ones we adopted for AGN BHs, but for the latter we only used the solar metallicity $Z = 0.02$, because the metallicity at the center of a massive galaxy tends to be solar or super-solar. In future work, we will also model the metallicity evolution in AGN disks.

The spin magnitudes of BHs in both isolated binaries and star clusters are drawn from a Maxwellian distribution with $\sigma_\chi = 0.1$, the same as for BHs in AGN disks. In the dynamical channel, spin orientations are assumed to be isotropic, because dynamical encounters randomize them with respect to the orbital plane; whereas for isolated binary systems we derive the final spin orientation accounting for the effect of natal kicks on the orbital plane (e.g., Mapelli 2021).

BHs in NSCs, GCs, and YSCs pair up dynamically, via three-body and binary-single encounters, as described by Mapelli et al. (2021). Here, we do not consider the contribution of primordial binaries to BBH mergers in NSCs, GCs, and YSCs (see, e.g., Mapelli et al. 2021, for a treatment of primordial binaries). Subsequently, dynamically formed BBHs evolve via three-body hardening and GW decay in their parent star clusters. When two BHs merge, we calculate the properties of the compact remnant (mass and spins) and its gravitational recoil. If the BH remnant remains in the star cluster, it can form a second-generation BBH and undergo further mergers (see Mapelli et al. 2021 for more details). We estimate the evolution of BBH mergers in isolated binaries, NSCs, GCs, and YSCs as described by Mapelli et al. (2022) and summarized in Appendix D.

Appendix B: Spin tilt

B.1. Gas-hardening scenario

Embedded objects can weakly perturb the surface-density profile of the AGN gaseous disk, resulting in gas torques that tend to align both the BH spin vectors, $\boldsymbol{\chi}_{1,2}$, and the binaries' orbital angular momentum vector, $\boldsymbol{L}$, with the angular momentum, $\boldsymbol{J}$, of the AGN disk (Lubow et al. 2015; Vajpeyi et al. 2022). According to Bogdanović et al. (2007), this process is particularly efficient if fully embedded BHs can accrete more than 1% of their mass due to gas accretion, so that $\Delta m_{\rm BH} \geq 0.01\, m_{\rm BH}$. Since we only consider BHs that are able to reach the migration trap and form a BBH, we estimate the mass variation of the binary as

$$\Delta m_{\rm BBH} = \Delta m_1 + \Delta m_2 \simeq \Sigma_{\rm gas}^{\rm trap} \pi a^2 \tag{B.1}$$

where $m_{1,2}$ are the primary and secondary BH masses, a is the BBH semimajor axis, and the superscript "trap" indicates that the quantities are computed at the radial location $R_{\rm trap}$. Then, we checked that

$$\Delta m_{\rm BBH} \geq 0.01\,(m_1 + m_2)\,. \tag{B.2}$$

This condition is always verified in our model since the migration trap is the location with the highest surface density $\Sigma_{\rm gas}$ in the disk. Hence, we modeled the spin tilt in our fiducial model as in the high-alignment scenario of Vajpeyi et al. (2022).

We sample $\cos(\theta_1)$ (θ_1 being the angle between $\boldsymbol{\chi}_1$ and $\boldsymbol{L}$) from a truncated Gaussian centered in 1 with standard deviation $\sigma_1 = 0.1$, so that the primary spin is aligned with the orbital angular momentum. We sample $\cos(\theta_2)$ (θ_2 being the angle between $\boldsymbol{\chi}_2$ and $\boldsymbol{L}$) from a truncated Gaussian centered in $\cos\theta_1$ with standard deviation $\sigma_2 = 0.1$, so that the spin of the secondary is aligned with that of the primary. This also implies $\boldsymbol{\chi}_2 \parallel \boldsymbol{L}$. For the azimuthal direction, we draw $\cos\phi_{1,2}$ from a uniform distribution.

We point out that alignment (and not anti-alignment) is efficient in AGNs because, as shown by King et al. (2005), counter-alignment is only possible if $J < 2\chi$, but in AGN disks J is typically very large because the gas is in Keplerian motion close to a massive SMBH, so we can safely assume that the counter-alignment condition is never met. Moreover, retrograde binaries (i.e. with $\boldsymbol{L} \cdot \boldsymbol{J} < 0$) are preferentially ionized or softened by tertiary encounters compared to prograde BBHs (Wang et al. 2021).

