Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **544,** 4576–4589 (2025) https://doi.org/10.1093/mnras/staf1896
Advance Access publication 2025 November 4

Black hole merger rates in AGN: contribution from gas-captured binaries

Connar Rowan [1,2★] Henry Whitehead [3] and Bence Kocsis [1,4]

[1]*Rudolf Peierls Centre for Theoretical Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK*
[2]*Niels Bohr International Academy, Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark*
[3]*Department of Physics, Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[4]*St Hugh's College, St Margaret's Rd, Oxford OX2 6LE, UK*

Accepted 2025 October 27. Received 2025 September 19; in original form 2025 January 17

ABSTRACT

Merging black hole (BH) binaries in active galactic nucleus (AGN) discs formed through two-body scatterings via the 'gas-capture' process may explain a significant fraction of BH mergers in AGN and a non-negligible contribution to the observed rate from LIGO-VIRGO-KAGRA. We perform Monte Carlo simulations of binary BH formation, evolution, and mergers across the observed AGN mass function using a novel physically motivated treatment for the gas-capture process derived from hydrodynamical simulations of BH–BH encounters in AGN. Our models suggest that gas-captured binaries could result in merger rates of $0.73 - 7.1\,\mathrm{Gpc^{-3}\,yr^{-1}}$. Mergers from AGN are dominated by AGN with supermassive BH masses of $\sim 10^7\,\mathrm{M_\odot}$, with 90 per cent of mergers occurring in the range $\sim 10^6\,\mathrm{M_\odot} - 10^8\,\mathrm{M_\odot}$. The merging mass distribution is flatter than the initial BH mass power law by a factor $\Delta\xi = 1.1 - 1.2$, as larger BHs align with the disc and form binaries more efficiently. Similarly, the merging mass ratio distribution is flatter therefore the AGN channel could explain high mass and unequal mass ratio detections such as GW190521 and GW190814. Using a simpler dynamical friction treatment for the binary formation process, the results are similar, where the primary bottleneck is the alignment time with the disc. The most influential parameters are the anticipated number of BHs and their mass function. Given the many uncertainties that remain in the AGN channel, we expect the true uncertainty extends beyond our predicted rates. None the less, we conclude that AGN remain an important channel for consideration, particularly for gravitational wave detections involving one or two high mass BHs.

Key words: gravitational waves – hydrodynamics – binaries: general – galaxies: nuclei – black hole mergers.

1 INTRODUCTION

Following the first gravitational wave (GW) detection in 2015 B. P. Abbott et al. (2016), we have been able to observationally detect BHs through both electromagnetic (e.g. Q. Z. Liu, J. van Paradijs & E. P. J. van den Heuvel 2006, 2007; C. J. Hailey et al. 2018; Event Horizon Telescope Collaboration 2019, 2022; T. A. Thompson et al. 2019; Gaia Collaboration 2024), and GW observations (e.g. B. P. Abbott et al. 2019a, 2020; T. Venumadhav et al. 2020; R. Abbott et al. 2020a, c), with over 170 more sources anticipated from the Laser Interferometer Gravitational-Wave Observatory (LIGO) public alert list for the O4 observing run.[1] The merging of BHs from gravitational wave emission requires the binary black hole (BBH) system to have an extremely small separation for the system to merge within a Hubble time (∼0.1 au for two 10 $\mathrm{M_\odot}$ BHs). As the separation of stellar binaries is typically far greater than this limit (E. Öpik 1924; A. A. Tokovinin 2000), what astrophysical mechanism(s) can bring BHs to such low separations is still an open question. Several possible explanations have been suggested: isolated stellar binary evolution and common envelope evolution (e.g. V. M. Lipunov, K. A. Postnov & M. E. Prokhorov 1997; K. Belczynski et al. 2010; M. Dominik et al. 2012, 2013, 2015; K. Belczynski et al. 2016; H. Tagawa, T. R. Saitoh & B. Kocsis 2018; M. Mapelli et al. 2021), three-body scatterings in nuclear and globular star clusters (e.g. H. Mouri & Y. Taniguchi 2002; M. C. Miller & D. P. Hamilton 2002; F. Antonini & H. B. Perets 2012; C. L. Rodriguez et al. 2015; C. L. Rodriguez, S. Chatterjee & F. A. Rasio 2016; J. Samsing & D. J. D'Orazio 2018; U. N. Di Carlo et al. 2020; B. Liu & D. Lai 2021), BBH mergers driven by gas in globular clusters (M. Rozner & H. B. Perets 2022) and in dense active galactic nucleus (AGN) discs (e.g. I. Bartos et al. 2017; N. C. Stone, B. D. Metzger & Z. Haiman 2017; B. McKernan et al. 2020a; H. Tagawa, Z. Haiman & B. Kocsis 2020a; B. McKernan, K. E. S. Ford & R. O'Shaughnessy 2020b; Y.-P. Li et al. 2021; J. Li, D. Lai & L. Rodet 2022; C. Rowan et al. 2023, 2024; J. Li et al. 2023; H. Whitehead et al. 2024a, b; V. Delfavero et al. 2025).

AGN provide a favourable environment for both BBH formation and mergers, due to the dense gaseous accretion disc orbiting the supermassive black hole (SMBH) at their centre (H. Tagawa et al. 2020a). Dynamical/accretion drag on objects crossing through the AGN disc can embed them within the geometrically thin disc for a considerably high number density of BHs (I. Bartos et al. 2017; G. Fabj et al. 2020). These BHs can then encounter one another

★ E-mail: connar.rowan@nbi.ku.dk

[1]see https://gracedb.ligo.org/superevents/public/O4/ for the current number of alerts.

and form stable binaries by dissipating their relative two-body energy via a complex interaction with the surrounding gas. The efficiency of this 'gas-assisted' binary formation mechanism has recently been validated by hydrodynamical simulations in our own works (e.g. C. Rowan et al. 2023, 2024; H. Whitehead et al. 2024a, b) and in e.g. J. Li et al. (2023), see also the analytical studies of S. DeLaurentiis, M. Epstein-Martin & Z. Haiman (2023), and M. Rozner, A. Generozov & H. B. Perets (2023). M. Dodici & S. Tremaine (2024) directly compared across previous simulation studies, finding good agreement for the island of parameter space in gas density and encounter impact parameter that leads to binary formation.

The evolution of a disc embedded binary is an open problem. In C. Rowan et al. (2023), we demonstrated that binaries can inspiral on short time-scales ($\lesssim 10^4$ yr), orders of magnitude shorter than the typical lifetime of an AGN, $t_{\rm AGN} = 10^7$ yr. The merger can be especially rapid for retrograde binaries (retrograde meaning the binary orbits counter to the orbit about the SMBH), due to enhanced accretion and gravitational torques from the gas that remove angular momentum. This picture is broadly consistent across other simulation studies with varying methodologies (e.g. C. Baruteau, J. Cuadra & D. N. C. Lin 2011; A. Secunda et al. 2019; J. Li et al. 2022, 2023; R. Li & D. Lai 2023, 2024). For prograde binaries that are highly circular and at larger separations, binaries have in some cases been found to out-spiral when an isothermal hydrodynamic treatment is used in 2D (e.g. Y.-P. Li et al. 2021).

With the ever increasing number of BBH merger detections from LIGO-VIRGO-KAGRA (LVK), we continually constrain the merger rate of these systems. Population studies and semi-analytic works (e.g. H. Bartko et al. 2009; F. Antonini & F. A. Rasio 2016; J. M. Bellovary et al. 2016; H. Tagawa et al. 2020a; B. McKernan et al. 2020b; K. E. S. Ford & B. McKernan 2022; M. P. Vaccaro et al. 2024) that predict the rate of mergers in AGN tend to suffer from a weakly motivated or highly simplified treatment of binary formation in gas, which naturally affects the anticipated rates. Arguably the most detailed of these studies is the 1D N-body simulations of H. Tagawa et al. (2020a, b, 2021). In that study a plethora of physical effects were accounted for, including: radial migration, binary-single interactions, merger kicks, and repeated mergers among many others. They found that gas-captured binaries make up the majority (> 85 per cent) of merging binaries in the AGN. If the binaries formed through the gas-formation mechanism do indeed represent the primary contribution to the BH mergers in AGN, it is vital that we accurately model the process. H. Tagawa et al. (2020a) assumed a binary was formed if the deceleration time-scale from dynamical friction (e.g. E. C. Ostriker 1999) was smaller than the binaries' Hill sphere crossing time. The validity of applying the Ostriker formula to this scenario is dubious, as the formalism assumes a uniform density and motion of the gas, whereas the embedded BHs will have their own very dense and rotating circum-single discs. This has been affirmed by detailed hydrodynamical studies (e.g. C. Rowan et al. 2023, 2024; J. Li et al. 2023; H. Whitehead et al. 2024a).

In this work, we build on two of our previous works C. Rowan et al. (2024) and H. Whitehead et al. (2024a) where a detailed capture criterion was derived from high resolution simulations of BH–BH scatterings. We apply this physically motivated and numerically verified criterion to a semi-analytic model of BHs in an AGN disc to test how the merger rates in the AGN channel change with the improved treatment of the gas-assisted binary formation mechanism, with a direct comparison to the simpler dynamical-friction-based formation criterion. We discuss the semi-analytic model and the inclusion of the formation criterion in Section 2 and present our results in Section 3 before concluding in Section 4. In Section 2.4, we describe the process of integrating the binary formation model from the hydrodynamical simulations to the Monte Carlo simulations of this work. References to equations in these papers are denoted explicitly in square brackets (e.g. equation [27]) for clarity.

2 METHODS

The core component of the AGN channel is the SMBH's accretion disc. Here we outline the properties of the disc, how the BHs become embedded into the disc, how gas can lead to formation, and finally how formed binaries can be driven to merge.

2.1 Disc setup

We consider the AGN disc model of E. Sirko & J. Goodman (2003) (hereafter SG discs). To observe the dependence on the mass of the AGN, we assume values for the SMBH mass $M_\bullet$ in the range $M_\bullet = 10^5\,{\rm M}_\odot - 10^9\,{\rm M}_\odot$ (consistent with the observed AGN mass range, e.g. J. E. Greene & L. C. Ho 2007, 2009) in uniform log-space, using 35 values.

The kinematic viscosity of the gas ν, which determines the angular momentum flow in SG discs, is related to the sound speed $c_{\rm s}$ and disc scale height H via

$$\nu = \alpha \beta^b c_{\rm s} H\,. \tag{1}$$

Here, α is the viscosity constant (N. I. Shakura & R. A. Sunyaev 1973) and $\beta \equiv P_{\rm gas}/(P_{\rm gas} + P_{\rm rad})$ is the ratio of the gas pressure $P_{\rm gas}$ to the total pressure, which includes the radiation pressure $P_{\rm rad}$. The parameter $b = \{0, 1\}$ acts as the switch between an α-disc ($b = 0$) and a β-disc ($b = 1$), see Z. Haiman, B. Kocsis & K. Menou (2009) for more on this viscosity treatment. We consider both the α-disc and β-disc treatment of the viscosity for our models of the AGN disc. We tack on the α, β labels to the SG abbreviation, such that 'SGα' corresponds to a Sirko-Goodman disc with the α viscosity treatment and 'SGβ' corresponds to a SG disc with the β viscosity treatment. In the SG model, the free parameters are the mass of the SMBH $M_\bullet$, the mean molecular mass $\mu_{\rm mol}$, the Eddington ratio $\epsilon = L/L_{\rm Edd}$ (where $L_{\rm Edd}$ is the Eddington luminosity) and viscosity constant α. Inline with our previous work (C. Rowan et al. 2023, 2024), we set $\alpha = 0.1$, $\mu_{\rm mol} = 0.6$, and $\epsilon = 0.2$.

We solve the disc equations explicitly, using the `pAGN` code of D. Gangardt et al. (2024) to obtain the radial disc profiles of density ρ and sound speed $c_{\rm s}$. As the binary formation criterion depends strongly on the gas density (see Section 2.7.2 below), it is essential to model the AGN discs in an accurate manner[2] that does not depend on being in a single disc regime (i.e. the gas pressure dominated regime assumed in C. Rowan et al. 2023). We model the discs out to where the Toomre Q parameter predicts gravitational instability, which occurs when

$$Q \equiv \frac{c_{\rm s}\Omega}{\pi G \Sigma} \lesssim 1\,, \tag{2}$$

where $\Omega = \sqrt{GM_\bullet/R^3}$ is the angular frequency of a Keplerian orbit, G is the gravitational constant, and Σ is the gas surface density. The R value that satisfies equation (2) provides the outer radius of the disc $R_{\rm disc}$ and the upper domain limit for our simulations. Over the range in $M_\bullet$, this value varies minimally around $R \sim 0.01$pc. We set

[2]We fix a minor error in `pAGN` where the molecular mass is fixed at unity in the ideal gas law.

the inner R boundary with $R_{\rm min} = \max(3R_\bullet, 10^{-4}{\rm pc})$, where $R_\bullet$ is the Schwarzschild radius.

