Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **510,** 531–550 (2022) https://doi.org/10.1093/mnras/stab3411
Advance Access publication 2021 November 26

Dynamical friction modelling of massive black holes in cosmological simulations and effects on merger rate predictions

Nianyi Chen [1]★ Yueying Ni,[1] Michael Tremmel [2] Tiziana Di Matteo,[1,3,4] Simeon Bird,[5] Colin DeGraf[1] and Yu Feng[6]

[1]*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[2]*Astronomy Department, Yale University, PO Box 208120, New Haven, CT 06520, USA*
[3]*NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[4]*OzGrav-Melbourne, Australian Research Council Centre of Excellence for Gravitational Wave Discovery, Melbourne, VIC 3122, Australia*
[5]*Department of Physics and Astronomy, University of California Riverside, Riverside, CA 90217, USA*
[6]*Berkeley Center for Cosmological Physics and Department of Physics, University of California, Berkeley, CA 94720, USA*

Accepted 2021 November 17. Received 2021 November 16; in original form 2021 April 12

ABSTRACT

In this work, we establish and test methods for implementing dynamical friction (DF) for massive black hole pairs that form in large volume cosmological hydrodynamical simulations that include galaxy formation and black hole growth. We verify our models and parameters both for individual black hole dynamics and for the black hole population in cosmological volumes. Using our model of DF from collisionless particles, black holes can effectively sink close to the galaxy centre, provided that the black hole's dynamical mass is at least twice that of the lowest mass resolution particles in the simulation. Gas drag also plays a role in assisting the black holes' orbital decay, but it is typically less effective than that from collisionless particles, especially after the first billion years of the black hole's evolution. DF from gas becomes less than 1 per cent of DF from collisionless particles for BH masses $>10^7$ $M_\odot$. Using our best DF model, we calculate the merger rate down to $z = 1.1$ using an $L_{\rm box} = 35$ Mpc h^{-1} simulation box. We predict $\sim$2 mergers per year for $z > 1.1$ peaking at $z \sim 2$. These merger rates are within the range obtained in previous work using similar resolution hydrodynamical simulations. We show that the rate is enhanced by factor of $\sim$2 when DF is taken into account in the simulations compared to the no-DF run. This is due to $>$40 per cent more black holes reaching the centre of their host halo when DF is added.

Key words: gravitational waves – methods: numerical – quasars: supermassive black holes.

1 INTRODUCTION

Supermassive black holes (SMBHs) are known to exist at the centre of the majority of massive galaxies (e.g. Soltan 1982; Kormendy & Richstone 1995; Magorrian et al. 1998; Kormendy & Ho 2013). As these galaxies merge (e.g. Lacey & Cole 1993; Lotz et al. 2011; Rodriguez-Gomez et al. 2015), the SMBHs that they host also go through mergers, resulting in the mass growth of the SMBH population (e.g. Begelman, Blandford & Rees 1980). SMBH mergers following their host galaxy mergers become an increasingly important aspect of SMBH growth for more massive black holes (BHs) in dense environments (e.g. Kulier et al. 2015). As a by-product of BH mergers, gravitational waves are emitted, and their detection opens up a new channel for probing the formation and evolution of early BHs in the Universe (e.g. Sesana, Volonteri & Haardt 2007a; Barausse 2012).

The gravitational wave detection by the Laser Interferometer Gravitational-Wave Observatory (LIGO; Abbott et al. 2016) proves the experimental feasibility of using gravitational waves for studying BH binaries. While LIGO cannot detect gravitational waves from binaries more massive than $\sim$100 $M_\odot$ (Mangiagli et al. 2019), long-baseline experiments are being planned for detections of more massive BH binaries. Specifically, the upcoming *Laser Interferometer Space Antenna* (*LISA*; Amaro-Seoane et al. 2017) mission will be sensitive to low-frequency (10^{-4}–10^{-1} Hz) gravitational waves from the coalescence of massive black holes (MBHs) with masses 10^4–10^7 $M_\odot$ up to $z \sim 20$. At even lower frequencies pulsar timing arrays (PTAs) are already collecting data and the Square Kilometre Array (SKA) in the next decade will be a major leap forward in sensitivity. PTA observations are likely to identify a number of continuous-wave sources representing the early inspiral phase of MBH binaries. PTAs experiments (e.g. Jenet et al. 2004, 2005) may also detect the inspiral of tight MBH binaries with mass $>10^8$ $M_\odot$. While massive BH binaries are the primary sources for PTAs and *LISA*, these two experiments probe different stages of massive BH evolution. PTAs are most sensitive to the early inspiral (orbital periods of years or longer) of nearby ($z < 1$) (massive) sources (Mingarelli et al. 2017). In contrast, *LISA* is sensitive to the inspiral, merger, and ringdown of MBH binaries at a wide range of redshifts (Amaro-Seoane et al. 2012). The two populations of MBH binaries probed by PTAs and *LISA* are linked via the growth and evolution of SMBH across cosmic time.

LISA will provide a unique way of probing the high-redshift Universe and understanding the early formation of the SMBHs, especially when combined with the soon-to-come observations of the

★ E-mail: nianyic@andrew.cmu.edu

electromagnetic (EM) counterparts (Natarajan et al. 2017; DeGraf & Sijacki 2020). For instance, they will potentially allow us to distinguish between different BH seeding mechanisms at high redshift (Ricarte & Natarajan 2018), to obtain information on the dynamical evolution of MBHs (Bonetti et al. 2019), and to gain information about the gas properties within the accretion disc (Derdzinski et al. 2019).

To properly analyse the upcoming results from the gravitational wave and the EM observations, we need to gain a thorough understanding of the physics of these MBH mergers with theoretical tools and be able to make statistical predictions on the binary population. In particular, it is important that the BH dynamics is modelled accurately, so that we can minimize the degeneracy with other physical properties of the merger, and gain accurate information about when and where BH coalescence is expected.

Hydrodynamical cosmological simulations provide a natural ground for studying the evolution and mergers of MBHs. In particular, large-volume cosmological simulations (e.g. Hirschmann et al. 2014; Vogelsberger et al. 2014; Schaye et al. 2015; Feng et al. 2016; Volonteri et al. 2016; Pillepich et al. 2018; Davé et al. 2019) have the statistical power to make merger rate predictions for the upcoming observations.

In order to accurately predict when BH mergers occur in these simulations, one must account for the long journey of the central BHs after the merger of their host galaxies: during galaxy mergers, the central SMBHs are usually separated by as much as a few tens of kpc. These SMBHs then gradually lose their orbital energy and sink to the centre of the new galaxy due to the dynamical friction (DF) exerted by the gas, stars, and dark matter around them (e.g. Chandrasekhar 1943; Ostriker 1999). When their separation reaches the sub-parsec scale, they form a binary and other energy-loss channels begin to dominate, such as scattering with stars (e.g. Quinlan 1996; Sesana, Haardt & Madau 2007b; Vasiliev, Antonini & Merritt 2015), gas drag from the circumbinary disc (e.g. Haiman, Kocsis & Menou 2009), or three-body scattering with a third BH (e.g. Bonetti et al. 2018).

However, due to limited mass and spatial resolution, large-scale cosmological simulations cannot feasibly include detailed treatment of the BH binary dynamics. Without any additional correction to the BH dynamics, the smoothed-away small-scale gravity prevents effective orbital decay of the BH after the orbit approaches the gravitational softening length. Once the binary reaches the innermost region of the remnant galaxy, the gravitational potential (close to the resolution limit) can be noisy. Such a noisy potential can scatter the BH around within the host galaxy, or in some cases even kick the BH to the outskirts of the galaxy if the BH mass is small. To avoid unexpected scattering of the BHs around the centre of the galaxy, large-volume cosmological simulations usually resort to pinning the BHs at the halo minimum potential (also known as repositioning). This repositioning algorithm has the undesirable effect of making the BHs merge rather efficiently once they reach the centre of the galaxy. Post-processing techniques have been used (e.g. Salcido et al. 2016; Kelley, Blecha & Hernquist 2017; Katz et al. 2020; Volonteri et al. 2020) to account for the additional DF effects on scales close to the gravitational smoothing scales of the BHs. This allows for an approximate estimation of the expected delay in the BH mergers. The post-processing calculations are mostly based on idealized analytical models, and therefore do not account for the variety of individual BH environments.

Because of the increased merger efficiency induced by BH repositioning and the limits of post-processing in DF calculations, emerging works have been adding subgrid modelling of DF self-consistently in cosmological simulations and removing the artificial repositioning approximation. Chapon, Mayer & Teyssier (2013) and Dubois et al. (2014) are the first large simulations to include the DF from gas, while Hirschmann et al. (2014) and Tremmel et al. (2017) account for DF from collisionless particles, and both have shown success in stabilizing the BHs at the halo centres. The DF modelling and its effect on the BH merger time-scale have been well tested in Tremmel et al. (2015) and Pfister et al. (2019) in the context of their relatively high-resolution simulations in a controlled single-halo environment, but they have also pointed out the failure of their model when the dark matter particle mass exceeds the BH mass, and so their models might not be directly applicable to lower resolution cosmological simulations. In the context of low-resolution cosmological simulations, the DF modelling is less well tested, and its effects on the BH evolution and merger rate are not fully explored.