B.2. No-gas-hardening scenario

Gas torques are not the only physical phenomenon influencing BHs spin tilt: gas turbulence (Chen & Lin 2023) and three-body encounters of BBHs with other objects (Tagawa et al. 2020a) tend to randomize the alignment of $\boldsymbol{\chi}_{1,2}$ relative to $\boldsymbol{L}$. The competing effects of the gaseous disk and dynamical encounters on BBHs determine the distribution of BBH spin orientations. We neglect three-body encounters in our fiducial model because GH significantly speeds up BBH inspiral, so that the timescale $t_{\rm del}$ between the BBH pair-up and merger (Section 2.4) is typically shorter than the typical timescale for three-body encounters $t_{\rm 3bb}$ (which we estimate as in Leigh et al.(2017, eq. 19)). Instead, when we neglect GH, the two timescales become comparable to each other. We assume three-body encounters to be the dominant effect in this scenario, similarly to what has been shown by Tagawa et al. (2020a) who also neglect GH. In this model, we set the spin tilts $\theta_{1,2}$ from a Gaussian distribution (as stated earlier), but with standard deviation $\sigma_1 = \sigma_2 = 10$, as in the isotropic scenario of Vajpeyi et al. (2022). Indeed, setting $\sigma_i = 10$ is equivalent to sampling $\cos\theta_i$ from a uniform distribution in the interval $[0\,, 1]$ and coincides with the isotropic case. Once again, we draw $\cos\phi_{1,2}$ from uniform distributions.

Appendix C: Secondary BH mass

In our model, we allow for $Ng-Mg$ mergers, where N and M are the generation of the primary and secondary BH mass, respectively. Following Zevin & Holz (2022), we assume that the probability that a given generation M is chosen for the secondary is proportional to the number of $1g$ BHs required to produce it:

$$p\,(M) \propto 2^{-(M-1)} \tag{C.1}$$

for $M \leq N$, so that a $1g$ primary BH ($N = 1$) will necessarily pair up with a $1g$ secondary ($M = 1$). For $M = 1$, the secondary BH mass m_2 is randomly drawn with probability distribution (O'Leary et al. 2016)

$$p\left(m_2 \,|\, m_1\right) \propto (m_1 + m_2)^4 \tag{C.2}$$

between $m_2^{\min} = 5\,M_\odot$ and $m_2^{\max} = m_1$.

For $M > 1$, we determined the secondary mass as follows. First, we generated a $1g$ BH determining its mass and spin as in Eq. C.2. Then, we let it go through a certain number of $Ng - 1g$ mergers[7] until it created an Mg remnant.

At each step, the primary will be the remnant of a previous merger event and its mass, m_1, and spin, χ_1, are computed according to Jiménez-Forteza et al. (2017). The secondary BH mass is sampled from

$$p(m_2 \,|\, m_1) \propto (m_1 + m_2^{\max})^4 \tag{C.3}$$

between $m_2^{\min} = 5\,M_\odot$ and $m_2^{\max}$.

Here, $m_2^{\max}$ is determined based on the value of m_1. In particular, if m_1 is smaller than the maximum mass $m_1^{\max}$ of a $1g$ BH in the input sample coming from the population synthesis code SEVN, the secondary mass cannot exceed the mass of the primary: $m_2^{\max} = m_1$. Otherwise, we set $m_2^{\max} = m_1^{\max}$. This is a modification of Eq. (C.2).

With this choice, if the primary has mass compatible with a $1g$ BH ($m_1 \leq m_1^{\max}$), we sample the secondary in Eq. (C.2) (model from O'Leary et al. 2016).

Otherwise, if the primary is a higher-generation BH ($m_1 > m_1^{\max}$), we keep the same analytical description as in eq. C.2 but we force the secondary to have mass compatible with a $1g$ BH ($m_2 \leq m_1^{\max}$). This calculation is quite fast because we only compute the remnant mass and spin at each step, neglecting the merger times.

Appendix D: COSMO$\mathcal{R}$ATE

We interface the catalogs produced in our computation with the code COSMO$\mathcal{R}$ATE (Santoliquido et al. 2020, 2021), which calculates the BBH merger rate evolution $\mathcal{R}(z)$ by using catalogs of BBH mergers simulated with FASTCLUSTER and by coupling them with the cosmic star formation rate $\psi(z)$ and metallicity evolution as

$$\mathcal{R}(z) = \int_{z_{\max}}^{z} \psi(z') \frac{\mathrm{d}t(z')}{\mathrm{d}z'} \left[\int_{Z_{\min}}^{Z_{\max}} \eta(Z)\, \mathcal{F}(z', z, Z)\, \mathrm{d}Z \right] \mathrm{d}z', \tag{D.1}$$

where $t(z')$ is the lookback-time at redshift z' and $\mathrm{d}t(z')/\mathrm{d}z' = [(1+z')\,H(z')]^{-1}$, with $H(z') = H_0 \left[(1+z')^3\,\Omega_{\mathrm{M}} + \Omega_\Lambda\right]^{1/2}$.