2.2 Objects crossing the disc

We assume all gas-assisted BBH formation and mergers can only occur within the disc therefore it is necessary to know how many BHs are on orbits that cross the disc. We calculate the number of objects that cross the disc within the simulation domain $R_{\rm min} \leq a \leq R_{\rm disc}$. The number density of stars n_*, and by extension their birthed BHs, in terms of the orbital semimajor axis a is taken to be a mass segregated function consistent with the central O-star distribution (H. Bartko et al. 2009; R. M. O'Leary, B. Kocsis & A. Loeb 2009; U. Keshet, C. Hopman & T. Alexander 2009), representing strongly mass segregated equilibrium for the heaviest central objects

$$n_*(a) \propto a^{-2.5}\,. \tag{3}$$

The distribution is assumed to be spherical with orbital inclinations i sampled uniformly in $\cos i$. The eccentricity of the orbits is taken from the thermal distribution $f(e) = 2e$. The maximal distance from the SMBH where the local dynamics are dominated by its presence is $R_{\rm inf} = GM_\bullet/\sigma_\bullet^2$, where $\sigma_\bullet$ is the velocity dispersion of objects in the central nuclear region. Using the $M-\sigma$ relation $M_\bullet = M_0(\sigma_\bullet/\sigma_0)^k$ (K. Gebhardt et al. 2000; L. Ferrarese & D. Merritt 2000; K. Gültekin et al. 2009), adopting $k = 4.384$, $M_0 = 3.097\times10^8\,{\rm M}_\odot$, and $\sigma_0 = 200\,{\rm km\ s^{-1}}$ from J. Kormendy & L. C. Ho (2013), this gives

$$R_{\rm inf} = \frac{GM_0}{\sigma_0^2}\left(\frac{M_\bullet}{M_0}\right)^{0.544}\,. \tag{4}$$

The fraction of BHs crossing the disc is then given by

$$f_{\rm cross} = \frac{1}{N_{\rm BH}} \iint\limits_{\substack{a(1-e)<R_{\rm disc}\\ a<R_{\rm inf}}} n(a)\,f(e)\,4\pi a^2 {\rm d}a\,{\rm d}e\,, \tag{5}$$

where we integrate over the number of BHs $N_{\rm BH}$ within $R_{\rm inf}$ that have a periapsis within the outer radial limit of the disc $R_{\rm disc}$. The number of BHs sampled in our Monte–Carlo scheme (see §2.6) is $f_{\rm cross}N_{\rm BH}$.

2.3 Population sample

Based on a standard power-law initial mass function (IMF), various models for the BH initial mass function (BIMF) are used in the Monte Carlo simulations of Section 2.6 to monitor their effect on the rates. The stellar population is represented by the initial stellar mass function of P. Kroupa (2001)

$$\frac{{\rm d}N_*}{{\rm d}m_*} \propto m_*^{-\gamma}\,. \tag{6}$$

Here, m_* is the zero-age main-sequence mass of stars with an assumed stellar mass range of $0.1\,{\rm M}_\odot \leq m_* \leq 140\,{\rm M}_\odot$ and N_* is their abundance. For high stellar masses relevant for forming BHs, the exponent γ has a range of $1.7 \leq \gamma \leq 2.35$, with the fiducial value taken to be $\gamma = 2.35$, consistent with E. E. Salpeter (1955). The stellar masses are sampled such that the total stellar mass within $R_{\rm inf}$ is $2M_\bullet$, in line with J. Binney & S. Tremaine (2008), B. Kocsis & J. Levin (2012), and I. Bartos et al. (2017). Three BH initial mass functions are considered, labelled according to their source.

2.3.1 BIMF 1

To make an accurate comparison to a highly relevant work, the BIMF of H. Tagawa et al. (2020a) is implemented. This distribution is derived from the stellar population synthesis simulations of K. Belczynski et al. (2010) corresponding to a solar metallicity and weak winds.

$$\frac{m_{\rm BH}(m_*/{\rm M}_\odot)}{{\rm M}_\odot} = \begin{cases} \text{no BH} & m_* < 20\,{\rm M}_\odot\,,\\ m_*/4 & 20\,{\rm M}_\odot \leq m_* < 40\,{\rm M}_\odot\,,\\ 10 & 40\,{\rm M}_\odot \leq m_* < 55\,{\rm M}_\odot\,,\\ m_*/13+5.77 & 55\,{\rm M}_\odot \leq m_* < 120\,{\rm M}_\odot\,,\\ 15 & 120\,{\rm M}_\odot \leq m_* \leq 140\,{\rm M}_\odot\,, \end{cases} \tag{7}$$

with an assumed stellar mass range of $0.1\,{\rm M}_\odot \leq m_* \leq 140\,{\rm M}_\odot$. By sampling the stellar masses from equation (6) and applying equation (7), the initial BH mass distribution is obtained. Unlike in H. Tagawa et al. (2020a), the BH mass function is not evolved here, i.e. only first generation mergers are considered.

2.3.2 BIMF 2

To also compare with the pre-existing binary rates of I. Bartos et al. (2017), we adopt their simplified power-law distribution for their binary mass as our BH mass function,

$$\frac{{\rm d}N_{\rm BH}}{{\rm d}M_{\rm BH}} \propto M_{\rm BH}^{-\beta}\,, \tag{8}$$

with the equivalent mass range of $5\,{\rm M}_\odot \leq M_{\rm bin} \leq 50\,{\rm M}_\odot$ and exponent $2.0 \leq \beta \leq 2.5$. The slope of mass distribution reflects the upper bound identified by B. P. Abbott et al. (2019b), which is more consistent with X-ray binary observations (e.g. F. Özel et al. 2010; C. S. Kochanek 2015), which have even steeper dependence. The simplicity of the function allows us to observe how the initial BH mass function compares to the merging mass function, discussed in Section 3.3. The same number of BHs as derived from the BIMF of H. Tagawa et al. (2020a) is maintained to compare directly the merger rates' dependence on the mass distribution of BHs.

2.3.3 BIMF 3

Our only current inference of the BIMF from observations is from X-ray binaries and GW observations. As our third BIMF, we adopt the mass distribution of E. J. Baxter et al. (2021). The function was derived by constraining the mass gap from the BH mass distribution from GW events (see R. Abbott et al. 2021) in tandem with stellar evolution theory, incorporating the predicted mass gap from the GW mass distribution. The function is given by

$$\frac{{\rm d}N_{\rm BH}}{{\rm d}M_{\rm BH}} \propto M_{\rm BH}^b\left[1+\frac{2a^2M_{\rm BH}^{1/2}(M_{\rm BHMG}-M_{\rm BH})^{a-1}}{M_{\rm BHMG}^{1-\frac{a}{2}}}\right], \tag{9}$$

where $M_{\rm BHMG}$ is the low mass edge of the mass gap. The value of $M_{\rm BHMG}$ is set to $47.7\,{\rm M}_\odot$ and the dimension-less constants are taken to be $a = 0.39$ and $b = -2.2$, as suggested in E. J. Baxter et al. (2021). We maintain the same minimum BH mass of $5\,{\rm M}_\odot$ for consistency and use the same number of BHs as the other BIMFs.

We show the normalized distributions of all IMFs, which we hereafter refer to as ${\rm BIMF_{Tagawa}}$, ${\rm BIMF_{Bartos}}$, and ${\rm BIMF_{Baxter}}$ in Fig. 1.

Qualitatively, ${\rm BIMF_{Bartos}}$ and ${\rm BIMF_{Baxter}}$ allow for larger BH masses up to $\sim 50\,{\rm M}_\odot$, whereas ${\rm BIMF_{Tagawa}}$ produces more BHs in the range of $10-12\,{\rm M}_\odot$ ($10-15\,{\rm M}_\odot$) for $\gamma = 2.35$ (1.7). The spike at $M_{\rm BH} = 10\,{\rm M}_\odot$ comes from the $40-55\,{\rm M}_\odot$ condition of equation

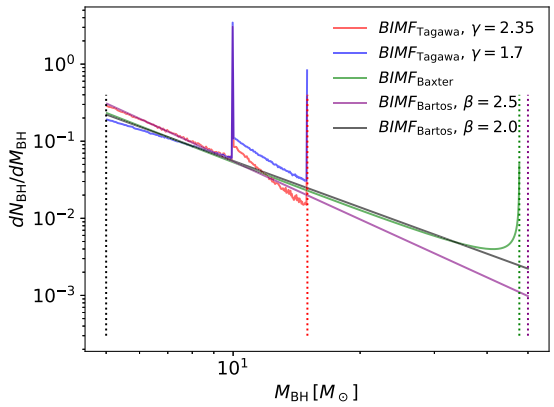

Figure 1. Normalized black hole initial mass functions $\mathrm{BIMF_{Tagawa}}$, $\mathrm{BIMF_{Bartos}}$, and $\mathrm{BIMF_{Baxter}}$ (equations 7, 8, and 9, respectively). $\mathrm{BIMF_{Tagawa}}$ is shown for $\gamma = \{1.7, 1.35\}$ and $\mathrm{BIMF_{Bartos}}$ for $\beta = \{2, 2.5\}$. The vertical lines of $\mathrm{BIMF_{Tagawa}}$ are a result of the $40\,\mathrm{M}_\odot \leq m_* < 55\,\mathrm{M}_\odot$ and $120\,\mathrm{M}_\odot \leq m_* \leq 140\,\mathrm{M}_\odot$ conditions of equation (7). The vertical cutoff of $\mathrm{BIMF_{Baxter}}$, represents the lower boundary of the BH mass gap.

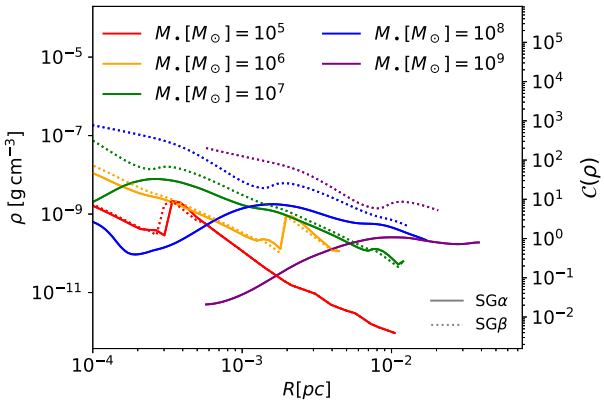

Figure 2. The radial density profiles of the AGN disc generated from `pAGN` as a function of radial distance in the AGN disc R along side the resulting modification $\mathcal{C}$ to the energy dissipation during encounter (equations 13–14). Results shown for the SGα and SGβ disc models with parameters $M_{\rm bin} = 20\,\mathrm{M}_\odot$ and $M_\bullet = \{10^5, 10^6, 10^7, 10^8, 10^9\}\,\mathrm{M}_\odot$.

(7). As very few stars are formed with masses $120 - 140\,\mathrm{M}_\odot$, the anticipated spike at $M_{\rm BH} = 15\,\mathrm{M}_\odot$ is far less significant.

2.4 Gas dissipation during the encounter

We utilize the semi-analytic prescriptions derived in our previous works C. Rowan et al. (2024) and H. Whitehead et al. (2024a) to model the energy dissipation of a BH–BH encounter. The orbital energy dissipation $\Delta E_{\rm bin}$ from gas drag during the first encounter is assumed to be described by equation [16] in C. Rowan et al. (2024). Its form is a power law with the depth of the first periapsis passage $r_{\rm per,1}$

$$\Delta E_{\rm bin}(r_{\rm per,1}) = -x\left(\frac{r_{\rm per,1}}{r_{\rm H}}\right)^{y} |E_{\rm H,bin}|\,, \tag{10}$$

where

$$|E_{\rm H,bin}| = \frac{G M_{\rm bin}\mu}{2r_{\rm H}} \tag{11}$$

is the absolute orbital energy of the binary in the centre-of-mass frame at a separation of one Hill radius, which we define as

$$r_{\rm H} = R\left(\frac{M_{\rm bin}}{6M_\bullet}\right)^{1/3} \tag{12}$$

to remain consistent with the parametrization of C. Rowan et al. (2024). We extend the relation to arbitrary densities based on the finding of our other paper H. Whitehead et al. (2024a) that dissipation (normalized to $E_{\rm H,bin}$) scales linearly with the gas density in the Hill sphere ρ_H. The two parameters that change ρ_H are the local sound speed and surface density as $\rho_H = \Sigma/(2H) = \Sigma\Omega/c_{\rm s}$. We assume ρ_H scales with the ambient density ρ according to the findings of H. Whitehead et al. (2024a). We modify equation (10) to account for the changing density in the Hill sphere at different radii in the AGN disc via an additional scaling

$$\frac{\Delta E_{\rm bin}(R, r_{\rm per,1})}{|E_{\rm H,bin}|} = -x\left(\frac{r_{\rm per,1}}{r_{\rm H}}\right)^{y} \mathcal{C}(R, M_\bullet)\,. \tag{13}$$