In this work, we carefully develop and test the subgrid modelling of DF from both gas and collisionless particles in the context of cosmological simulations with resolution similar to the aforementioned large-volume, low-resolution hydrodynamical simulations (i.e. with a spatial resolution of $\sim$1 kpc and mass resolution of $M_{\rm DM} \sim 10^7\,{\rm M}_\odot$). We evaluate the models both by looking at individual BH dynamics, growth and mergers, and by statistically comparing the behaviour of different models in terms of the mass growth and merger statistics. In particular, we focus on how various models affect the BH merger rate in the cosmological simulations, which is essential for making merger rate predictions for the *LISA* mission.

This paper proceeds as follows. In Section 2, we describe the numerical code and the Gaussian-constrained technique we use to study large SMBHs within a small volume. In Section 3, we talk about the different dynamical models for BH mergers that we study and test in this work. Section 4 is dedicated to investigating the effect of the different models on the evolution of individual BHs, while Section 5 studies the differences statistically. Finally, in Section 6, we show merger rate predictions with a model chosen based on the results of the previous sections, and compare with previous simulations at similar resolution.

2 THE SIMULATIONS

2.1 The numerical code

We use the massively parallel cosmological smoothed particle hydrodynamics (SPH) simulation software, MP-GADGET(Feng et al. 2016), to run all the simulations in this paper. The hydrodynamics solver of MP-GADGET adopts the new pressure–entropy formulation of SPH (Hopkins 2013). We apply a variety of subgrid models to model the galaxy and BH formation and associated feedback processes already validated against a number of observables (e.g. Feng et al. 2016; Waters et al. 2016; Di Matteo et al. 2017; Wilkins et al. 2017; Bhowmick et al. 2018; Huang et al. 2018; Ni et al. 2018; Tenneti et al. 2018; Marshall et al. 2020, 2021; Ni, Di Matteo & Feng 2021). Here we review briefly the main aspects of these. In the simulations, gas is allowed to cool through radiative processes (Katz, Hernquist & Weinberg 1999), including metal cooling. For metal cooling, we follow the method in Vogelsberger et al. (2014), and scale a solar metallicity template according to the metallicity of gas particles. Our star formation (SF) is based on a multiphase SF model (Springel & Hernquist 2003) with modifications following Vogelsberger et al. (2013). We model the formation of molecular hydrogen and its effects on SF at low metallicity according to the prescription of Krumholz & Gnedin (2011). We self-consistently estimate the fraction of

Table 1. Constrained simulations.

Name	$L_{\rm box}$ (h^{-1} Mpc)	$N_{\rm part}$	$M_{\rm DM}$ ($h^{-1}\,{\rm M}_\odot$)	$M_{\rm dyn,seed}$ ($M_{\rm DM}$)	$\epsilon_{\rm g}$ (h^{-1} kpc)	BH dynamics	Merging criterion
NoDF_4DM	10	176^3	1.2×10^7	4	1.5	Gravity	Distance
NoDF_4DM_G	10	176^3	1.2×10^7	4	1.5	Gravity	Distance & grav.bound
DF_4DM	10	176^3	1.2×10^7	4	1.5	Gravity+DF	Distance
Drag_4DM_G	10	176^3	1.2×10^7	4	1.5	Gravity+drag	Distance & grav.bound
DF+drag_4DM_G	10	176^3	1.2×10^7	4	1.5	Gravity+DF+drag	Distance & grav.bound
DF_4DM_G	10	176^3	1.2×10^7	4	1.5	Gravity+DF	Distance & grav.bound
DF_2DM_G	10	176^3	1.2×10^7	2	1.5	Gravity+DF	Distance & grav.bound
DF_1DM_G	10	176^3	1.2×10^7	1	1.5	Gravity+DF	Distance & grav.bound
DF(T15)_4DM_G	10	176^3	1.2×10^7	4	1.5	Gravity+DF(T15)	Distance & grav.bound
DF_HR_4DM_G	10	256^3	4×10^6	4	1.0	Gravity+DF	Distance & grav.bound
DF_HR_12DM_G	10	256^3	4×10^6	12	1.0	Gravity+DF	Distance & grav.bound

molecular hydrogen gas from the baryon column density, which in turn couples the density gradient to the SF rate. We include Type II supernova wind feedback (the model used in BlueTides; Okamoto et al. 2010; Feng et al. 2016) in our simulations, assuming that the wind speed is proportional to the local one-dimensional dark matter velocity dispersion.

BHs are seeded with an initial seed mass of $M_{\rm seed}=5\times10^5\,{\rm M}_\odot\,h^{-1}$ in haloes with mass more than $10^{10}\,{\rm M}_\odot\,h^{-1}$ if the halo does not already contain a BH. We model BH growth and active galactic nuclei (AGN) feedback in the same way as in the MassiveBlack I and II simulations, using the BH subgrid model developed in Springel, Di Matteo & Hernquist (2005) and Di Matteo, Springel & Hernquist (2005) with modifications consistent with BlueTides. The gas accretion rate on to the BHs is given by Bondi accretion rate,

$$\dot{M}_{\rm B}=\alpha\frac{4\pi G^2M_{\rm BH}^2\rho}{\left(c_{\rm s}^2+v_{\rm rel}^2\right)^{3/2}}, \tag{1}$$

where $c_{\rm s}$ and ρ are the local sound speed and density of the cold gas, $v_{\rm rel}$ is the relative velocity of the BH to the nearby gas, and $\alpha=100$ is a numerical correction factor introduced by Springel et al. (2005). This can also be eliminated (without affecting the values of the accretion rate significantly) in favour of a more detailed modelling of the contributions in the cold and hot phase accretion (Pelupessy, Papadopoulos & van der Werf 2006).

We allow for super-Eddington accretion in the simulation (e.g. Volonteri & Rees 2005; Volonteri, Silk & Dubus 2015), but limit the accretion rate to two times the Eddington accretion rate:

$$\dot{M}_{\rm Edd}=\frac{4\pi GM_{\rm BH}m_{\rm p}}{\eta\sigma_{\rm T}c}, \tag{2}$$

where $m_{\rm p}$ is the proton mass, $\sigma_{\rm T}$ the Thompson cross-section, c is the speed of light, and $\eta=0.1$ is the radiative efficiency of the accretion flow on to the BH. Therefore, the BH accretion rate is determined by

$$\dot{M}_{\rm BH}={\rm Min}(\dot{M}_{\rm B},2\dot{M}_{\rm Edd}). \tag{3}$$

The SMBH is assumed to radiate with a bolometric luminosity $L_{\rm Bol}$ proportional to the accretion rate $\dot{M}_{\rm BH}$:

$$L_{\rm Bol}=\eta\dot{M}_{\rm BH}c^2, \tag{4}$$

with $\eta=0.1$ being the mass-to-light conversion efficiency in an accretion disc according to Shakura & Sunyaev (1973). 5 per cent of the radiated energy is thermally coupled to the surrounding gas that resides within twice the radius of the SPH smoothing kernel of the BH particle. This scale is typically about ∼1–3 per cent of the virial radius of the halo.

The cosmological parameters used are from the 9-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*; Hinshaw et al. 2013) ($\Omega_0=0.2814$, $\Omega_\Lambda=0.7186$, $\Omega_{\rm b}=0.0464$, $\sigma_8=0.82$, $h=0.697$, $n_{\rm s}=0.971$). For our fiducial resolution simulations, the mass resolution is $M_{\rm DM}=1.2\times10^7\,{\rm M}_\odot\,h^{-1}$ and $M_{\rm gas}=2.4\times10^6\,{\rm M}_\odot\,h^{-1}$ in the initial conditions. The mass of a star particle is $M_*=1/4M_{\rm gas}=6\times10^5\,{\rm M}_\odot\,h^{-1}$. The gravitational softening length is $\epsilon_{\rm g}=1.5$ ckpc h^{-1} in the fiducial resolution for both DM and gas particles. The detailed simulation and model parameters are listed in Tables 1 and 2.

2.2 Gaussian constrained realization

MBHs at high redshift typically reside in rare density peaks, which are absent in the small uniform box (∼10 Mpc h^{-1}) simulations. In order to test the dynamics for more massive BHs (with $M_{\rm BH}>10^8\,{\rm M}_\odot$) in our small volume simulation, we apply the constrained realization (CR) technique[1] to impose a relatively high-density peak in the initial condition (IC), with peak height $\nu=4\sigma_0$ on scale of $R_{\rm G}=1$ Mpc h^{-1}.

The prescription for the CR technique was first introduced by Hoffman & Ribak (1991) as an optimal way to construct samples of constrained Gaussian random fields. This formalism was further elaborated and extended by van de Weygaert & Bertschinger (1996) as a more general type of convolution format constraints. The CR technique imposes constraints on different characteristics of the linear density field. It can specify density peaks in the Gaussian random field with any desired height and shape, providing an efficient way to study rare massive objects with a relatively small box and thus lower computational costs (e.g. Ni et al. 2021). In this study, we specify a $4\sigma_0$ density peak in the IC of our 10 Mpc h^{-1} box, boosting the early formation of haloes and BHs to study the dynamics of massive BHs. Before applying the peak height constraint, the highest density peak has $\nu=2.4\sigma_0$ and the largest BH has mass $<6\times10^7\,{\rm M}_\odot$ at $z=3$ in our fiducial model (`DF_4DM_G` in Table 1). After applying the $4\sigma_0$ constraint, the largest BH has mass $3\times10^8\,{\rm M}_\odot$ at $z=3$ in the same box.