For the AGN channel, we assume solar metallicity $Z = 0.02$. We compute the fraction $\mathcal{F}(z', z, Z)$ of BBH that form at redshift z' and merge at redshift z by interpolating the density $n_{\mathrm{AGN}}(z')$ of active SMBHs with $\dot{M} \geq 0.2\dot{M}_{\mathrm{Edd}}$, shown in Fig. 2, and keeping into account the merger timescale t_{merg}. We compute the merger efficiency $\eta(Z)$ as the ratio between the total number of BBH mergers with merger timescale lower than the Hubble timescale ($t_{\mathrm{merg}} \leq 13.4\,\mathrm{Gyr}$) and the total simulated stellar mass (Eq. 6).

We repeat the computation for other FASTCLUSTER channels as illustrated in Mapelli et al. (2022). We assume metallicity spread $\sigma_Z = 0.3$.

The COSMO$\mathcal{R}$ATE algorithm can produce catalogs of BBH mergers at different redshifts. In Section 3.3 we compare the outputs of five channels at low redshift $z \leq 0.1$. The full five-channel catalog up to redshift $z = 15$ is used for the computation of mixing fractions illustrated in Appendix E.

Appendix E: Mixing fractions

We compare our models against the 56 high-purity GW events analyzed by Abbott et al. (2023b) using a hierarchical Bayesian approach as in Mapelli et al. (2022). We shortly describe the process here for the reader's comfort.

Given a number N_{obs} of GW observations $\mathcal{H} = \{h_k\}_{k=1}^{N_{\mathrm{obs}}}$ described by an ensemble of parameters θ, the posterior distribution of the hyper-parameters λ associated with the models is

$$p\left(\lambda, N_\lambda \,|\, \mathcal{H}\right) = e^{-\mu_\lambda} \pi(\lambda, N_\lambda) \prod_{k=1}^{N_{\mathrm{obs}}} N_\lambda \int_\theta \mathcal{L}^k(h^k \,|\, \theta) p(\theta \,|\, \lambda)\, d\theta, \tag{E.1}$$

where $\theta = \left\{M_{\mathrm{chirp}}, m_1 + m_2, \chi_{\mathrm{eff}}, \chi_{\mathrm{p}}, z\right\}$ are the GW parameters, N_λ is the number of events predicted by the astrophysical model, μ_λ is the predicted number of detections associated with the model and the GW detector, $\pi(\lambda, N_\lambda)$ is the prior distribution on λ and N_λ, and $\mathcal{L}^k(h^k \,|\, \theta)$ is the likelihood of the k-th detection. The predicted number of detections is given by

$$\mu_\lambda = N_\lambda \int_\theta p(\theta \,|\, \lambda) p_{\mathrm{det}}(\theta)\, d\theta, \tag{E.2}$$

where p_{det} is the probability of detecting a source with parameters θ and can be inferred by computing the optimal signal-to-noise ratio and comparing it to a detection threshold, as described in Bouffanais et al. (2021). Also, we marginalize over N_λ using a prior $\pi(N_\lambda) \sim 1/N_\lambda$ (Fishbach et al. 2018). This implies that our mixing fractions do not depend on the merger rate of each channel.

We computed the mixing fractions $\{f_{\mathrm{iso}}, f_{\mathrm{AGN}}, f_{\mathrm{GC}}, f_{\mathrm{YSC}}, f_{\mathrm{NSC}}\}$ which weight the contributions of our five channels to the overall distribution (i = isolated, AGN, NSC, GC, YSC):

$$p(\theta \,|\, \lambda) = \sum_{\mathrm{i}} f_{\mathrm{i}}\, p(\theta \,|\, \mathrm{i}, \lambda), \tag{E.3}$$

The mixing fractions are defined such that

$$f_{\mathrm{iso}} + f_{\mathrm{AGN}} + f_{\mathrm{GC}} + f_{\mathrm{YSC}} + f_{\mathrm{NSC}} = 1. \tag{E.4}$$

Based on this definition, the mixing fraction for each channel is approximately the fraction of merger events associated with that specific channel. This definition of the mixing fraction assumes that all GWTC-3 events originate from the five channels we considered here, so we are neglecting the possibilities of BBH mergers from primordial BHs, triples, and multiples as well as any other possible evolution channel.

[7] We chose to consider only $Ng - 1g$ mergers rather than $Ng - Mg$ for simplicity.