The function $\mathcal{C}$ is given by

$$\mathcal{C}(R, M_\bullet) = \left(\frac{\rho(R, M_\bullet)}{\rho_0}\right)^{w}, \tag{14}$$

where $\rho(R, M_\bullet)$ is the ambient density of the disc at the radial location R of the binary. The normalization $\rho_0 \simeq 6.5 \times 10^{-10}\,\mathrm{g\,cm^{-3}}$ is set by the original simulations in C. Rowan et al. (2024). Inline with H. Whitehead et al. (2024a), we set $w = 1$. The same following methodology is applied from section 3.5 of C. Rowan et al. (2024) to reconstruct equation [33] using equation (13). We adopt the fiducial values in C. Rowan et al. (2024) of $x = 1.3 \times 10^{-4}$, $y = -0.4$.[3] The full capture criterion is then

$$\left[\frac{E_{\rm 2H}}{|E_{\rm H,bin}|} < \left(\frac{E_{\rm 2H}}{|E_{\rm H,bin}|}\right)_{\rm crit}\right] \wedge \left[p_{\rm 1H} < 0.68 r_{\rm H}\right], \tag{15}$$

$$\left(\frac{E_{\rm 2H}}{|E_{\rm H,bin}|}\right)_{\rm crit} = \underbrace{\mathcal{C}(R, M_\bullet) 10^{-1.74\frac{p_{\rm 1H}}{r_{\rm H}} - 2.52}}_{|\Delta E_{\rm bin}| \text{for encounter}} - \underbrace{10^{1.31\frac{p_{\rm 1H}}{r_{\rm H}} - 4.34}}_{\text{max stable } E_{\rm bin}}, \tag{16}$$

where $E_{\rm 2H}$ is the binary energy in the centre-of-mass frame and $p_{\rm 1H} = |\hat{\mathbf{v}}_{\rm rel} \times \mathbf{\Delta R}|$ is the impact parameter of the encounter at a separation of $\Delta R = r_{\rm H}$ where the relative velocity unit vector is $\hat{\mathbf{v}}_{\rm rel}$, see equation [30] in C. Rowan et al. (2024). This is calculated using a numerical simulation of the single–single scattering process starting from a given set of initial conditions at a separation of $2r_{\rm H}$ without gas. The second term represents the maximum energy for the binaries to remain bound in the simulations of C. Rowan et al. (2024) for more than two encounters (all binaries subsequently hardened after two encounters). We show the value of $\mathcal{C}$ as a function of R for five different $M_\bullet$ according to the `pAGN` disc profiles in Fig. 2, showing that the anticipated dissipation $\Delta E_{\rm bin}/E_{\rm H,bin}$ can vary by 3–4 orders of magnitude due to the density variation across the AGN disc. The increase in ρ and $\mathcal{C}$ for the SGβ disc model is the result of lower viscosities in the radiation-dominated zone, requiring a greater amount of mass at each radius in order to achieve the same Eddington luminosity fraction.

The primary determinant for binary capture is whether the energy condition of equation (15) is satisfied. At separation of $2r_{\rm H}$, the relative velocity of the two objects is taken to be the combined magnitude of the ambient velocity dispersion in the disc, $\sigma_{\rm disp}$ and

[3]Though we find in C. Rowan et al. (2024) that $y = -0.6$ for 3 times the ambient density ($\rho_0 \simeq 6.5 \times 10^{-10}\,\mathrm{g\,cm^{-3}}$), it is still within the confidence interval of the error bounds and we retain the same $y = -0.4$ scaling for simplicity. H. Whitehead et al. (2024a) suggests that y is fixed for our fiducial density and below.

the Keplerian shear over the radial separation of the BHs ΔR, i.e. $\sigma_{\rm Kep} = |R\,\Delta R\,d\Omega/{\rm d}R| = \frac{3}{2}\Omega\Delta R$. The energy $E_{\rm 2H}$ is then

$$E_{\rm 2H} = \frac{1}{2}\mu v_{\rm rel}^2 - \frac{GM_{\rm bin}\mu}{2r_{\rm H}}\,. \tag{17}$$

Using $v_{\rm rel}^2 = \sigma_{\rm Kep}^2 + \sigma_{\rm disp}^2$, the average encounter energy is

$$E_{\rm 2H} = \frac{1}{2}\mu(\sigma_{\rm Kep}^2 + \sigma_{\rm disp}^2) - \frac{GM_{\rm bin}\mu}{2r_{\rm H}}\,. \tag{18}$$

Expressing equation (18) in units of $E_{\rm H,bin}$ (as required by equation 15), this gives

$$\frac{E_{\rm 2H}}{|E_{\rm H,bin}|} = \frac{r_{\rm H}(\sigma_{\rm Kep}^2 + \sigma_{\rm disp}^2)}{GM_{\rm bin}} - 1\,. \tag{19}$$

We take the maximum approach velocity to be the shear over, $\Delta R = 2r_{\rm H}$, i.e $\sigma_{\rm Kep,max} = 3\Omega r_H$. This gives the maximum possible encounter energy as

$$\begin{aligned}\frac{E_{\rm 2H}}{|E_{\rm H,bin}|_{\rm max}} &= \frac{r_{\rm H}\Omega^2(9r_{\rm H}^2 + H^2)}{GM_{\rm bin}} - 1 = 9\frac{M_\bullet}{M_{\rm bin}}\frac{r_{\rm H}^3}{R^3} + \frac{r_{\rm H}H^2}{R^3} - 1\\ &= \frac{1}{2} + \left(\frac{M_{\rm bin}}{6M_\bullet}\right)^{1/3}\frac{H^2}{R^2}\,.\end{aligned} \tag{20}$$

The ambient velocity dispersion of objects in the disc $\sigma_{\rm disp} = Hv_{\rm Kep}/R$ is assumed to be equal to the local sound speed, so their mean vertical motion is of the order of H. This is a conservative assumption as it assumes BHs that have not yet aligned have zero chance of forming binaries via gas-assisted captures and binaries in the disc do not further settle to the midplane. For the assumed parameters, $\sigma_{\rm disp}^2/\sigma_{\rm Kep}^2$ ranges from $10^{-2} - 10^{-4}$.

This formalism does not account for any eccentricity in the BH orbits around the SMBH due to gas effects when embedded objects open up a gap (R. Sari & P. Goldreich 2004), additional gravitational focusing on approach to $r = 2r_{\rm H}$ or the long-term depletion of gas due to gap opening prior the encounter.[4] Using equations (32) and (33) of Z. Pan & H. Yang (2021) (see also B. Kocsis, N. Yunes & A. Loeb 2011), the gap opening criteria is only marginally satisfied for $M_\bullet \lesssim 10^5\,{\rm M}_\odot$ and $M_\bullet \lesssim 5\times10^5\,{\rm M}_\odot$ for $M_{\rm BH} = 10\,{\rm M}_\odot$ and $50\,{\rm M}_\odot$ respectively (assuming $\alpha = 0.1$). Therefore excluding this effect likely does not affect the predictions of the simulations. Embedded BHs may also have modified eccentricities and or encounter energies from scatterings with other objects prior to the encounter. The two-body relaxation time-scale (calculated according to S. Tremaine et al. 2002; S. Naoz et al. 2022) of the system is $\{0.05, 1.5, 50\}$Myr for $M_\bullet = \{10^5, 10^7, 10^9\}\,{\rm M}_\odot$. This calculation assumes the size of the system is 0.01 pc, the RMS BH mass is 10 $\rm M_\odot$ and the confinement of objects to the disc increases the density and reduces the two-body relaxation time-scale by a factor $\sim (H/0.01\,{\rm pc})^2$. Taking the lifetime of the AGN to be $t_{\rm AGN} = 10$ Myr, scattering from a third object could be important in the low $M_\bullet$ regime, however note this is an underestimate as it assumes all objects in the system have been embedded in the disc and ignores stars/neutron stars. The resulting additional velocity dispersion from binary-single scatterings could potentially reduce the merger rate via an increase in the value of $\sigma_{\rm disp}$ in equation (19)–(20). We assess the relative importance of binary-single scatterings in our simulations further in Section 2.7.1.

In Fig. 3 we show the maximum allowable encounter energies $(E_{\rm 2H}/E_{\rm crit})$ that lead to successful binary formation for select values of $p_{\rm 1H}$, assuming three values of $M_\bullet$.

[4]The simulations of C. Rowan et al. (2024) evolved the BHs for just a few AGN orbits before their encounter.

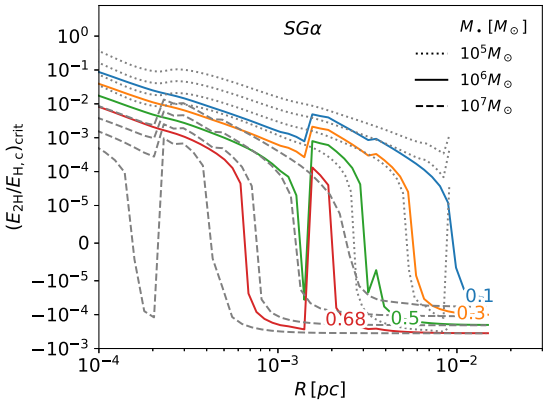

Figure 3. Maximum initial encounter energy $E_{\rm 2H}$ of a binary that leads to a successfully formed binary for different impact parameters $p_{\rm 1H}$, as labelled on the curves, as a function of radial distance in the AGN disc R. Results shown for an SGα disc with parameters $M_{\rm bin} = 20\,{\rm M}_\odot$ and $M_\bullet = [10^5, 10^6, 10^7]\,{\rm M}_\odot$. At lower R, the lower velocity dispersion and higher gas density allows BHs with larger initial encounter energies to dissipate enough energy to stay bound. Closer encounters at low impact parameters can extend binary formation to larger R.

The contours of $(E_{\rm 2H}/E_{\rm crit})$ with $p_{\rm 1H}$ at low R form a steep power law where initially unbound BHs can still be captured into binaries, driven by the first term in equation (16). Then, as C decreases at higher R, the second term in equation (16) dominates and only encounters with increasingly negative energies will form stable binaries, tending towards an asymptotic value dependent on $p_{\rm 1H}$. Decreasing $p_{\rm 1H}$ values lead to higher dissipation values and allows binaries to be more easily formed at higher energies further out in the AGN disc (for the same typical encounter energy). The transition point from the power law to the plateau occurs at lower R for smaller $M_\bullet$. In the low $M_\bullet$ regime, the density profile of the AGN exhibits an additional minimum due to the opacity break in the inner region of the disc, leading to a region of inefficient binary formation, suggesting a region where the density of gas-captured binaries could be lower than anticipated otherwise. However, the number of BHs in this region, particularly for low $M_\bullet$, is likely not significant.

2.5 The formation function

To determine the rate of BH mergers, it is necessary to define the statistical fraction of BH encounters that will lead to binary formation $f_{\rm form}(R, M_{\rm bin}, M_\bullet)$. Here, encounters are defined as events where the separation of two BHs is less than the binary Hill radius. In Fig. 4, $f_{\rm form}$ is shown as a function of R for a range of $M_\bullet$, sampling the initial mass distribution from BIMF$_{\rm Tagawa}$.

The ambient velocity dispersion term $\sigma_{\rm disp}$ is sampled from a random Gaussian distribution with standard deviation $\sigma \equiv c_{\rm s}$ centred on zero. The upper bound of ΔR in accordance with equation (20) is taken to be $2r_{\rm H}$. This is also a conservative assumption as C. Rowan et al. (2023), H. Whitehead et al. (2024a), and C. Rowan et al. (2024) indicated BHs can be focused into minimum encounter separations smaller than $r_{\rm H}$ for $\Delta R > 2r_{\rm H}$ depending on the disc density. The values of ΔR and $p_{\rm 1H}$ are sampled uniformly[5] between $0.01r_{\rm H} - 2r_{\rm H}$ and $0 - r_{\rm H}$ respectively for approaching BHs, where the lower ΔR

[5]In practice there is a correlation between $p_{\rm 1H}$ and ΔR, however as it is unclear how this should be affected by the ambient velocity dispersion $\sigma_{\rm disp}$, they are sampled randomly.

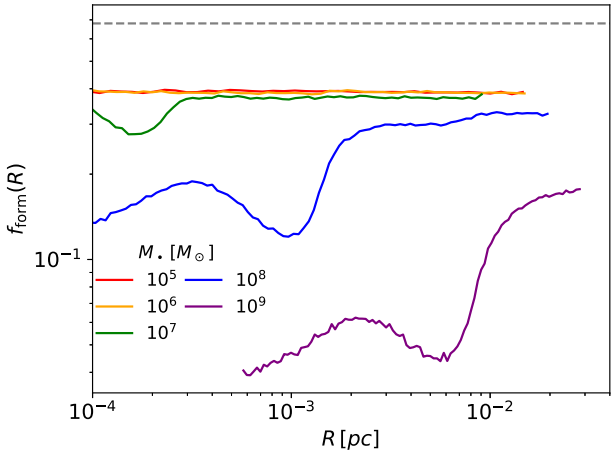

Figure 4. Fraction of encounters with impact parameters $p_{\rm 1H} < r_{\rm H}$ that lead to successfully formed binaries as a function of radial distance in the SMBH disc for different $M_\bullet$, assuming an SGα disc. The function is shown assuming the BH mass function of BIMF$^{\gamma=2.35}_{\rm Tagawa}$ and a uniform distribution of $p_{\rm 1H}$ and ΔR is assumed. Results show AGN with higher $M_\bullet$ have a reduced formation probability for higher R.

value is given a non-zero value to avoid divergence in the calculation of the time between encounters later in Section 2.6.1.