3 BH DYNAMICS

3.1 BH dynamical mass

In our simulations, the seed mass of the BHs is $5\times10^5\,{\rm M}_\odot\,h^{-1}$, which is 20 times smaller than the fiducial dark matter particle mass at $1.2\times10^7\,{\rm M}_\odot\,h^{-1}$. Such a small mass of the BH relative to the dark

[1] https://github.com/yueyingn/gaussianCR

Table 2. Unconstrained simulations.

Name	$L_{\rm box}$ (h^{-1} Mpc)	$N_{\rm part}$	$M_{\rm DM}$ (h^{-1} M$_\odot$)	$M_{\rm dyn,seed}$ ($M_{\rm DM}$)	$\epsilon_{\rm g}$ (h^{-1} kpc)	BH dynamics	Merging criterion
L15_repos_4DM	15	256^3	1.2×10^7	4	1.5	Reposition	Distance
L15_NoDF_4DM	15	256^3	1.2×10^7	4	1.5	Gravity	Distance
L15_NoDF_4DM_G	15	256^3	1.2×10^7	4	1.5	Gravity	Distance & grav.bound
L15_DF_4DM	15	256^3	1.2×10^7	4	1.5	Gravity+DF	Distance
L15_DF_4DM_G	15	256^3	1.2×10^7	4	1.5	Gravity+DF	Distance & grav.bound
L15_DF(T15)_4DM_G	15	256^3	1.2×10^7	4	1.5	Gravity+DF(T15)	Distance & grav.bound
L15_DF+drag_4DM_G	15	256^3	1.2×10^7	4	1.5	Gravity+DF+drag	Distance & grav.bound
L35_NoDF_4DM_G	35	600^3	1.2×10^7	4	1.5	Gravity+DF	Distance & grav.bound
L35_DF+drag_4DM_G	35	600^3	1.2×10^7	4	1.5	Gravity+DF+drag	Distance & grav.bound

matter particles will result in very noisy gravitational acceleration on the BHs, and causes instability in the BH's motion and drift from the halo centre. Moreover, as shown in previous works (e.g. Tremmel et al. 2015; Pfister et al. 2019), under the low $M_{\rm BH}/M_{\rm DM}$ regime, it is challenging to effectively model DF in a subgrid fashion.

To alleviate dynamical heating by the noisy potential due to the low $M_{\rm BH}/M_{\rm DM}$ ratio, we introduce a second mass tracer, the dynamical mass $M_{\rm dyn}$, which is set to be comparable to $M_{\rm DM}$ when the BH is seeded. This mass is used in force calculation for the BHs, including the gravitational force and DF, while the intrinsic BH mass $M_{\rm BH}$ is used in the accretion and feedback process. $M_{\rm dyn}$ is kept at its seeding value $M_{\rm dyn,seed}$ until $M_{\rm BH} > M_{\rm dyn,seed}$. After that $M_{\rm dyn}$ grows following the BH's mass accretion. With the boost in the seed dynamical mass, the sinking time-scale will be shortened by a factor of $\sim M_{\rm BH}/M_{\rm dyn}$ compared to the no-boost case. Because of the boost in mass, we note that while our model can be safely applied to the more massive BHs, there is limitation in the seed BH dynamics. On the other hand, we also note that the bare BH sinking time-scale estimated in the no-boost case could overestimate the true sinking time, as the high-density stellar bulges sinking together with the BH are not fully resolved (e.g. Antonini & Merritt 2012; Biernacki, Teyssier & Bleuler 2017; Dosopoulou & Antonini 2017).

The boost we need to prevent dynamical heating depends on the dark matter particle mass $M_{\rm DM}$ (if we have high enough resolution the boost is no longer necessary), so we parametrize the dynamical mass in terms of the dark matter particle mass, $M_{\rm dyn,seed} = k_{\rm dyn} M_{\rm DM}$, instead of setting an absolute seeding dynamical mass for all simulations. We expect that as we go to higher resolutions where $M_{\rm DM}$ is comparable to $M_{\rm BH,seed}$, the dynamical seed mass should converge to the BH seed mass, if we keep $k_{\rm dyn}$ constant. We study the effect of setting different $k_{\rm dyn}$ by running three simulations with the same resolution and DF models, but various $k_{\rm dyn}$ ratios. They are listed in Table 1 as `DF_4DM_G`, `DF_2DM_G`, and `DF_1DM_G`, with $k_{\rm dyn} = 4$, 2, and 1, respectively.

To explore the effects of the BH seed dynamical mass on the motion and mergers of the BH, we test a variety of $M_{\rm dyn,seed}$ values in our simulations. The comparison between different $M_{\rm dyn,seed}$ can be found in Appendix A.

3.2 Modelling of black hole dynamics

3.2.1 Reposition of the black hole

Before introducing our DF implementations, we first describe a baseline model utilized by many large-volume cosmological simulations: the reposition model. As the name suggests, the reposition model of BH dynamics places the BH at the location of a local gas particle with minimum gravitational potential at each time-step, in order to avoid the unrealistic motion of the BHs due to limited mass and force resolution. This is particularly preferred for large-volume, low-resolution cosmological simulations (e.g. Springel et al. 2005; Sijacki et al. 2007; Booth & Schaye 2009; Schaye et al. 2015; Pillepich et al. 2018), where the BH mass is smaller than a star or gas particle mass and the BH can be inappropriately scattered around by two-body forces and the noisy local potential.

This simple fix of repositioning, however, comes with many disadvantages. For example, it may lead to higher accretion and feedback of the BHs, as they sink to the high-density regions too quickly. As was shown in Wurster & Thacker (2013) and Tremmel et al. (2017), repositioning also leads to burstier feedback of the BHs, which is more likely to quench star formation in the host galaxies. Moreover, repositioning leads to ill-defined velocity and non-smooth trajectories of the BH particles. Because of the ill-defined velocity and extremely short orbital decay time, such methods cannot be reliably used for merger rate predictions without careful post-processing calculations to account for the orbital decays.

In our study, we use the reposition model as a reference for the BH statistics, as it is still widely adopted in many existing simulations. We want to compare the DF models with the reposition model and quantify the effect of repositioning on BH mass growth and merger rate compared with the DF models.

3.2.2 Dynamical friction from collisionless particles

When the BH travels through a continuous medium or a medium consisting of particles with smaller masses than the BH, it attracts the surrounding mass towards itself, leaving a tail of overdensity behind. Dynamical friction (DF) is the resulting gravitational force exerted on to the BH by this tail of overdensity (e.g. Chandrasekhar 1943; Binney & Tremaine 2008). DF causes the orbits of SMBHs to decay towards the centre of massive galaxies (e.g. Governato, Colpi & Maraschi 1994; Kazantzidis et al. 2005), and enables the BHs to stay at the high-density regions where they could go through efficient accretion and mergers.

We follow equation (8.3) in Binney & Tremaine (2008) for the acceleration of the BH due to DF:

$$\boldsymbol{F}_{\rm DF} = -16\pi^2 G^2 M_{\rm BH}^2 m_{\rm a} \log(\Lambda) \frac{\boldsymbol{v}_{\rm BH}}{v_{\rm BH}^3} \int_0^{v_{\rm BH}} {\rm d}v_{\rm a} v_{\rm a}^2 f(v_{\rm a}), \tag{5}$$

where $M_{\rm BH}$ is the black hole mass, $\boldsymbol{v}_{\rm BH}$ is the velocity of the black hole relative to its surrounding medium, $m_{\rm a}$ and $v_{\rm a}$ are the masses and velocities of the particles surrounding the BH, and $\log(\Lambda) = \log(b_{\rm max}/b_{\rm min})$ is the Coulomb logarithm that accounts for the effective range of the friction between $b_{\rm min}$ and $b_{\rm max}$ (we will specify how we set these parameters later). $f(v_{\rm a})$ is the velocity distribution of the surrounding particles (unless we explicitly state otherwise,

all variables involving the BH's surrounding particles are calculated using stars and dark matter particles). Here we have assumed an isotropic velocity distribution of the particles surrounding the BH, so that we are left with a 1D integration.

We test two different numerical implementations of the DF in our simulations: one with a more aggressive approach that likely overestimates the effective range of DF, but could be more suitable for large-volume simulations (we refer to it as DF(fid) in places where we carry out explicit comparisons between the two DF models, and drop the 'fid' in all other places); the other with a more conservative method that aims to only account for the DF below the gravitational softening length, and is well tested for smaller volume, high-resolution simulations (Tremmel et al. 2015) (we refer to it as DF(T15)).

We begin by introducing the DF(fid) model. In this model, we further follow the derivation in Binney & Tremaine (2008), and approximate $f(v_a)$ by the Maxwellian distribution, so that equation (5) reduces to

$$\boldsymbol{F}_{\rm DF,fid} = -4\pi\rho_{\rm sph}\left(\frac{GM_{\rm dyn}}{v_{\rm BH}}\right)^2 \log(\Lambda_{\rm fid})\mathcal{F}\left(\frac{v_{\rm BH}}{\sigma_v}\right)\frac{\boldsymbol{v}_{\rm BH}}{v_{\rm BH}}. \tag{6}$$

Here $\rho_{\rm sph}$ is the density of dark matter and star particles within the SPH kernel (we will sometimes refer to these particles as 'surrounding particles') of the BH. All other definitions follow those of equation (5), except that we have substituted $M_{\rm BH}$ with $M_{\rm dyn}$ following the discussion in Section 3.1. We caution the reader that throughout our paper, this boost in mass could lead to an overestimation of the DF, and therefore an underestimation of the DF decay time-scale, especially for seed BHs. The function $\mathcal{F}$ defined as

$$\mathcal{F}(x) = {\rm erf}(x) - \frac{2x}{\sqrt{\pi}}{\rm e}^{-x^2},\ x = \frac{v_{\rm BH}}{\sigma_v} \tag{7}$$

is the result of analytically integrating the Maxwellian distribution, where σ_v is the velocity dispersion of the surrounding particles.

The subscript 'fid' in $\log(\Lambda)$ means that this definition of Λ is specific to the DF(fid) model, with

$$\Lambda_{\rm fid} = \frac{b_{\rm max,fid}}{(GM_{\rm dyn})/v_{\rm BH}^2},\ b_{\rm max,fid} = 10\,{\rm ckpc}\,h^{-1}. \tag{8}$$

Note that here we have defined $b_{\rm max}$ as a constant roughly equal to six times the gravitational softening. As there is no general agreement on the distance above which DF is fully resolved, we tested several values ranging from $\epsilon_{\rm g}$ to $20\epsilon_{\rm g}$. We found that values above $2\epsilon_{\rm g}$ are effective in sinking the BH, although a smaller $b_{\rm max}$ tends to result in more drifting BHs at higher redshift. By using this definition, we are likely overestimating the effective range of DF. However, we find this overestimation necessary in the early stage of BH growth to stabilize the BH motion.

We also implement a more localized version of DF following Tremmel et al. (2015) that we call DF(T15). Under the DF(T15) model, the DF is expressed as

$$\boldsymbol{F}_{\rm DF,T15} = -4\pi\rho(v < v_{\rm BH})\left(\frac{GM_{\rm dyn}}{v_{\rm BH}}\right)^2 \log(\Lambda_{\rm T15})\frac{\boldsymbol{v}_{\rm BH}}{v_{\rm BH}}, \tag{9}$$

where $\boldsymbol{v}_{\rm BH}$ is the velocity of the black hole relative to its surrounding medium. Note that different from equation (5), the surrounding density here only accounts for the particles moving slower than the BH with respect to the environment. More formally,

$$\rho(v < v_{\rm BH}) = \frac{M(< v_{\rm BH})}{M_{\rm total}}\rho_{\rm T15}, \tag{10}$$

where $M_{\rm total}$ is the total mass of the nearest 100 DM and stars, $M(<v_{\rm BH})$ is the fractional mass counting only DM and star particles with velocities smaller than the BH, and $\rho_{\rm T15}$ is the density calculated from the nearest 100 DM/star particles (note that in comparison, the SPH kernel contains 113 gas particles but far more collisionless particles; see Fig. 5). By using $\rho(v < v_{\rm BH})$ in place of $\rho_{\rm sph}\mathcal{F}$, we are approximating the velocity distribution of surrounding particles by the distribution of the nearest 100 collisionless particles. Another major difference from the DFsph model is the Coulomb logarithm, where in this model we define

$$\Lambda_{\rm T15} = \frac{b_{\rm max,T15}}{(GM_{\rm dyn})/v_{\rm BH}^2},\ b_{\rm max,T15} = \epsilon_{\rm g}. \tag{11}$$

The choice of a lower $b_{\rm max}$ is consistent with the localized density and velocity calculations, and by doing so we have assumed that DF is fully resolved above the gravitational softening.

3.2.3 *Gas drag*

In addition to the DF from dark matter and stars, the BH can also lose its orbital energy due to the DF from gas (to distinguish from DF from dark matter and stars, we will refer to the gas DF as 'gas drag' hereafter). Ostriker (1999) first came up with the analytical expression for the gas drag term from linear perturbation theory, and showed that in the transonic regime the gas drag can be more effective than the DF from collisionless particles. Although later studies show that Ostriker (1999) likely overestimates the gas drag for gas with Mach numbers slightly above unity (e.g. Escala et al. 2004; Chapon et al. 2013), simulations with gas drag implemented still demonstrate that this is an effective channel for BH energy loss during orbital decays (e.g. Chapon et al. 2013; Dubois et al. 2013; Pfister et al. 2019).

In order to investigate the relative effectiveness of DF and gas drag, we also include gas drag on to BHs in our simulations following the analytical approximation from Ostriker (1999):

$$\boldsymbol{F}_{\rm drag} = -4\pi\rho\left(\frac{GM_{\rm dyn}}{c_{\rm s}^2}\right)^2 \times \mathcal{I}(\mathcal{M})\frac{\boldsymbol{v}_{\rm BH}}{v_{\rm BH}}, \tag{12}$$

where $c_{\rm s}$ is the sound speed, $\mathcal{M} = \frac{|\boldsymbol{v}_{\rm BH}-\boldsymbol{v}_{\rm gas}|}{c_{\rm s}}$ is the Mach number, and $\mathcal{I}(\mathcal{M})$ is given by

$$\mathcal{I}_{\rm subsonic} = \mathcal{M}^{-2}\left[\frac{1}{2}\log\left(\frac{1+\mathcal{M}}{1-\mathcal{M}}\right) - \mathcal{M}\right], \tag{13}$$

$$\mathcal{I}_{\rm supersonic} = \mathcal{M}^{-2}\left[\frac{1}{2}\log\left(\frac{\mathcal{M}+1}{\mathcal{M}-1}\right) - \log\Lambda_{\rm fid}\right], \tag{14}$$

where $\log\Lambda_{\rm fid}$ is the Coulomb logarithm defined similarly to the collisionless DF.

3.3 Merging criterion

In all of our simulations, we set the merging distance to be $2\epsilon_{\rm g}$, because the BH dynamics below this distance is not well resolved due to our limited spatial resolution. We conserve the total momentum of the binary during the merger.

Under the baseline repositioning treatment of the BH dynamics, the velocity of the BH is not a well-defined quantity. Therefore, in cosmological simulations with repositioning, the distance between the two BHs is often the only criterion imposed during the time of mergers (e.g. BlueTides – Feng et al. 2016; Illustris – Vogelsberger et al. 2013; and IllustrisTNG – Pillepich et al. 2018). One problem with using only the distance as a merging criterion is that it can

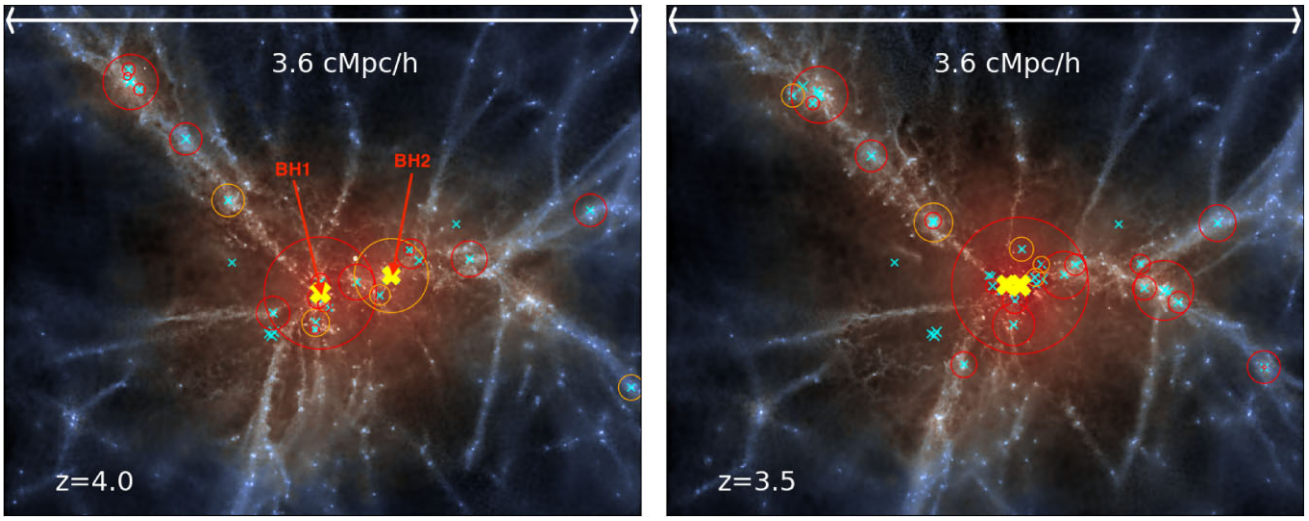

Figure 1. Visualization of $4\sigma_0$ density peak of the DF_4_DM_G simulation at $z = 4.0$ and $z = 3.5$. The brightness corresponds to the gas density, and the warmness of the tone indicates the mass-weighted temperature of the gas. We plot the BHs (cross) with mass $>10^6\,M_\odot$, as well as the haloes (subhaloes) hosting them (red circles correspond to central haloes and orange circles correspond to subhaloes). The circle radius shows the virial radius of the halo; haloes are identified by AMIGA Halo Finder (AHF). This density peak hosts the two largest BHs in our simulations (yellow cross), and they are going through a merger along with the merger of their host haloes between $z = 4$ and $z = 3$. For the BH and merger case studies, we will use examples from the circled haloes/BHs shown in this figure.

spuriously merge two passing by BHs with high velocities, when in reality they are not gravitationally bound and should not merge just yet (or may never merge). Although some similar resolution simulations such as EAGLE (Crain et al. 2015; Schaye et al. 2015) also check whether two BH particles are gravitationally bound, the BHs still do not have a well-defined orbit and sinking time due to the discrete positioning.