The formation function is weakly dependent on $M_\bullet$, with a decrease of roughly an order of magnitude in the formation probability from $M_\bullet = 10^5\,{\rm M}_\odot$ to $10^9\,{\rm M}_\odot$. The formation probability becomes flat for $M_\bullet < 10^7\,{\rm M}_\odot$ as the gas dissipation becomes highly efficient at forming binaries from typically lower relative velocities ($E_{\rm 2H}/E_{\rm H,c} \propto M_\bullet^{2/3}$). Similarly, the formation probability is larger at higher radii in the AGN disc. The initial increases in $f_{\rm form}$ result from the maxima in the density (ρ) profiles, which shifts to higher R for larger $M_\bullet$.

2.6 BH binary merger rates using a Monte Carlo approach

In order for an isolated black hole to merge, it must satisfy four conditions within $t_{\rm AGN}$: (i) the BHs must align with the disc, (ii) encounter another BH, (iii) successfully form a binary, and (iv) successfully merge. To estimate the overall merger rate, we first consider the time-scale for a single BH to go through each of these stages in its evolution.

2.6.1 The time-scales of the system

We assume the alignment time-scale is derived according to I. Bartos et al. (2017). Given some initial vertical velocity v_z at the point of disc crossing and a typical velocity reduction Δv_z upon crossing the disc from dynamical friction, the general expression for the characteristic time-scale of disc alignment is

$$t_{\rm align} \simeq \frac{t_{\rm orb}}{2}\frac{v_z}{\Delta v_z}\,. \tag{21}$$

The $t_{\rm orb} = 2\pi R^{3/2}(GM_\bullet)^{-1/2}$ term is the orbital period, and the factor 2 accounts for two crossings per $t_{\rm orb}$. The fractional change in velocity of an object is equated to the ratio of the mass accreted during its crossing of the disc and its own mass such that $\Delta v_z/v_z = \Delta M_{\rm cross}/M_{\rm BH}$. The accreted mass is assumed to be that within its Bondi–Hoyle–Lyttleton radius $r_{\rm BHL} = 2GM_{\rm BH}/(\Delta v^2 + c_s^2)$, where $\Delta v = v_{\rm orb}((1-\cos(i))^2 + \sin^2(i))^{1/2} = 2v_{\rm orb}\sin(\frac{i}{2})$ is the relative velocity of the BH to the gas, which orbits the SMBH with velocity $v_{\rm orb} = \sqrt{GM_\bullet/R}$. The crossing mass is then[6] $\Delta M_{\rm cross} = \Delta v t_{\rm cross} r_{\rm BHL}^2 \pi\Sigma/(2H)$ with crossing time $t_{\rm cross} \approx 2H/(v_{\rm orb}\sin i)$. Putting all this together gives an alignment time of

$$\begin{aligned} t_{\rm align} &= \frac{t_{\rm orb}}{2}\frac{\cos(i/2)(\Delta v^2 + c_s^2)^2}{4G^2M_{\rm BH}\pi\Sigma} \\ &= \frac{t_{\rm orb}M_\bullet^2}{2M_{\rm BH}M_{\rm d}}\cos\left(\frac{i}{2}\right)\left[\sin^2\left(\frac{i}{2}\right) + \frac{H^2}{4R^2}\right]^2, \end{aligned} \tag{22}$$

where the identity $\sin(i)/\sin(i/2) = 2\cos(i/2)$ has been applied and we define $M_{\rm d} = 2\pi R^2\Sigma$. Note that the second term in the parenthesis $H^2/4R^2$ is much smaller unless the orbit is close to being fully embedded in the disc where $\sin(i/2) \approx \frac{1}{2}\sin i = H/2R$. Note the strong dependence on the velocity term $(\Delta v^2 + c_s^2)^2$ to the fourth power, which makes it increasingly difficult to embed objects for larger SMBH masses for a fixed R or $M_{\rm d}$. Note that the potential of the AGN disc or stellar population is not accounted for.

The encounter time-scale $t_{\rm enc}$ is given by:

$$t_{\rm enc} = \frac{1}{n_{\rm BH}r_{\rm H}z_{\rm H}v_{\rm rel}} \approx \frac{2/3}{n_{\rm BH}\Omega r_{\rm H}^3} = \frac{4}{n_{\rm BH}G^{1/2}M_\bullet^{-1/2}M_{\rm bin}R^{3/2}}\,, \tag{23}$$

where $n_{\rm BH}$ is the volume number density of BHs, $z_{\rm H} = \min(H, r_{\rm H})$ is the vertical cross-section for the encounter in the case $r_{\rm H} < H$.[7] We assume the relative velocity is equivalent to the velocity shear across the entire Hill sphere $v_{\rm rel} = \frac{3}{2}\Omega\, r_{\rm H}$ (since objects can approach from inside and outside the BHs orbit). In practice $v_{\rm rel}$ is likely larger due to the assumed velocity dispersion of BHs and the ability for an approaching BH to be perturbed from higher/lower radii to a radial separation of $r_{\rm H}$ due to gravitational focusing (e.g.T. C. N. Boekholt, C. Rowan & B. Kocsis 2023). Once a BH encounters another in the disc, the formation likelihood is given by $f_{\rm form}$. This then modifies $t_{\rm enc}$ to give the effective formation time-scale

$$t_{\rm form} = \frac{t_{\rm enc}}{f_{\rm form}}\,. \tag{24}$$

Determining the value of $f_{\rm form}$ requires knowledge of the expected number and masses of BHs at a particular point of the disc, which we discuss in Section 2.7.2 below.

Perhaps the most uncertain time-scale is the merger time-scale. It has been shown that while retrograde binaries can reliably merge (e.g. J. Li et al. 2022; C. Rowan et al. 2023; R. Li & D. Lai 2023; R. Li & D. Lai 2024), prograde binaries can in some cases outspiral (e.g. Y.-P. Li et al. 2021; A. M. Dempsey et al. 2022). These binaries are typically given zero initial eccentricity, though it has been shown that eccentricity persists long after the initial formation. Additionally it has been shown that hotter circumbinary discs (more typical when the isothermal assumption is relaxed) lead to reliable inspiral (e.g. C. Baruteau et al. 2011; Y.-P. Li et al. 2022). In this work, like most other population studies (e.g. H. Tagawa et al. 2020a; M. Mapelli et al. 2021; K. E. S. Ford & B. McKernan 2022), it is assumed that the binaries formed are reliably hardened by gas. The merger time-scale itself is significantly smaller than the AGN lifetime (e.g. A. Cavaliere & P. Padovani 1989; M. G. Haehnelt & M. J. Rees 1993). Nevertheless, given the uncertainty of the inspiral

[6] Note that in reality $r_{\rm BHL}$ should be replaced by $\min(r_{\rm BHL}, r_{\rm H})$, but $r_{\rm BHL}/r_{\rm H} < \frac{3^{1/3}}{4\sin^2(i/2)}(M_{\rm BH}/M_\bullet)^{2/3} \ll 1$ for $i \gg 3^{1/6}(M_{\rm BH}/M_\bullet)^{1/3}$. The final alignment may be prolonged if $H/R < 3^{1/6}(M_{\rm BH}/M_\bullet)^{1/3}$. Note that near the disc's self-gravitating boundary $H/R \sim M_{\rm d}/2M_\bullet$ [see equation (2) and $c_{\rm s} = \Omega H$], implying a prolonged relaxation if $M_{\rm d} < (M_{\rm BH}/M_\bullet)^{1/3}M_\bullet$.

[7] Similar to the conditions required to open a gap in the disc, the vertical cross-section $z_{\rm H}$ is only less than $r_{\rm H}$ for low $M_\bullet$, here typically $M_\bullet < 5\times10^5\,{\rm M}_\odot$.

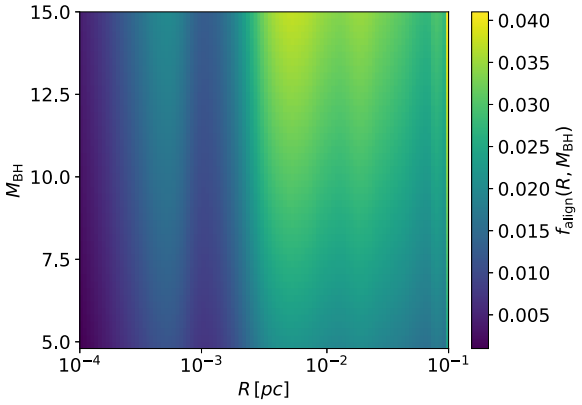

Figure 5. Fraction $f_{\rm align}$ of BHs with mass $M_{\rm BH}$ at radius R aligning with an SGα AGN disc for fiducial parameters $M_\bullet = 4 \times 10^6\,{\rm M}_\odot$, $\mathrm{BIMF}^{\gamma=2.35}_{\rm Tagawa}$, and $t_{\rm AGN} = 10^7$yr. Generated by evaluating equation (22) over uniform $\cos i$. Figure shows higher mass BHs can align quicker, with a larger fraction aligning typically at larger radii.

rate still present in the literature, the maximal merger time-scale in I. Bartos et al. (2017) of $t_{\rm merge} \sim 10^5$yr is used corresponding to $M_\bullet = 10^6\,{\rm M}_\odot$ at 0.01 pc. At higher $M_\bullet$ or lower R, the inspiral rate is shorter, but the value is maintained for all binaries as pessimistic assumption. Even at this upper bound, the merger time-scale is still two orders of magnitude shorter than the AGN lifetime, leaving $t_{\rm align}$ and $t_{\rm enc}$ as the more impactful time-scales for calculating merger rates.

The full time-scale of a BH to merge, including the dependencies is

$$t_{\rm tot} = t_{\rm align}(R, M_{\rm BH}) + t_{\rm form}(R, M_{\rm BH}, n_{\rm BH}) + t_{\rm merge}\,. \tag{25}$$

The merger rate is then the number of BHs which satisfy $t_{\rm tot} < t_{\rm AGN}$ divided by $2t_{\rm AGN}$, where the factor 2 accounts for double counting.

2.7 Resolving dependencies

2.7.1 BH number density

The complexities of the calculation lie within the dependencies of $n_{\rm BH}$ and $f_{\rm form}$. Starting with the former, the number density of BHs in the disc $n_{\rm BH}$ is determined from the initial sample of $\{M_{{\rm BH},i}\}$ and $\{R_i\}$ (see Section 2.3) that satisfy $t_{\rm align} < t_{\rm AGN}$. To smooth out the stochasticity in the sampling from the calculation of $n_{\rm BH}$, a probability grid, $f_{\rm align}(R, M_{\rm bin})$, in bins of $M_{\rm BH}$ and R is constructed by sampling over the full range of $M_{{\rm BH},i}$ and R_i, calculating their alignment time using equation (22). For each bin in $M_{\rm BH}$ and R, the alignment time-scale is sampled over the full range of $\cos i$ and the probability of that BH to align is the number of instances where the alignment time condition is met as a fraction of the number of inclination samples. The grid is constructed in 100 bins of R in log space and 50 bins of $M_{\rm BH}$ in linear space. For a fiducial setup of $M_\bullet = 4 \times 10^6\,{\rm M}_\odot$, the probability of alignment across $M_{\rm BH}$ and R is shown in Fig. 5. As predicted by equation (22), the figure indicates BHs at smaller R embed themselves more easily in the AGN disc.

The number density is similarly calculated as a grid in $M_{\rm BH}$ and R by sampling over the full initial distributions of $\{M_{{\rm BH},i}\}$ and $\{R_i\}$, binning them into the same bins for $p_{\rm align}$ and then adding the probability for that BH to align with the disc. The number density is also represented as a grid in $M_{\rm BH}$ and R to keep track of the mass distribution at each radius, as this affects the calculation of $f_{\rm form}$. When evaluating $t_{\rm enc}$ in equation (23), the number density is the sum of the number densities across the mass bins.

We can compare the typical time-scale of binary-single interactions in the disc $t_{\rm bs} = \frac{1}{n_{\rm BH} r_{\rm H} z_{\rm H} v_{\rm rel}}$ to the merger time-scale $t_{\rm merge}$. We show how this time-scale varies with $M_\bullet$ and R in Fig. 7.

The ratio of $t_{\rm bs}/t_{\rm merge}$ indicates that binary-single encounters likely play a subdominant role in the outer edges of the AGN disc, which as we will see is where the majority of our mergers occur (See Section 3.3). The shorter binary-single time-scales when adopting the more top heavy stellar mass function of $\mathrm{BIMF}^{\gamma=1.7}_{\rm Tagawa}$ are driven by the higher number of BHs in the initial NSC population, which drives a higher object number density in the disc. Higher SMBH masses also have greater embedded disc populations, resulting in a higher frequency of binary-single scatterings at higher $M_\bullet$. The binary-single time-scale becomes more frequent in the inner portions of the disc, becoming comparable or even shorter than the merger time-scale for our model with $M_\bullet = 10^9$ and $\mathrm{BIMF}^{\gamma=1.7}_{\rm Tagawa}$, therefore our assumption that binary-single scatterings do not affect the merger likelihood in this range of the parameter space breaks down. Across the majority of our parameter space, this remains a safe assumption (note that our other BIMFs assume the same number of initial BHs as $\mathrm{BIMF}^{\gamma=2.35}_{\rm Tagawa}$).