When we turn-off the repositioning of the BHs to the nearby minimum potential, the BHs will have well-defined velocities at each time-step (this is true whether or not we add the DF). This allows us to apply further merging criteria based on the velocities and accelerations of the BH pair, and thus avoid earlier mergers of the gravitationally unbound pairs. Also, as the BH pairs now have well-defined orbits all the way down to the numerical merger time, we will be able to directly measure binary separation and eccentricity from the numerical merger, and use the measurements as the initial condition for post-processing methods without having to assume a constant initial value (e.g. Kelley et al. 2017).

We follow Bellovary et al. (2011) and Tremmel et al. (2017), and use the criterion

$$\frac{1}{2}|\boldsymbol{\Delta v}|^2 < \boldsymbol{\Delta a}\boldsymbol{\Delta r} \tag{15}$$

to check whether two BHs are gravitationally bound. Here $\boldsymbol{\Delta a}$, $\boldsymbol{\Delta v}$, and $\boldsymbol{\Delta r}$ denote the relative acceleration, velocity, and position of the BH pair, respectively. Note that this expression is not strictly the total energy of the BH pair, but an approximation of the kinetic energy and the work needed to get the BHs to merge. Because in the simulations the BH is constantly interacting with surrounding particles, on the right-hand side we use the overall gravitational acceleration instead of the acceleration purely from the two-body interaction.

4 CASE STUDIES OF BH MODELS

Given the variety of models we have described so far, we first study the effect of different BH dynamics models by looking at the individual BH evolution and BH pairs using the constrained simulations. The details of these simulations and specific dynamical models are shown in Table 1. For all the constrained simulations, we use the same initial conditions, which enable us to do a case-by-case comparison between different BH dynamical models.

For the case studies, we choose to study the growth and merger histories of the two largest BHs and a few surrounding BHs within the density peak of our simulations. The haloes and BHs at the $4\sigma_0$ density peak in DF_4DM_G are shown in Fig. 1. The haloes and subhaloes shown in circles are identified with AMIGA Halo Finder (AHF; Knollmann & Knebe 2009). The haloes are centred at the minimum-potential gas particle within the halo, and the sizes of the circles correspond to the virial radius of the halo. Throughout the paper, we will always define the halo centres by the position of the minimum-potential gas particle, and we note that the offset between the minimum-potential gas and the halo centre given by AHF (found via density peaks) is always less than 1.5 ckpc h^{-1}. The cyan crosses are BHs with mass larger than $10^6\,M_\odot\,h^{-1}$, and the yellow crosses are the two largest BHs in the simulation. From the plot, we can see that in the DF_4DM_G simulation, most of the BHs already reside in the centre of their hosting haloes at $z = 4$, although we also see some cases of wandering BHs outside of the haloes.

4.1 Black hole dynamics modelling

To compare different dynamical models, we look at the distance between the BH and the halo centre $\Delta r_{\rm BH}$ (we will sometimes refer to this distance as 'drift' hereafter), the BH mass, and the velocity along the x direction through the entire history of BH2 from Fig. 1.

We evaluate the BH drift with two approaches: at each time-step, we find the minimum potential gas particle within 10 ckpc h^{-1} of the BH and calculate the distance between this gas particle and the BH. This is a quick evaluation of the drift that allows us to trace the BH motion at each time-step, but it fails to account for orbits larger than 10 ckpc h^{-1}, and the minimum-potential gas particle may not reside in the same halo as the BH. Therefore, for each snapshot

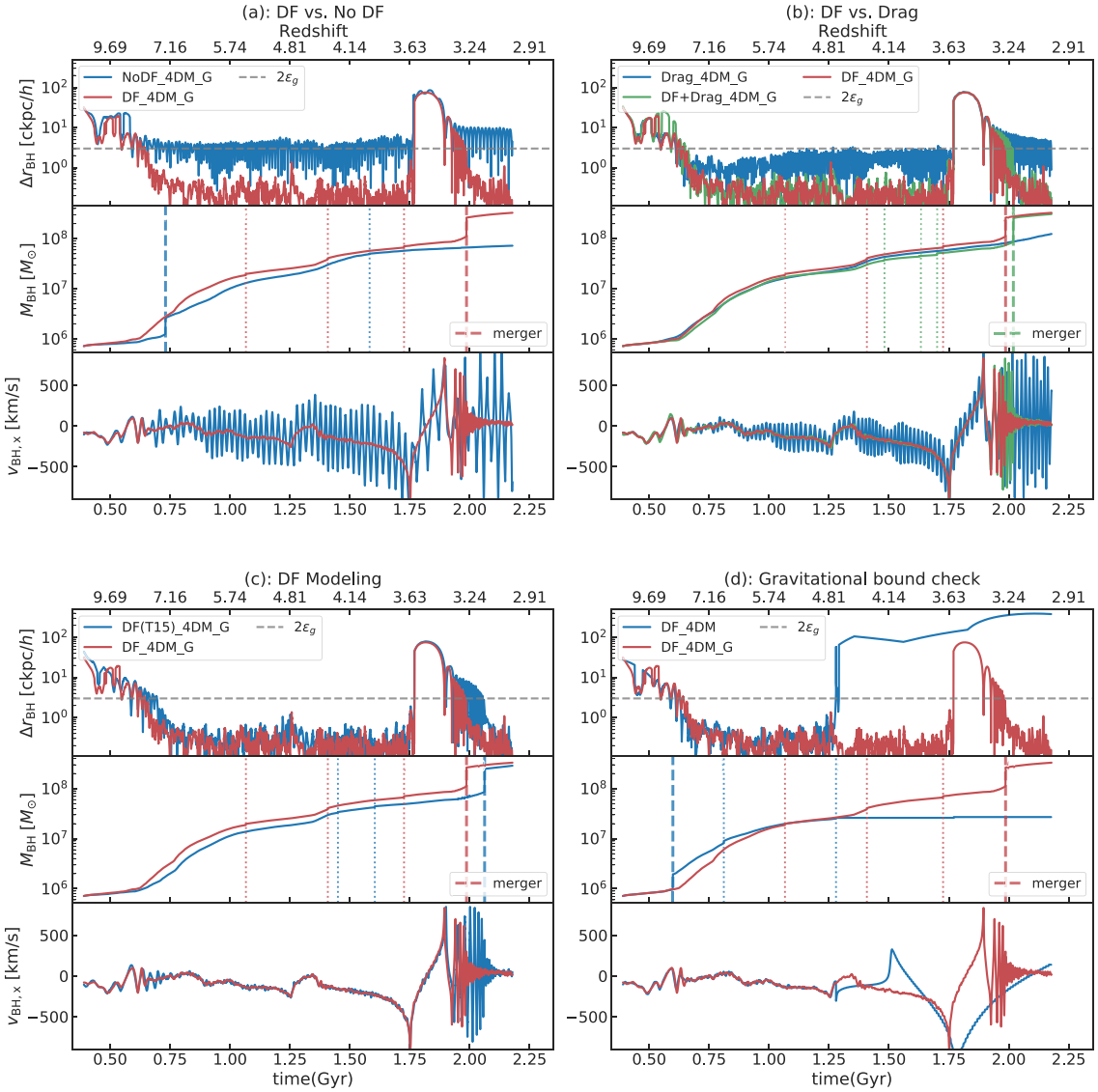

Figure 2. The evolution of BH2 in Fig. 1 under different BH dynamics prescriptions. We show the distance to halo centre (top), BH mass (middle), and the x-component of the BH velocity (bottom). Mergers are shown in vertical lines [thick dashed lines are major mergers ($q > 0.3$), and thin dotted lines are minor mergers]. (a) Comparison between no-DF and DF models. DF clearly helps the BH sink to the halo centre and stay there. (b) Effects of DF from stars and dark matter compared with gas drag. DF has a stronger effect throughout, except that in the very early stage the drag-only model is comparable to the DF-only model. (c) Comparison between the DF(fid) and DF(T15) model. In general, the DF(fid) model results in a more stable BH motion and faster sinking, but the difference is small. (d) BH dynamics with and without the gravitational bound check during mergers. Without the gravitational bound check, the BHs can merge while still moving with large momenta, and thereby get kicked out of the halo by the injected momentum.

we saved, we define the drift more carefully by running the halo finder and calculate the distance between the BH and the centre of its host halo. Whenever the BH is further than 9 ckpc h^{-1} from the minimum-potential gas particle, we take the distance from the two nearest snapshots and linearly interpolate in time between them. Otherwise we use the distance to the local minimum-potential gas particle calculated at each time-step.

4.1.1 DF and no correction

Before calibrating our DF modelling, we first demonstrate the effectiveness of our fiducial DF model, `DF_4DM_G`, by comparing it with the no-DF run `NoDF_4DM_G` (note that throughout the paper, no-DF means no correction to the BH dynamics of any form besides the resolved gravity). We keep all parameters fixed except for the BH dynamics modelling. The details of these simulations can be found in Table 1.

In Fig. 2(a), we show the evolution of BH2 in Fig. 1 under the no-DF and the fiducial DF models. Without any correction to the BH dynamics, even the largest BH in the simulation does not exhibit efficient orbital decay throughout its evolution: the distance from the halo centre is always fluctuating above $2\epsilon_g$. This is because the BH does not experience enough gravity on scales below the softening length, and cannot lose its angular momentum efficiently. Now when we add the additional DF to compensate for the missing small-scale gravity, the BH is able to sink to within 1 ckpc h^{-1} of the halo centres in <200 Myr and remain there.

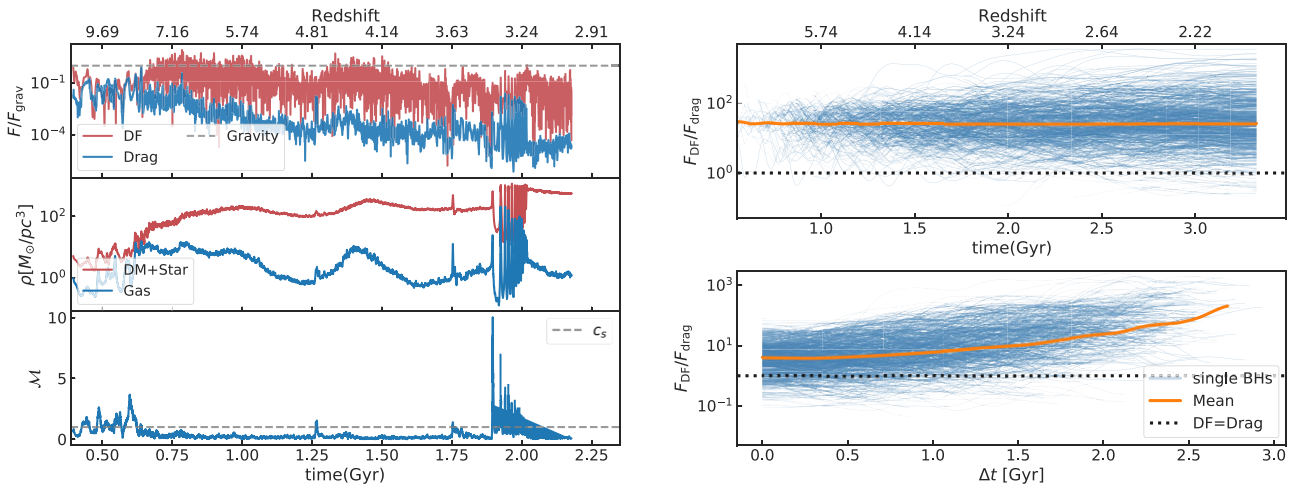

Figure 3. Comparisons between DF and hydro drag. Left: comparison for a single BH. In the top panel, we show the magnitude of the DF (red) and gas drag (blue) relative to gravity for the same BH, in the `DF+Drag_4DM_G` run. During the early stage of the BH evolution, DF and gas drag have comparable effect, while after $z = 7.5$ the gas drag becomes less and less important, as the gas density decreases relative to the stellar density (middle), and the BH velocity goes into the subsonic regime (lower). Right: ratio between DF and gas drag for all BHs. We plot the ratio both as a function of redshift (top) and as a function of time after a BH is seeded (bottom). The orange lines represent the logarithmic mean of the scatter. The F_{DF}/F_{drag} ratio depends strongly on the evolution time of the BH: the longer the BH evolves, the less important the drag force is. However, there is not a strong correlation between redshift and the F_{DF}/F_{drag} ratio.

The 90 ckpc h^{-1} peak in the drift of the BH marks the merger between BH1 and BH2 in Fig. 1, when the host halo of BH2 merges into the host of BH1, and the halo centre is redefined near the merger. After the halo merger, DF is able to sink the BH to the new halo centre and allows it to merge with the BH in the other halo, whereas in the no-DF case we do not see the clear orbital decay of the BHs after the merger of their host halo until the end of the simulation.

Besides the drift, we also show the x-component of the BH's velocity relative to its surrounding collisionless particles (lower panel). Here we show one component instead of the magnitude to better visualize the velocity oscillation. With DF turned on, the velocity of the BH is more stable, as the BH's orbit has already become small and is effectively moving together with the host halo. Without DF, the BH tends to oscillate with large velocities around the halo centre without losing its angular momentum.

The different dynamics of the BH can also affect accretion due to differences in density and velocities, so we also look at the BHs' mass growth in the two scenarios (middle panel). The mass growths of the two BHs are similar under the two models, although when subjected to DF, the BHs have more and earlier mergers. Even though the BH mass is less sensitive to the dynamics modelling, the merger rate predictions can be affected significantly as we will discuss later.

Note that for our no-DF model, we have also boosted the dynamical mass to $4 \times M_{DM}$ at the early stage to prevent scattering by the dark matter and star particles. However, even after the boost, the BHs cannot lose enough angular momentum to be able to stay at the halo centre. This means that even though dynamical heating is alleviated through the large dynamical mass, the subresolution gravity is still essential in sinking the BH to the host halo centre.

4.1.2 Dynamical friction and gas drag

In the previous subsection, we have only included collisionless particles (DM+star) when modelling the dynamical friction (DF), now we will look into the effects of DF of gas (gas drag) in comparison with the collisionless particles in the context of our simulations.

From equations (6) and (12), the relative magnitudes of DF and drag mainly depend on two components: the relative density of DM+stars versus gas, and the values of $\mathcal{F}(x)$ and $\mathcal{I}(\mathcal{M})$. Ostriker (1999) has shown that when a BH's velocity relative to the medium falls in the transonic regime (i.e. near the local sound speed), $\mathcal{I}$ is a few times higher than $\mathcal{F}$, while in the subsonic and highly supersonic regimes $\mathcal{I}$ is smaller or equal to $\mathcal{F}$. Therefore, we would expect the gas drag to be larger when the BH is in the early sinking stage with a relatively high velocity and a high gas fraction.

In Fig. 3, the left-hand panel shows the comparison between the magnitude of DF and gas drag through different stages of the BH evolution, as well as the factors that can alter the effectiveness of the gas drag. In the very early stages ($z > 7.5$) of BH evolution, DF and gas drag have comparable effects, while after $z = 7.5$ the gas drag becomes significantly less important and almost negligible compared with DF. The reason follows what we have discussed earlier: the gas density decreases relative to the stellar density (shown in the middle panel), and the BH's velocity relative to the surrounding medium goes into the subsonic regime as a result of the orbital decay (shown in the lower panel). Around $z = 3.5$, there is a boost in the BH's velocity due to disruption during a major merger with a larger galaxy and BH. The effect of gas is again raised for a short period of time (although still subdominant compared to the DF).

In Fig. 2(b), we plot the BH evolution for the DF-only (`DF_4DM_G`), drag-only (`Drag_4DM_G`), and DF+drag (`DF+Drag_4DM_G`) simulations. Both the drag-only and DF-only models are effective in sinking the BH at early times ($z > 7$). However, at lower redshifts, the gas drag is not able to sink the BH by itself, whereas DF is far more effective in stabilizing the BH at the halo centre. For this reason, in low-resolution cosmological simulations, DF from collisionless particles is necessary to prevent the drift of the BHs out of the halo centre.

To further illustrate the relative importance between DF and gas drag for the entire BH population, we examine the dependencies of

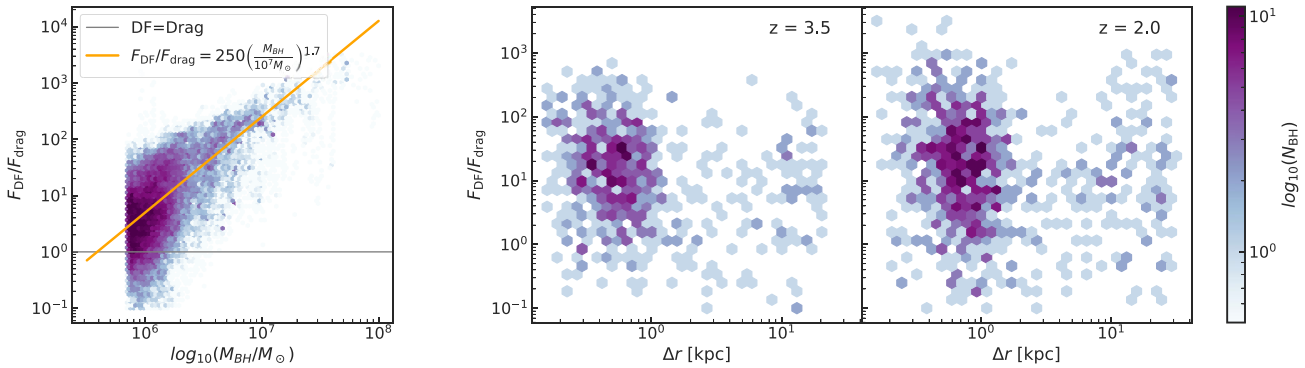

Figure 4. Left: scattering relation between the F_{DF}/F_{drag} ratio and the BH mass. For each BH, we sample its mass at uniformly distributed time bins throughout its evolution, and we show the scattered density of all samples. DF has significantly larger effects over gas drag on larger BHs. We fit the scatter to a power law shown in the orange line. Right: scattering relation between the F_{DF}/F_{drag} ratio and the BHs' distance to the halo centre. Comparing with the BH mass, we do not see a clear dependence of the F_{DF}/F_{drag} ratio on the distance to halo centre. For BHs at all locations within the halo, DF is in general larger than the gas drag.

the F_{DF}/F_{drag} on variables related to the BH evolution for all BHs in the `DF+Drag_4DM_G` simulation. First, in the right-hand panel of Fig. 3, we show the time evolution of F_{DF}/F_{drag}. The top panel shows the ratio as a function of cosmic time, while the bottom panel shows the ratio as a function of each BH's seeding time. The DF/Drag ratio has a wide range for different BHs, but overall DF is becoming larger relative to the gas drag as the BH evolves. From the mean value of the DF/drag ratio, we see that when the BHs are first seeded, DF is only a few times larger than the gas drag. After a few Gyr of evolution, DF becomes 2–3 orders of magnitude larger than the gas drag. However, there is not a strong correlation between redshift and the F_{DF}/F_{drag} ratio. Given the resolution of our simulations, we do not resolve some high-density gas clumps nor the stellar clusters in our simulations, and therefore the comparison between DF and drag is only limited to simulations of a comparable resolution. The result from very high-resolution simulations may be different from ours.