2.7.2 Formation probability

The BH formation function $f_{\rm form}$ is a function of both R according equation (16) but also the anticipated binary mass by equation (20). For a given BH we compute the mass averaged value of the formation function $\langle f_{\rm form}\rangle_{\rm M}$ by sampling over the mass distribution at the radial position of the BH, utilizing the mass distribution of $n_{\rm BH}(R, M_{\rm BH})$. Since smaller objects are less likely to align in the AGN disc, the mass distribution of embedded BHs skews to higher masses compared to the initial distribution.

2.7.3 Comparing to dynamical friction

We will compare our results to another suite of Monte Carlo simulations where $f_{\rm form}$ is determined assuming a simplified gas dynamical friction treatment. There, we incorporate the H. Tagawa et al. (2020a) formation prescription into the Monte Carlo simulations in the calculation of the formation time-scale. Under this prescription, the probability that a BH–BH scattering successfully forms a binary is given by

$$f_{\rm form,Tag} = \min(1, t_{\rm pass}/t_{\rm DF})\,, \tag{26}$$

where $t_{\rm pass}$ is the time taken for the objects to traverse the Hill sphere,

$$t_{\rm pass} = r_{\rm H}/v_{\rm rel}\,, \tag{27}$$

and $t_{\rm DF} = v_{\rm rel}/a_{\rm DF}$ is the gas dynamical friction time-scale from the deceleration given by (E. C. Ostriker 1999):

$$a_{\rm DF}(v_{\rm rel}) = -\frac{4\pi G^2 M_{\rm BH}\rho}{v_{\rm rel}^2} f(v_{\rm rel}/c_{\rm s})\,, \tag{28}$$

$$f(x) = \begin{cases} \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) - x & 0 < x \leq 1\,, \\ \frac{1}{2}\ln(x^2-1) + 3.1 & x > 1\,. \end{cases} \tag{29}$$

The modified formation function in equation (26) is used in equation (24) in the same manner as before, with identical sampling of the relative velocity (Section 2.6).

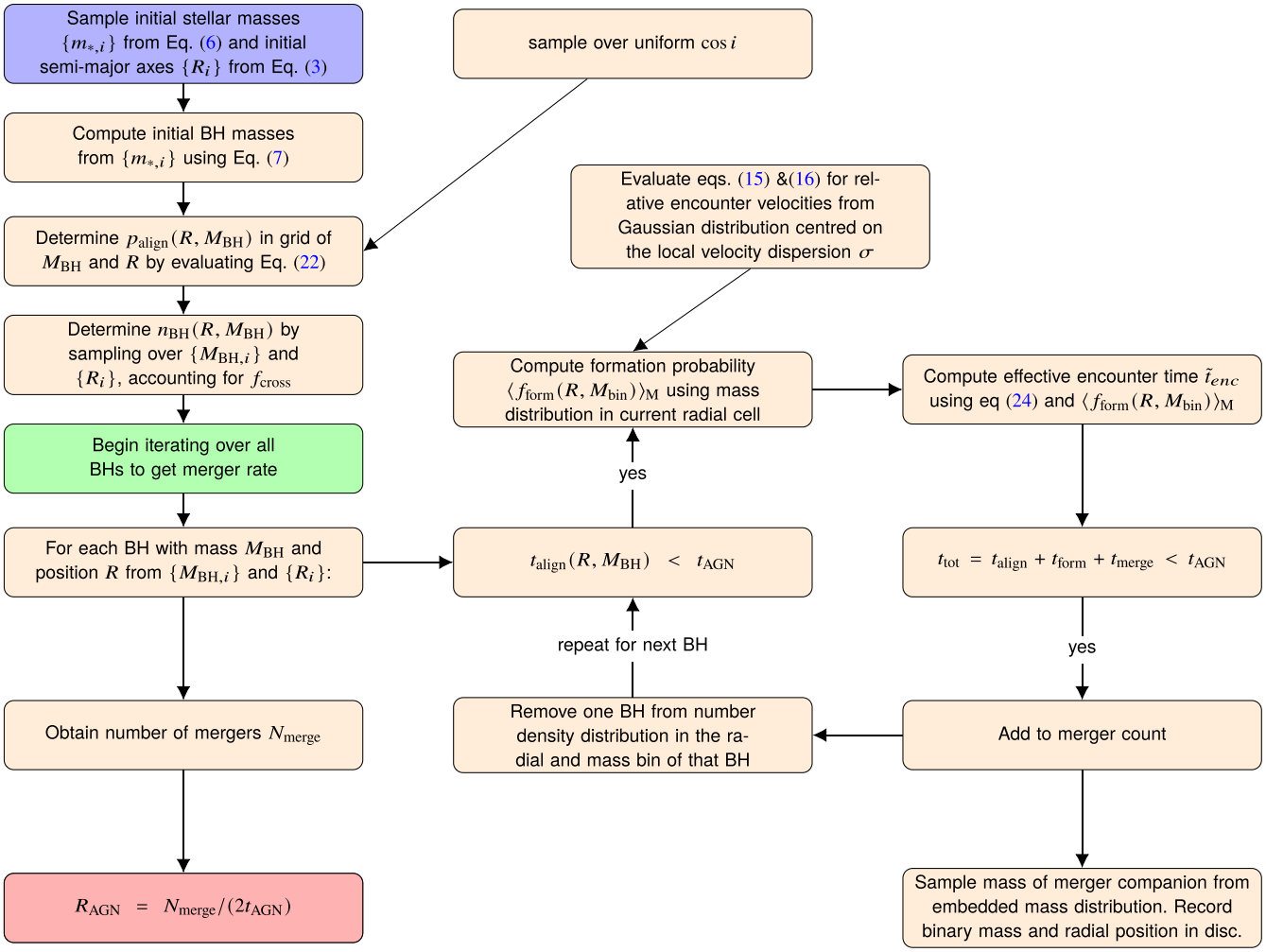

Figure 6. Summary of the fiducial semi-analytic procedure to determine black hole merger rates, from an initial sample of stars in the central stellar cluster.

2.8 Knock on effects

Extending the single BH calculation to a sample across the entire BH population requires accounting for knock on effects from the outcome of each calculation, i.e. was there a merger. To account for the finite number of BHs and the time dependence of $n_{\rm BH}$ the contribution of one BH is removed from the number density for each merger. Specifically, if a BH satisfies the time constraints of equation (25), a random BH merging partner is sampled from the current distribution of masses predicted by $n_{\rm BH}$ and its contribution to the number density is removed. The masses and position in the disc is recorded for all mergers, which is required to compute the anticipated merger rate from GW detectors. Though this is not a formal implementation of the time dependence of $n_{\rm BH}$ and one cannot comment on the change in merger rate over the AGN lifetime, it accounts for the overall reduction in the number of mergers within $t_{\rm AGN}$. Additionally, neglecting this effect would lead to overcounting higher mass binary encounters/mergers, since higher mass BHs have shorter encounter time-scales. From equation (23) there is an overall $M_{\rm bin}$ dependence of $\sim M_{\rm bin}^{-1}$ (assuming $z_{\rm H} = r_{\rm H}$, which is true for the vast majority of R, $M_{\rm bin}$, and $M_\bullet$).

However, the merger rate dependence on this assumption is small, as discussed in Section 3.2. For clarity, a detailed flow chart to summarize the overall Monte–Carlo Scheme is shown in Fig. 6.

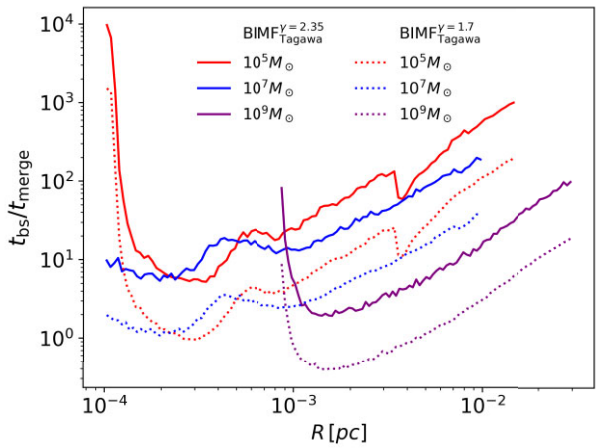

Figure 7. The ratio of the binary-single scattering time-scale $t_{\rm bs}$ to the merger time-scale $t_{\rm merge}$ for various SMBH masses and our fiducial BIMF. The reduction in $t_{\rm bs}$ time-scales between BIMF$^{\gamma=2.35}_{\rm Tagawa}$ and BIMF$^{\gamma=1.7}_{\rm Tagawa}$ stems from the approximately order of magnitude increase in number of initial black holes in the nuclear star cluster.

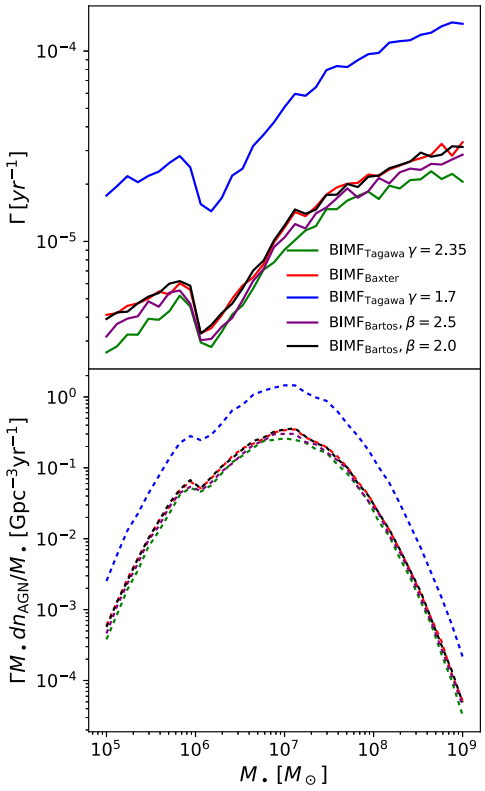

Figure 8. *Top:* The BH merger rate Γ per year per AGN with mass $M_\bullet$ for the SGα models. *Bottom:* The resulting merger rate across $M_\bullet$ weighted by the mass distribution of AGN (equation 30). The different colours represent the assumed BH initial mass function (see Section 2.3). The graph indicates the observable rates should be dominated by AGN with $M_\bullet \sim 10^7\,{\rm M}_\odot$.

3 RESULTS

3.1 $M_\bullet$ dependence

The merger rate *per* AGN per year (Γ) and observable merger rate density distribution (considering only the SGα models for now) over $M_\bullet$ is shown in Fig. 8. The merger rate distribution assumes the AGN number density in the Universe follows the function given by J. E. Greene & L. C. Ho (2007, 2009)

$$\frac{{\rm d}n_{\rm AGN}}{{\rm d}M_\bullet} = \frac{34\,000\ {\rm Gpc^{-3}}}{M_\bullet}\,10^{-(\log_{10}(M_\bullet/M_\odot)-0.67)^2/1.22}. \tag{30}$$

We do not propagate errors associated with this relation as they are insignificant compared to resulting range in merger rates from our models, which span an order of magnitude. The rate of BH mergers is dependent on SMBH mass, ranging from order $\sim 10^{-6}\,{\rm yr^{-1}}$ at $M_\bullet = 10^5\,{\rm M}_\odot$ to $\sim 10^{-4}\,{\rm yr^{-1}}$ at $M_\bullet = 10^9\,{\rm M}_\odot$ for a single AGN. These results are in agreement with the lower range predicted by H. Tagawa et al. (2020a) ($\Gamma \sim 10^{-4} - 10^{-3}{\rm yr^{-1}}$, for $M_\bullet = 10^7\,{\rm M}_\odot$). The rates are around 10–50 times larger than the results of I. Bartos et al. (2017). As I. Bartos et al. (2017) considers only pre-existing binary mergers (i.e. the binaries did not form inside the AGN disc), the steeper dependence on $M_\bullet$ here is a result of the binary formation function and encounter time-scale, which have additional $M_\bullet$ dependence. At low masses, merger rates are restricted by the number of BHs in the system, the lowest number being ~ 100 BHs for $M_\bullet = 10^5\,{\rm M}_\odot$. As $M_\bullet$ increases, the number of available BHs within $R_{\rm inf}$ increases, at a faster rate than $f_{\rm cross}$ can limit the embedded number of BHs. The increased BH population in combination with the AGN mass function results in a peak in the merger rate at $M_\bullet \sim 10^7\,{\rm M}_\odot$. Beyond this, the scarcity of AGN with $M_\bullet > 10^7\,{\rm M}_\odot$ limits the contribution to the merger rate despite $\Gamma(M_\bullet)$ being larger for these more massive AGNs.