In the left-hand panel of Fig. 4, we show the scattering relation between the F_{DF}/F_{drag} ratio and the BH mass M_{BH}. We see a strong correlation between the F_{DF}/F_{drag} ratio and the BH mass: DF has significantly larger effects over gas drag on larger BHs, although the range of the ratio is large at the low-mass end. We fit a power law to the median of the scatter:

$$\frac{F_{DF}}{F_{drag}} = 250 \left(\frac{M_{BH}}{10^7\,M_\odot} \right)^{1.7}, \tag{16}$$

which roughly characterize the effect of the two forces on BHs of different masses. From this relation we see that for BHs with masses $>10^7\,M_\odot$, gas drag is in general less than 1 per cent of DF. Finally, the right-hand panels show the relation between the F_{DF}/F_{drag} ratio and the BH's distance to the halo centre: there is not a strong dependency on the BH's position within the halo.

4.1.3 Comparisons with the T15 model

For the collisionless particles, we test and study two different implementations for the DF: DF(fid) and DF(T15) (see Section 3 for detailed descriptions). In Section 3, we pointed out three main differences between them: different kernel sizes (SPH kernel versus nearest 100 DM+star), different definitions of b_{max} (10 ckpc versus 1.5 ckpc h^{-1}), and different approximation of the surrounding velocity distribution (Maxwellian versus nearest 100-sample distribution).

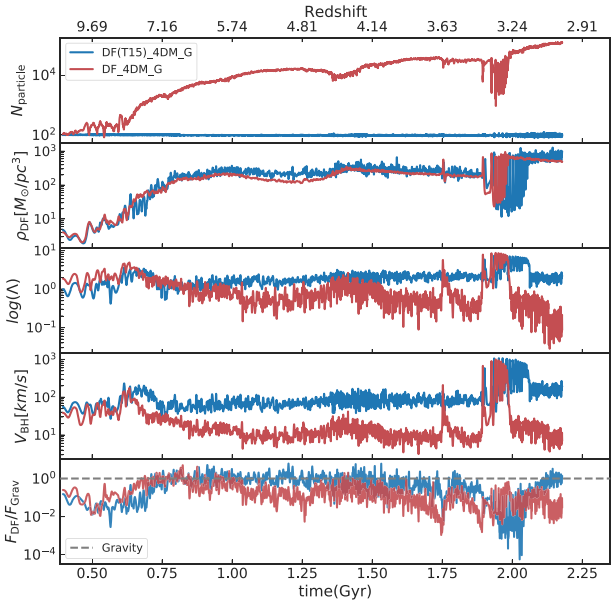

Figure 5. Comparison between different components in the two DF models, DF(fid) (red) and DF(T15) (blue) (see Section 3 for descriptions). We show the number of stars and dark matter particles included in the DF density and velocity calculation (top panel), the density used for DF calculation (second panel), the Coulomb logarithm used in the two methods (third panel), the velocity of the BH relative to the surrounding particles (fourth panel; note that the 'surrounding particles' are defined differently for the two models), and the magnitude of DF relative to gravity (bottom panel). The higher DF in the DF(fid) model at $z > 8$ is due to the larger Coulomb logarithm. After $z \sim 7$, the higher density of DF(T15) due to more localized density calculation counterbalances its lower $\log(\Lambda)$, resulting in similar DF between $z = 8$ and $z = 3.5$. During the halo merger at $z = 3.5$, the DF(fid) model included particles from the target halo into the density calculation, and therefore yields larger DF during the merger.

Essentially, these differences mean that DF(fid) is a less localized implementation than DF(T15). Now we would like to evaluate the effectiveness of these two implementations and show how different factors affect the final DF calculation.

Fig. 5 shows the relevant quantities in the DF computation for the two methods. The two kernels both contain ∼100 dark matter

and star particles at high redshift ($z > 8$), but after that the SPH kernel (defined to include the nearest 113 gas particles) begins to include more and more stars and dark matter. The mass fraction of stars in the SPH kernel dominate over that of dark matter by ∼10 times for a BH at the centre of the galaxy. The larger kernel of DF(fid) has two effects: first, the DF density will be smoother over time; second, during halo mergers, the DF(fid) kernel can 'see' the high-density region of the larger halo, which results in a higher DF near mergers compared to DF(T15). This is confirmed by the second panel, where we show the density for DF calculation from the two kernels. The densities calculated from the two kernels are similar in magnitude throughout the evolution, although the DF(T15) kernel yields slightly larger density due to its smaller size. Around the BH merger, the density in DF(fid) is larger due to its inclusion of the host halo's central region.

The third panel shows the Coulomb logarithm in the two models. Recall that $\Lambda = \frac{b_{\rm max}}{(GM_{\rm BH})/v_{\rm BH}^2}$, and so the Coulomb logarithm depends on the BH's mass, its velocity relative to the surrounding particles, and the value of $b_{\rm max}$. From Fig. 2(c), the mass of the DF(T15) BH is slightly smaller, but the mass difference is small compared with the six times difference in $b_{\rm max}$. Given $b_{\rm max} = 10$ ckpc h^{-1} in DF(fid) and $b_{\rm max} = 1.5$ ckpc h^{-1} in DF(T15), we would expect the Coulomb logarithm to be larger for the former. However, there is yet another tweak: the $v_{\rm BH}^2$ term turns out to be significantly larger in the DF(T15) model (fourth panel). Note that in the DF(T15) model $v_{\rm BH}^2$ is calculated using only 100 surrounding particles, and for the high-density region we are considering here, the velocity of the nearest 100 particles is very noisy in time. As we will show in Appendix B, for smaller BHs the difference in $v_{\rm BH}^2$ is not as large, and usually DF(fid) has a larger $\log\Lambda$ due to its larger $b_{\rm max}$.

In Fig. 2(c), we show the evolution of the BH under these two models. At high redshift ($z > 8$), due to the large $\log(\Lambda)$, the BH in the DF(fid) simulation sinks slightly faster to the halo centre. Between $z = 8$ and $z = 3.5$, both models have similar DF (as discussed in the previous paragraph) and the motion and mass accretion are also similar. Then at $z = 3.5$, within the host halo of the BH major merger, DF in DF(fid) is again larger because the density kernel includes more particles from the high-density region in the target halo, and this leads to an earlier merger time.

Overall, the performance of the two models is similar. However, as we have seen in the velocity calculation of the BHs relative to the surrounding particles, DF(T15) could be too localized for simulations of our resolution ($\epsilon_{\rm g} \sim 1\,{\rm kpc}\,h^{-1}$) and is sometimes subject to numerical noise. Therefore, in our subsequent statistical runs we pick DF(fid) as our fiducial model, and will drop the 'fid' in its name hereafter.

4.1.4 *Gravitationally bound merging criterion*

The merging criterion can affect not only the merging time, but also the dynamics and evolution of the BHs. Naively, we might expect the distance-only merging to produce more massive BHs, because BHs are merged more easily. However, in many cases this is not true, and we will illustrate here through one example.

Fig. 2(d) shows the evolution of the same BH with the same DF prescription, but different merging criteria. We note a drastic difference in the BH's trajectories: while the BH in the gravitationally bound merger case is staying at the centre of its host halo, the BH in the distance-only merger flies out of its host after a merger. This is because with the distance-only model, it is possible for one BH to have a very large velocity at the time of the merger, since we do not limit the BH's velocity. By momentum conservation, the BH with a larger velocity can transfer the momentum to the other BH (and the merger remnant) that might have already sunk to the halo centre. The sunk BH then drifts out of the halo centre after a merger due to the large momentum injection. This is especially common in simulations where the BH's dynamical mass is boosted, because the injected momentum is also boosted with mass and a smaller BH in a satellite galaxy can easily kick a larger BH out. If we add on the gravitational bound check, there will be more time for the BHs to lose their angular momentum, and so the injected momentum is far less, and in most cases does not kick each other out of the central region.

4.2 Black hole mergers

Having seen the effect of different dynamical models on the evolution of individual BHs, next we will discuss how the dynamics, together with different BH merging criteria, affect the evolution and mergers of the BHs. In particular, we want to study their merging time and trajectories before and after the mergers. Similar to the previous subsection, we will draw our examples from the two haloes shown in Fig. 1.