Table 1. The fraction of BHs in the Monte Carlo SGα simulation with BIMF$^{\gamma=2.35}_{\rm Tagawa}$ that make it through each stage of the merger pathway. Starting with the objects that align $t_{\rm align} < t_{\rm AGN}$ (F1), then the number of *aligned objects* that form binaries $t_{\rm align} + t_{\rm form} < t_{\rm AGN}$ (F2), and the number of formed objects that merge $t_{\rm align} + t_{\rm form} + t_{\rm merge} < t_{\rm AGN}$ (F3). Indicating that the majority of BHs do not merge due to the alignment time.

$M_\bullet$ ($\rm M_\odot$)	F1 (per cent)	F2 (per cent)	F3 (per cent)
10^5	26	43	> 99
10^7	1.4	65	> 99
10^9	0.21	80	> 99

To investigate the main bottleneck of the merger process, we can consider the fraction of objects that make it through each stage of the merger pathway. Starting with the objects that align $t_{\rm align} < t_{\rm AGN}$ (F1), then the number of *aligned objects* that form binaries $t_{\rm align} + t_{\rm form} < t_{\rm AGN}$ (F2), and the number of formed objects that merge $t_{\rm align} + t_{\rm form} + t_{\rm merge} < t_{\rm AGN}$ (F3). Assuming BIMF$^{\gamma=2.35}_{\rm Tagawa}$, these values are shown in Table 1 for $M_\bullet = \{10^5, 10^7, 10^9\}\,{\rm M}_\odot$.

The relative fractions imply that (per BH) the primary bottleneck is the alignment time (F1) by approximately an order of magnitude therefore constraining the inclination and radial distribution of BHs in the initial distribution is also crucial. Though F1 is the primary bottleneck, the formation time-scale (unlike the merger time-scale) is not negligibly small as a notable fraction of embedded BHs fail to form a binary within $t_{\rm AGN}$.

There is no significant (order of magnitude) difference in the overall merger rates from our fiducial model with BIMF$_{\rm Baxter}$ and BIMF$_{\rm Bartos}$ (Table 2). However we find a strong dependence on γ for BIMF$_{\rm Tagawa}$, with the lower value of $\gamma = 1.7$ leading to a rate increase of about an order of magnitude. This stems from having many BHs from the more top heavy stellar mass distribution ($\sim$5 times more BHs) and the resulting top heavy BIMF compared with BIMF$^{\Gamma=2.35}_{\rm Tagawa}$. The increase in merger rate is driven by a higher embedded BH density $n_{\rm BH}$ since higher mass BHs more easily embed themselves within $t_{\rm AGN}$ and have shorter encounter time-scales $t_{\rm enc} \propto n_{\rm BH}^{-1} M_{\rm bin}^{-1}$. As the range in $M_{\rm BH}$ ($M_{\rm bin}$) for BIMF$_{\rm Tagawa}$ is only $5\,{\rm M}_\odot - 15\,{\rm M}_\odot$ ($10\,{\rm M}_\odot - 30\,{\rm M}_\odot$), this suggests the increase in overall BH number is the dominant factor in the increased rates for $\gamma = 1.7$. Allowing for larger initial BH masses increases the rates slightly (up to a factor of a few, e.g. BIMF$_{\rm Baxter}$) compared to BIMF$^{\gamma=2.35}_{\rm Tagawa}$. This comparison highlights the mass bias definitively as the number of BHs between these two models is fixed. Given the comparatively low number of BHs in the high mass range $> 20\,{\rm M}_\odot$, the increase in the rates by a factor of 2–3 suggests that a high percentage of these BHs will partake in a merger and represent a significant portion of total merging BHs.

3.2 Observable rates

The merger rate density $\mathcal{R}_\rho$ in ${\rm Gpc^{-3}yr^{-1}}$ is given by

$$\mathcal{R}_\rho = \int \Gamma_{\rm AGN}(M_\bullet)\frac{{\rm d}n_{\rm AGN}}{{\rm d}M_\bullet}{\rm d}M_\bullet\,. \tag{31}$$

For our fiducial model BIMF$^{\gamma=2.35}_{\rm Tagawa}$ (SGα), this results in a rate of $0.73\,{\rm Gpc^{-3}\,yr^{-1}}$, more than an order of magnitude lower than the value from current LVK data.

Table 2. Table of results from Monte Carlo runs with different BIMFs and variations of initial conditions. *From left to right:* The assumed disc viscosity model, the formation function, the BIMF function, the merger rate density $\mathcal{R}_\rho$ in Gpc^3 yr^{-1}, the rate density in the binary mass range $5\,M_\odot \leq M_{\rm bin} \leq 20\,M_\odot$, density rate for $20\,M_\odot < M_{\rm bin} \leq 50\,M_\odot$, density rate for $50\,M_\odot < M_{\rm bin} \leq 100\,M_\odot$, ratio of density rates in second to first mass range, predicted detection rate of events for advanced LIGO, $\Gamma_{\rm LIGO}$. The (f) denotes the fiducial model which is used to test the merger rate change when the BH density is assumed to be constant (final row). The ranges of merger rates is shown in row 11/12 when BIMF$^{\gamma=1.7}_{\rm Tagawa}$ is/isn't included and compared to the currently available observed range from LIGO-VIRGO-KAGRA (R. Abbott et al. 2023) in row 13.

Disc	$f_{\rm form}$	BIMF	$\mathcal{R}_\rho$	$\mathcal{R}_\rho$[5–20]	$\mathcal{R}_\rho$[20–50]	$\mathcal{R}_\rho$[50–100]	$\mathcal{R}_\rho$[20–50]/[5–20]	$\Gamma_{\rm LIGO}$ (yr^{-1})
SGα	Rowan	BIMF$^{\gamma=2.35}_{\rm Tagawa}$(f)	0.73	0.55	0.18	–	0.33	2.56
SGα	Rowan	BIMF$^{\gamma=1.7}_{\rm Tagawa}$	3.92	2.26	1.65	–	0.73	16.9
SGα	Rowan	BIMF$_{\rm Baxter}$	0.90	0.23	0.48	0.19	2.09	25.2
SGα	Rowan	BIMF$^{\beta=2.5}_{\rm Bartos}$	0.80	0.34	0.40	0.06	1.18	12.5
SGα	Rowan	BIMF$^{\beta=2.0}_{\rm Bartos}$	0.91	0.24	0.52	0.15	2.16	23.4
SGβ	Rowan	BIMF$^{\gamma=2.35}_{\rm Tagawa}$	1.31	1.00	0.32	–	0.32	4.69
SGβ	Rowan	BIMF$^{\gamma=1.7}_{\rm Tagawa}$	7.1	4.1	3.1	–	0.76	31.7
SGβ	Rowan	BIMF$_{\rm Baxter}$	1.64	0.45	0.87	0.32	1.93	42.0
SGβ	Rowan	BIMF$^{\beta=2.5}_{\rm Bartos}$	1.45	0.64	0.72	0.10	1.13	22.0
SGβ	Rowan	BIMF$^{\beta=2.0}_{\rm Bartos}$	1.63	0.45	0.93	0.25	2.07	37.4
		Range	0.73–7.1	0.23–4.1	0.18–3.1	0.06–0.32	0.32–2.16	2.56–42.0
		Range [w.o BIMF$^{\gamma=1.7}_{\rm Tagawa}$]	0.73–1.64	0.23–0.64	0.18–0.93	0.06–0.32	0.32–2.16	2.56–42.0
		Observed	17.9–44	13.3–39	2.5–6.8	0.1–0.4	0.09–0.29	
SGα	Tagawa	BIMF$^{\gamma=2.35}_{\rm Tagawa}$	0.73	0.56	0.18	–	0.32	2.63
SGα	Tagawa	BIMF$^{\gamma=1.7}_{\rm Tagawa}$	3.9	2.27	1.64	–	0.72	17.2
SGα	Tagawa	BIMF$_{\rm Baxter}$	0.93	0.25	0.49	0.19	1.96	25.2
SGα	Tagawa	BIMF$^{\beta=2.5}_{\rm Bartos}$	0.82	0.35	0.41	0.06	1.17	12.5
SGα	Tagawa	BIMF$^{\beta=2.0}_{\rm Bartos}$	0.89	0.23	0.51	0.14	2.22	23.4
SGα, $n_{\rm BH}$ = const	Rowan	BIMF$^{\gamma=2.35}_{\rm Tagawa}$	0.79	0.56	0.21	–	0.38	2.64

To calculate the rate of BH mergers per year from Earth, we adopt the observational horizon distance $D_{\rm h}$ for Advanced LIGO for which a binary with mass $M_{\rm bin}$ is detectable at a signal-to-noise ratio of 8 (see M. Dominik et al. 2015)

$$D_{\rm h}(M_{\rm bin}) = 0.45\left(\frac{M_{\rm bin}}{2.8\,M_\odot}\right)^{5/6}\,{\rm Gpc}\,. \tag{32}$$

The comoving volume $V_{\rm c}$ in which we can detect a merger with binary mass $M_{\rm bin}$ is given by

$$V_{\rm c}(M_{\rm bin}) = \frac{4}{3}\pi\left(\frac{D_{\rm h}(M_{\rm bin})}{2.26}\right)^3 (1+z)^{-3}\,, \tag{33}$$

where z is redshift, that we neglect for the purpose of this study and set to zero. The observed rate $\Gamma_{\rm LIGO}$ per year is then

$$\Gamma_{\rm LIGO} = \iint V_{\rm c}(M_{\rm bin})\frac{{\rm d}R_{\rm AGN}(M_{\rm bin}, M_\bullet)}{{\rm d}M_{\rm bin}}\frac{{\rm d}n_{\rm AGN}}{{\rm d}M_\bullet}{\rm d}M_{\rm bin}\,{\rm d}M_\bullet\,. \tag{34}$$

The merger rate per binary mass term, ${\rm d}R_{\rm AGN}/{\rm d}M_{\rm bin}$ is evaluated using the merging binary mass distribution from each $M_\bullet$ value put through our model. For lower $M_\bullet$ masses that have a low merger number, we repeat the analysis until we have a merger sample of at least 1000 BHs. We display the merger rate densities, LIGO merger rate and the merger rates for binary masses in the ranges $5 \leq M_{\rm bin} \leq 20$, $20 < M_{\rm bin} \leq 50$, and $50 < M_{\rm bin} \leq 100$ in Table 2 for all three BIMFs and varying initial conditions. The varied initial conditions include: maintaining a fixed BH density over $t_{\rm AGN}$ and a rerun of the entire SGα suite assuming the binary formation criterion of H. Tagawa et al. (2020a) (equation 26).

Our models give us a merger rate of $\Gamma_{\rm LIGO} = 2.56$–$42.0\,{\rm yr}^{-1}$ with a local density rate of $0.73 - 7.1\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}$. While not within the range of 17.9–$44\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}$ from LIGO-VIRGO-KAGRA (R. Abbott et al. 2023), it suggests the contribution from the AGN channel is non-negligible. The merger rate increases by approximately a factor 2 going from the SGα to SGβ discs. This is a direct result of an increased gas density in the inner radial regions of $M_\bullet = 10^7\,M_\odot - 10^9\,M_\odot$ (see Fig. 2), which reduces the alignment time (equation 22) and the effective formation time-scale. We find the merger rate is *highly* dependent on the number of BHs and the most sensitive parameter to the rates across the simulations in this work. Hence, we encourage further constraint of the expected number of BHs within $R_{\rm inf}$ in the Universe. When larger initial BH masses are permitted in the initial distribution, as in BIMF$_{\rm Baxter}$, the rates become dominated by larger binary masses ($M_{\rm bin} > 20\,M_\odot$), despite BHs of masses $> 10\,M_\odot$ being fewer in number. This result remains consistent across the SGα, SGβ models, and the two different formation functions. This hardening of the merging mass function in AGN was found in Y. Yang et al. (2019), in the context of pre-existing binary mergers, driven by migration traps and mass biased alignment time. Here, we again have the biased alignment time ($t_{\rm align} \sim M_{\rm BH}^{-1}$, equation 22). In addition to this, larger BHs have a considerably shorter BBH formation time-scale since their Hill radius is larger and $\langle f_{\rm form}\rangle_{\rm M}$ is far larger. Hence high mass mergers could make up a significant fraction of the anticipated rates in the AGN channel, as shown by Table 2. While we have adopted a physically well-motivated formation criterion, large uncertainties still remain in many aspects of this and other population studies of AGN merging

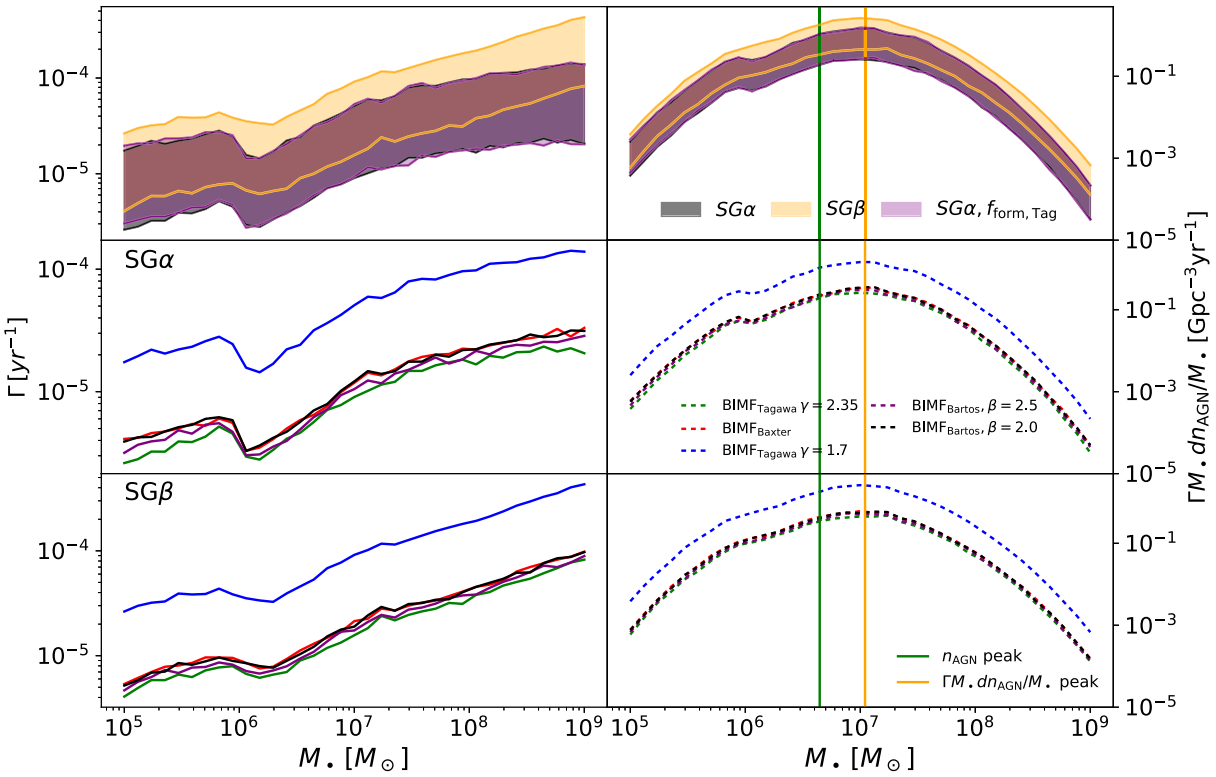

Figure 9. *Left column:* The BH merger rate Γ per year per AGN with mass $M_\bullet$. *Right column:* The merger rate across $M_\bullet$ weighted by the mass distribution of AGN (equation 30). The vertical green and yellow lines indicate the peak of the AGN mass function and the rate distribution over $M_\bullet$, respectively. *Top row:* The range in rates across all BIMFs for each disc viscosity prescription (SGα, SGβ) and simplified formation function ($f_{\rm form,Tag}$). *Middle row:* the rates for the *SG*α disc model. *Bottom row:* the rates for the *SG*β disc model. The graph indicates the observable rates should be dominated by AGN with $M_\bullet \sim 10^7\,{\rm M}_\odot$. The SG$\beta$ discs show a greater merger rate, particularly in the high $M_\bullet$ range.

populations. This includes the BIMF, binary-single scatterings, the effects of turbulence in the AGN disc, AGN lifetime among others. Therefore, the true uncertainty in the rates/merger statistics from the AGN channel is higher than the reported figures of this work and others. None the less, this does not affect the primary goal in this work of assessing their sensitivity to the binary formation likelihood.

The values of Γ and $\mathcal{R}_\rho$ as a function of $M_\bullet$ for our SGα, SGβ, and test suite with the $f_{\rm form,Tag}$ and SGα assumption are compared in Fig. 9. Between the two disc models, the SGβ discs show an increased merger rate across $M_\bullet$, as anticipated from the lower viscosity and higher density. The merger rates diverge further from the SGα models as $M_\bullet$ increases, stemming from the boundary of the radiation zone ($\beta \ll 1$) moving further out in R, covering a larger portion of the domain. The rates calculated assuming $f_{\rm form,Tag}$ remain nearly identical to the original SGα models across the range in $M_\bullet$.

3.3 Merger properties

We show the distributions of mass, mass ratio, and radial position of merging binaries generated by our model averaged over all $M_\bullet$, weighted by the AGN mass function in Figs 10, 11, and 12, respectively.

The mass distribution reaffirms the aforementioned mass hardening effect of the AGN channel. For example, consider BIMF$_{\rm Baxter}$, which demonstrates a near flat profile in M_1 compared to the original profile $\sim M_{\rm BH}^{-2.2}$. Considering the two simple power laws of BIMF$_{\rm Bartos}$, we find the exponent for the power-law distribution of M_1 is flattened by a factor $\Delta\xi \simeq 1.1-1.2$ compared to the BIMF, i.e $M_1 \propto M_{\rm BH}^{\xi_0+\Delta\xi}$, where ξ_0 is the exponent of the BIMF. This relatively good agreement with the value $\Delta\xi \simeq 1.3$ predicted by Y. Yang et al. (2019). This is also reflected in the q distribution. For flatter BIMFs with higher limits on $M_{\rm BH}$ (i.e. BIMF$_{\rm Baxter}$, BIMF$_{\rm Bartos}$), the q distribution becomes less steep and can even flip to favour higher mass ratio mergers, consistent with the recent findings of V. Delfavero et al. (2025) who account for additional physics (migration, repeated mergers). Qualitatively, this means less numerous high mass BHs can very easily encounter an merge with abundant low mass BHs. This has optimistic prospects for second (or more) generation mergers as the merged BH can more easily repeat the formation and merger process thanks to its larger mass, potentially further hardening the merging mass distribution. Such a scenario could easily explain highly unequal mass ratio or massive binary mergers such as GW190814 and GW190521 (e.g. R. Abbott et al. 2020b, c). Comparing the rates in the three mass bins of Table 2, we find mergers in AGN may account for $1.7-40$ per cent mergers in the range $5\,{\rm M}_\odot < M_{\rm bin} \leq 20\,{\rm M}_\odot$, $\sim 3-100$ per cent mergers in the range $20\,{\rm M}_\odot < M_{\rm bin} \leq 50\,{\rm M}_\odot$, and $\sim 15-100$ per cent of mergers in $20\,{\rm M}_\odot < M_{\rm bin} \leq 50\,{\rm M}_\odot$. Therefore within our assumptions, AGN may not be the primary source of observed low mass BH mergers, but could account for a large fraction of high mass mergers. Hence, the relative contribution of the AGN channel could potentially be constrained using the ratio of merger rates in different mass bins. Using the ratio between the first and second mass bins (Table 2), we find a preference for a higher ratio in the AGN channel than the observed rate. This would again indicate a large contribution to the rates from other channels for low mass mergers with a steeper

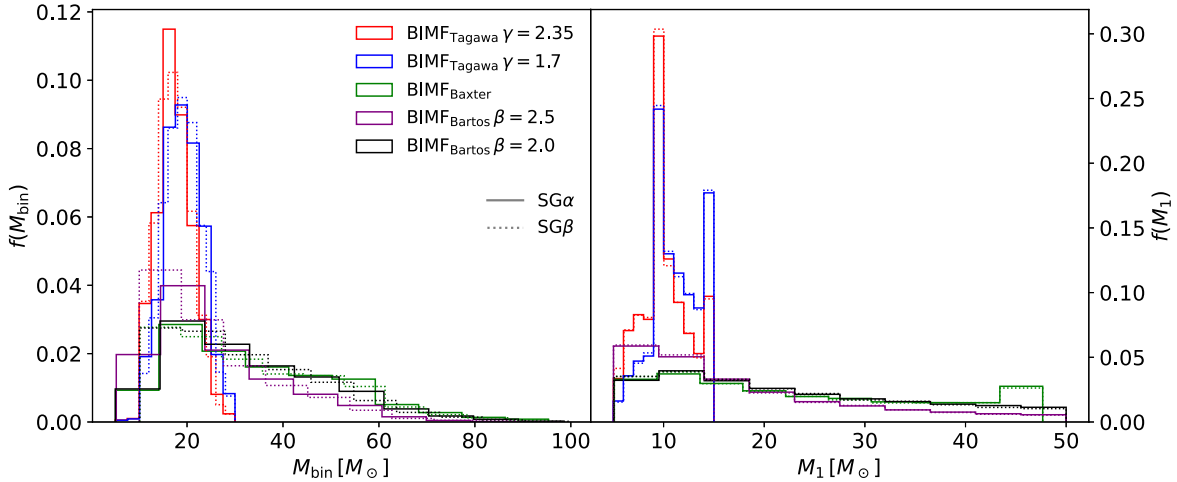

Figure 10. *Left:* Mass distribution of the merging binary mass $M_{\rm bin}$ for each BIMF outlined in Section 2.3, represented by different colours. *Right:* Mass distribution of the primary black hole mass M_1 of mergers. The results show a significant hardening of the merging BH mass function compared to the BIMF, indicating larger black holes have a much greater chance to form binaries and merge.

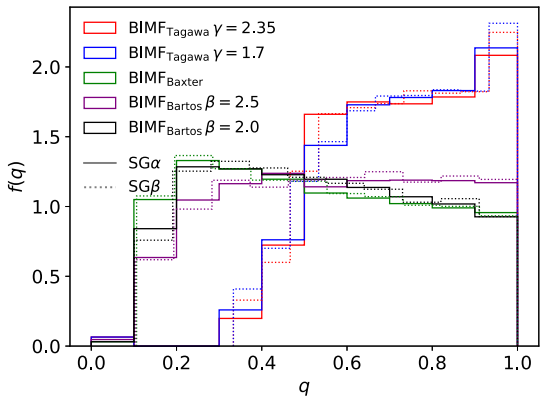

Figure 11. Mass ratio distribution $q = M_2/M_1$ for our merging binaries for each BIMF (colour coded). Demonstrating the AGN channel can easily produce high mass ratio mergers.

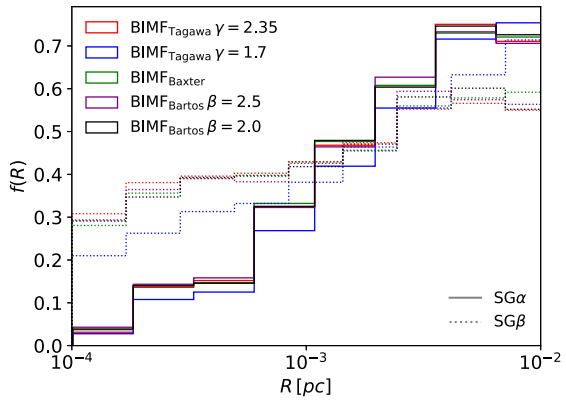

Figure 12. Distribution of radial positions in the disc R for our merging binaries for each BIMF (colour coded). Merger rate peaks is greatest around the outer radial bound of the simulations, outside of which the AGN disc becomes gravitationally unstable under the Sirko-Goodman model.

mass distribution. As more GW detections are made and the BIMF in AGN better constrained, it could then be possible to constrain more rigorously the relative contribution from the AGN channel from the merging mass distribution.

The radial distribution of mergers (Fig. 12) varies less than an order of magnitude across R, peaking at the outer radial bound of the simulations. At lower radii, we are limited by a low number of BHs in each cylindrical volume and increased Keplerian shear $\sigma_{\rm Kep}$. The positive slope of the curve is consistent with H. Tagawa et al. (2020a), who find the majority of gas-capture encounters occur in the outer regions of the AGN disc. Note we have neglected the possibility of repeated mergers and radial migration, which we expect will change this radial profile in reality.

4 SUMMARY AND CONCLUSIONS

In this work, we predicted the merger rate of BBHs in AGN formed via BH–BH scatterings in an AGN disc using the physically motivated formation prescription based on C. Rowan et al. (2024) (see also H. Whitehead et al. 2024a). The primary goal was to test whether implementing the prescription, derived from high resolution hydrodynamical simulations, alters the rates significantly compared to simplified dynamical friction models. Using a range of initial black hole mass functions, our models yield merger rate densities ranging from $0.73 - 7.1\,{\rm Gpc^{-3}\,pc^{-1}}$. These rates corroborate the range predicted from many analytical studies (e.g. I. Bartos et al. 2017; Y. Yang et al. 2019; H. Tagawa et al. 2020a; B. McKernan et al. 2020b; V. Delfavero et al. 2025). We note that V. Delfavero et al. (2025) find the merger rate peaks at higher $M_\bullet$ and posit this could stem from their assumption that BHs that pass within $r_{\rm H}$ will reliably become bound independently of $M_\bullet$, where we find the formation function drops by approximately an order of magnitude in the range $10^7 - 10^9\,{\rm M_\odot}$. Therefore we encourage more detailed studies in the future to include a formation function when simulating BH scatterings in AGN discs. We find that the rates are very similar (within 10 per cent) when the gas-capture process is modelled using an analytic dynamical friction treatment of the gas-assisted binary formation process, affirming the results of earlier studies that use this simplified model.

The mass distribution (both the binary and primary mass) of merging BHs is significantly more top heavy compared to the initial BH mass distribution, due to a favourability for high mass BHs to align with the disc and form binaries via gas-dissipation. Therefore the merger rates and masses are sensitive to the assumed BIMF, where more top heavy BIMFs lead to increased rates. This bias also leads to a more flat mass ratio distribution, thus the AGN channel can easily explain the high mass and high mass ratio detections from GW observatories. This sensitivity highlights the need to better constrain the BIMF, which may be different to what is expected from a typical stellar mass function (e.g. A. S. Jermyn et al. 2022; A. J. Dittmann, A. S. Jermyn & M. Cantiello 2023). Mergers are more numerous closer to the gravitationally unstable region of the AGN disc (higher R), although the distribution becomes flatter when a pressure-dependent disc viscosity is assumed. We find the overall merger rate to lie within $\lesssim 5-40$ per cent that of LIGO-VIRGO-KAGRA, suggesting a non-negligible contribution from the AGN channel. However, due to the top heavy merging mass function, the AGN channel can potentially be the source of $\sim 3-100$ per cent mergers in the range $20\,\mathrm{M}_\odot < M_{\rm bin} \leq 50\,\mathrm{M}_\odot$ and $\sim 15-100$ per cent of mergers in $50\,\mathrm{M}_\odot < M_{\rm bin} \leq 100\,\mathrm{M}_\odot$. Due to the mass bias of the gas-assisted binary formation mechanism, we encourage future studies to account for repeated mergers, which may further increase the merger rate for $M_{\rm bin} > 30\,\mathrm{M}_\odot$. We posit that reductions in the observed merger rate uncertainty and better constraints of the BIMF in AGN could allow us to constrain the relative contribution from AGN using the relative merger rates from low and high mass binaries.

Though we have used a well-motivated formation criterion for our binaries, our model is still subject to several large assumptions. The formation function assumes that the hydrodynamics of the encounter are isothermal. Our recent 2D non-isothermal hydrodynamical work H. Whitehead et al. (2024b) has shown that this is an oversimplification, where gas heating can be significant during the encounter, reducing the gas mass in the Hill sphere and potentially reducing the typical value of $f_{\rm form}$. We have pessimistically assumed all BHs are formed in the spherical stellar cluster and the merging systems must first align with the disc. In reality the parent stars could align first (and faster) before the BH is born (T. Panamarev et al. 2018). We assume there is no migration in the disc, which could speed up the rate of encounters. Binary-single scatterings with a tertiary BH or star that could potentially ionize the BBH or induce a merger are also ignored, alongside BBH formations from three-body scatterings. Recently, hydrodynamical simulations of binary-single scatterings in AGN discs have shown that gas can shrink the scale of the triple system (e.g. C. Rowan et al. 2025; M. Wang, Q. Wu & Y. Ma 2025a; M. Wang et al. 2025b), typically leaving the remaining binary harder than prior to the encounter with a single. In many cases, embedded binary-single scatterings are shown to induce the prompt merger of two of the BHs during the triple interaction on the scale of $\sim 10^2$ yr compared with the isolated inspiral time-scale of $\sim 10^5$ yr assumed in this study. Therefore whether binary-single scatterings help or hinder the merger rate is an open question. Binary-single interactions with stars are also not considered in this work as we do not track the stellar population, the effects of which may also help or hinder an inspiralling BH binary.

Another uncertainty lies in the length and frequency of an AGN duty cycle, which may also have a dependence on the SMBH mass. If the lifetime of AGN is longer, then this leads to an increase in the total number of mergers per AGN, however the average rate across it's lifetime is likely to decrease (e.g. V. Delfavero et al. 2025). Given the number of AGN in the sky is fixed, longer AGN lifetimes would lead to a decrease in the overall merger rate anticipated from the AGN channel. Furthermore, there is some observational evidence for a mass bias in the AGN lifetime (e.g. A. Merloni 2004; P. Guetzoyan et al. 2025), with more massive $M_\bullet > 10^7\,\mathrm{M}_\odot$ AGN having shorter lifetimes. This may shift the peak of our merger rate distribution (Fig. 9) to a higher $M_\bullet$. Given the large uncertainties at various stages of the AGN channel, we no doubt expect the true uncertainty of the merger rates to be higher than the range given by our models.

Our results indicate that the number of BHs produced/present around the AGN disc is the most influential parameter on the BH rates and therefore encourage future projects to constrain this value and its possible dependence on $M_\bullet$ itself. We conclude that BHs merging in AGN through the gas capture mechanism are a non-negligible contributor to the observed rates and a potentially dominant channel for high mass and comparable binary component mass mergers.

ACKNOWLEDGEMENTS

(i) This work was supported by the Science and Technology Facilities Council Grant Number ST/W000903/1.

DATA AVAILABILITY

The data used to create this work will be shared upon reasonable request.

REFERENCES

Abbott B. P. et al., 2016, Phys. Rev. Lett., 116, 061102
Abbott B. P. et al., 2019a, Phys. Rev. X, 9, 031040
Abbott B. P. et al., 2019b, ApJ, 882, L24
Abbott B. P. et al., 2020, ApJ, 892, L3
Abbott R. et al., 2020a, Phys. Rev. D, 102, 043015
Abbott R. et al., 2020b, Phys. Rev. Lett., 125, 101102
Abbott R. et al., 2020c, ApJ, 896, L44
Abbott R. et al., 2021, SoftwareX, 13, 100658
Abbott R. et al., 2023, Phys. Rev. X, 13, 011048
Antonini F., Perets H. B., 2012, ApJ, 757, 27
Antonini F., Rasio F. A., 2016, ApJ, 831, 187
Bartko H. et al., 2009, ApJ, 697, 1741
Bartos I., Kocsis B., Haiman Z., Márka S., 2017, ApJ, 835, 165
Baruteau C., Cuadra J., Lin D. N. C., 2011, ApJ, 726, 28
Baxter E. J., Croon D., McDermott S. D., Sakstein J., 2021, ApJ, 916, L16
Belczynski K., Dominik M., Bulik T., O'Shaughnessy R., Fryer C., Holz D. E., 2010, ApJ, 715, L138
Belczynski K. et al., 2016, A&A, 594, A97
Bellovary J. M., Mac Low M.-M., McKernan B., Ford K. E. S., 2016, ApJ, 819, L17
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton Univ. Press, Princeton
Boekholt T. C. N., Rowan C., Kocsis B., 2023, MNRAS, 518, 5653
Cavaliere A., Padovani P., 1989, ApJ, 340, L5
DeLaurentiis S., Epstein-Martin M., Haiman Z., 2023, MNRAS, 523, 1126
Delfavero V. et al., 2025, ApJ, 989, 67
Dempsey A. M., Li H., Mishra B., Li S., 2022, ApJ, 940, 155
Di Carlo U. N. et al., 2020, MNRAS, 498, 495
Dittmann A. J., Jermyn A. S., Cantiello M., 2023, ApJ, 946, 56
Dodici M., Tremaine S., 2024, ApJ, 972, 193
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shaughnessy R., 2012, ApJ, 759, 52
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shaughnessy R., 2013, ApJ, 779, 72
Dominik M. et al., 2015, ApJ, 806, 263
Event Horizon Telescope Collaboration 2019, ApJ, 875, L6
Event Horizon Telescope Collaboration 2022, ApJ, 930, L12

Fabj G., Nasim S. S., Caban F., Ford K. E. S., McKernan B., Bellovary J. M., 2020, MNRAS, 499, 2608
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Ford K. E. S., McKernan B., 2022, MNRAS, 517, 5827
Gaia Collaboration 2024, A&A, 686, L2
Gangardt D., Trani A. A., Bonnerot C., Gerosa D., 2024, MNRAS, 530, 3689
Gebhardt K. et al., 2000, ApJ, 539, L13
Greene J. E., Ho L. C., 2007, ApJ, 667, 131
Greene J. E., Ho L. C., 2009, ApJ, 704, 1743
Guetzoyan P., Aird J., Georgakakis A., Coil A. L., Barlow-Hall C., Hickox R. C., Rankine A. L., Terrazas B. A., 2025, MNRAS, 536, 79
Gültekin K. et al., 2009, ApJ, 698, 198
Haehnelt M. G., Rees M. J., 1993, MNRAS, 263, 168
Hailey C. J., Mori K., Bauer F. E., Berkowitz M. E., Hong J., Hord B. J., 2018, Nature, 556, 70
Haiman Z., Kocsis B., Menou K., 2009, ApJ, 700, 1952
Jermyn A. S., Dittmann A. J., McKernan B., Ford K. E. S., Cantiello M., 2022, ApJ, 929, 133
Keshet U., Hopman C., Alexander T., 2009, ApJ, 698, L64
Kochanek C. S., 2015, MNRAS, 446, 1213
Kocsis B., Levin J., 2012, Phys. Rev. D, 85, 123005
Kocsis B., Yunes N., Loeb A., 2011, Phys. Rev. D, 84, 024032
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kroupa P., 2001, MNRAS, 322, 231
Li J., Lai D., Rodet L., 2022, ApJ, 934, 154
Li J., Dempsey A. M., Li H., Lai D., Li S., 2023, ApJ, 944, L42
Li R., Lai D., 2023, MNRAS, 522, 1881
Li R., Lai D., 2024, MNRAS, 529, 348
Li Y.-P., Dempsey A. M., Li S., Li H., Li J., 2021, ApJ, 911, 124
Li Y.-P., Dempsey A. M., Li H., Li S., Li J., 2022, ApJ, 928, L19
Lipunov V. M., Postnov K. A., Prokhorov M. E., 1997, Astron. Lett., 23, 492
Liu B., Lai D., 2021, MNRAS, 502, 2049
Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2006, A&A, 455, 1165
Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2007, A&A, 469, 807
Mapelli M. et al., 2021, MNRAS, 505, 339
McKernan B., Ford K. E. S., O'Shaugnessy R., Wysocki D., 2020a, MNRAS, 494, 1203
McKernan B., Ford K. E. S., O'Shaughnessy R., 2020b, MNRAS, 498, 4088
Merloni A., 2004, MNRAS, 353, 1035
Miller M. C., Hamilton D. P., 2002, MNRAS, 330, 232
Mouri H., Taniguchi Y., 2002, ApJ, 566, L17
Naoz S., Rose S. C., Michaely E., Melchor D., Ramirez-Ruiz E., Mockler B., Schnittman J. D., 2022, ApJ, 927, L18
O'Leary R. M., Kocsis B., Loeb A., 2009, MNRAS, 395, 2127
Öpik E., 1924, Publ. Tartu Astrofiz. Obs., 25, 1
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
Panamarev T., Shukirgaliyev B., Meiron Y., Berczik P., Just A., Spurzem R., Omarov C., Vilkoviskij E., 2018, MNRAS, 476, 4224
Rodriguez C. L., Morscher M., Pattabiraman B., Chatterjee S., Haster C.-J., Rasio F. A., 2015, Phys. Rev. Lett., 115, 051101
Rodriguez C. L., Chatterjee S., Rasio F. A., 2016, Phys. Rev. D, 93, 084029
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
Rowan C., Whitehead H., Fabj G., Saini P., Kocsis B., Pessah M., Samsing J., 2025, MNRAS, 539, 1501
Rozner M., Perets H. B., 2022, ApJ, 931, 149
Rozner M., Generozov A., Perets H. B., 2023, MNRAS, 521, 866
Salpeter E. E., 1955, ApJ, 121, 161
Samsing J., D'Orazio D. J., 2018, MNRAS, 481, 5445
Sari R., Goldreich P., 2004, ApJ, 606, L77
Secunda A., Bellovary J., Mac Low M.-M., Ford K. E. S., McKernan B., Leigh N. W. C., Lyra W., Sándor Z., 2019, ApJ, 878, 85
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Sirko E., Goodman J., 2003, MNRAS, 341, 501
Stone N. C., Metzger B. D., Haiman Z., 2017, MNRAS, 464, 946
Tagawa H., Saitoh T. R., Kocsis B., 2018, Phys. Rev. Lett., 120, 261101
Tagawa H., Haiman Z., Kocsis B., 2020a, ApJ, 898, 25
Tagawa H., Haiman Z., Bartos I., Kocsis B., 2020b, ApJ, 899, 26
Tagawa H., Kocsis B., Haiman Z., Bartos I., Omukai K., Samsing J., 2021, ApJ, 907, L20
Thompson T. A. et al., 2019, Science, 366, 637
Tokovinin A. A., 2000, A&A, 360, 997
Tremaine S. et al., 2002, ApJ, 574, 740
Vaccaro M. P., Mapelli M., Périgois C., Barone D., Artale M. C., Dall'Amico M., Iorio G., Torniamenti S., 2024, A&A, 685, A51
Venumadhav T., Zackay B., Roulet J., Dai L., Zaldarriaga M., 2020, Phys. Rev. D, 101, 083030
Wang M., Wu Q., Ma Y., 2025a, preprint (arXiv:2507.07715)
Wang M., Ma Y., Li H., Wu Q., Li Y.-P., Lei X., Wu J., 2025b, ApJ, 983, 114
Whitehead H., Rowan C., Boekholt T., Kocsis B., 2024a, MNRAS, 531, 4656
Whitehead H., Rowan C., Boekholt T., Kocsis B., 2024b, MNRAS, 533, 1766
Yang Y., Bartos I., Haiman Z., Kocsis B., Márka Z., Stone N. C., Márka S., 2019, ApJ, 876, 122

This paper has been typeset from a TEX/LATEX file prepared by the author.