4.2.1 *Effect of dynamical friction modelling*

We first look at how different dynamical models affect the time-scale of BH orbital decay and mergers. We pick two cases of mergers: one is an early merger at $z > 5$ when the BHs have not outgrown their dynamical masses; the other is a later merger at $z \sim 3.3$ when both BHs are larger than their seed dynamical masses (the major merger between BH1 and BH2 in Fig. 1). Following Tremmel et al. (2015), we also compute the dynamical friction (DF) time for the two mergers using equations (12)–(15) from Taffoni et al. (2003):

$$t_{\rm DF} = 0.6 \times 1.67\,{\rm Gyr} \times \frac{r_{\rm c}^2 V_{\rm h}}{GM_{\rm s}} \log^{-1}\left(1 + \frac{M_{\rm vir}}{M_{\rm s}}\right)\left(\frac{J}{J_{\rm c}}\right)^{\alpha}, \quad (17)$$

where $M_{\rm s}$ is the mass of the smaller BH (which we treat as the satellite), $M_{\rm vir}$ is the virial mass of the host halo of the larger BH (found by AHF), $V_{\rm h}$ is the circular velocity at the virial radius of the host, and $r_{\rm c}$ is the radius of a circular orbit with the same energy as the satellite BH's initial orbit. The last term $\left(\frac{J}{J_{\rm c}}\right)^{\alpha}$ is the correction for orbital eccentricity, where J is the angular momentum of the satellite, $J_{\rm c}$ is the angular momentum of the circular orbit with the same energy as the satellite, and α is given by

$$\alpha\left(\frac{r_{\rm c}}{R_{\rm vir}}, \frac{M_{\rm s}}{M_{\rm vir}}\right) = 0.475\left[1 - \tanh\left(10.3\left(\frac{M_{\rm s}}{M_{\rm vir}}\right)^{0.33} - 7.5\left(\frac{r_{\rm c}}{R_{\rm vir}}\right)\right)\right]. \quad (18)$$

In our calculation the virial radius, velocity, and mass are obtained from the AHF outputs, and the circular radius, orbit energy, and angular momentum are calculated by fitting the halo density profile to the Navarro–Frenk–White (NFW) profile.

Fig. 6 shows distances between two merging BHs in the no-DF, DF(fid), DF(T15), and gas drag models in the early and later stages of their evolution. For the early merger, the effect of the frictional forces (DF and drag) is not very big but still noticeable. The DF and gas drag have similar effects on the orbital decay at higher redshifts, consistent with our discussion in Section 4.1.2. The DF(T15) model

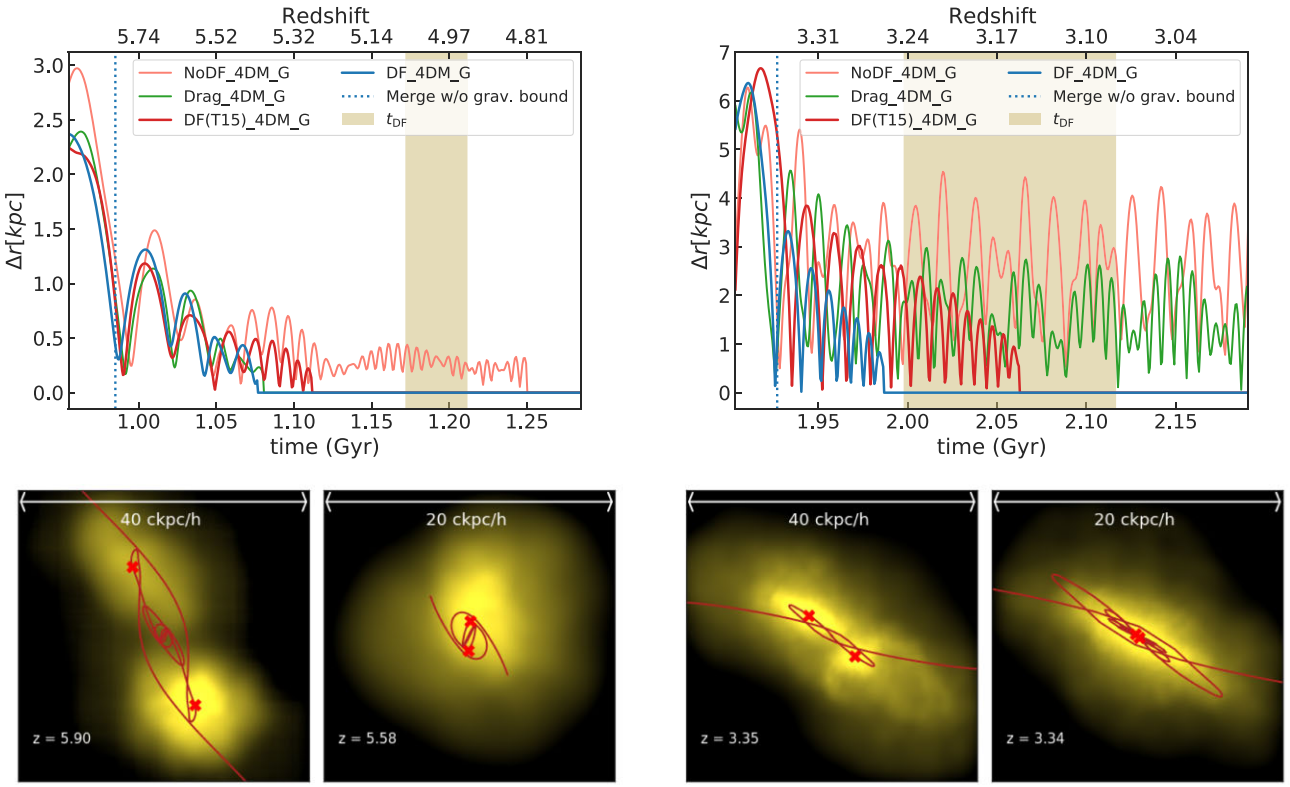

Figure 6. The comparison between the distance of two merging BHs in the no-correction, DF(fid), DF(T15), and gas drag models in the early stage (left) and later stage (right) of the BH evolution. For early mergers, the effect of the frictional forces (DF and drag) is not very prominent but still noticeable. The DF and gas drag both allow the BHs to merge faster compare to the no-DF case. For the later merger happening in a denser environment, the effect of DF is clear. However, the gas drag does not have a big effect on the BH at this late stage compared with the no-DF case. The lower panels show the merging BHs within their host galaxies and their trajectories towards the merger in the DF_4DM_G run. The left-hand images show the early phase of the orbital decay, and the right-hand images show the later phase when the orbits get smaller.

sinks the BH a little slower than the DF(fid) model, but the difference is within 50 Myr. All three friction models allow the BHs to merge faster compare to the no-DF case by ∼150 Myr.

For the later merger, which takes place in a denser environment, the effect of DF is clearer: the DF allows the BHs to sink within the gravitational softening of the particles in <200 Myr. Without DF the BH's orbit does not have a clear decay below 2 kpc and does not merge at the end of our simulation. Furthermore, the gas drag does not have a big effect on the BH at this late stage compared with the no-correction case. This follows from our discussion in Section 4.1.2 that gas drag is much less effective at lower redshift compared to DF.

In both plots, the yellow shaded region is the DF time from the analytical calculation in equation (17). Here we draw a band instead of a single line, because the BH's orbit is not a strict ellipse, and the BH is continuously losing energy. We calculate t_{DF} at multiple points between the first and second peak in the BH's orbit (e.g. between $z = 5.9$ and $z = 5.7$ in the earlier case), and plot the range of those t_{DF}. For both mergers, the analytical prediction is less than 150 Myr later than the merger of the (fid) model. We note that the Taffoni et al. (2003) analytical t_{DF} is a fit to the NFW profiles, and the previous numerical and analytical comparisons on the BH DF (e.g. Tremmel et al. 2015; Pfister et al. 2019) are performed in idealized NFW haloes with a fixed initial BH orbit. In our case, the halo profiles and BH orbits are not directly controlled, and therefore deviation from the analytical prediction is expected. We will study such deviations statistically later in Section 5.3.

4.2.2 *Effect of gravitational bound check*

In Section 3.3, we introduced two criteria that we use to perform BH mergers in our simulations: we can merge two BHs when they are close in distance, and we can also require that the two BHs are gravitationally bounded in addition to the distance check.

In Fig. 6, we show the difference in BHs' merging time with and without the gravitational bound criterion. The vertical dashed line marks the time that the two BHs in the DF_4DM_G simulation would merge if there was not the gravitational bound check. Without the gravitational bound check, the orbit of the BHs is still larger than 1 kpc when they merge, whereas with the gravitational bound check, the orbit size generally decays to less than 300 pc when the BHs merge. The merger without gravitational bound check generally makes the merger happen earlier by a few hundred Myr (we will study the orbital decay time statistically in the next section). Therefore, for more accurate merger rate predictions and the correct accretion and feedback, it is necessary to apply the gravitational bound check during BH mergers whenever the BH has a well-defined velocity.

5 BLACK HOLE STATISTICS

After looking at individual cases of BH evolution, we now turn to the whole SMBH population in the simulations with different modelling of BH dynamics. For statistics comparison, instead of using the $L_{box} = 10$ Mpc h^{-1} constrained realizations where we added 4σ density peaks to our small volumes (see Section 2.2), we now use

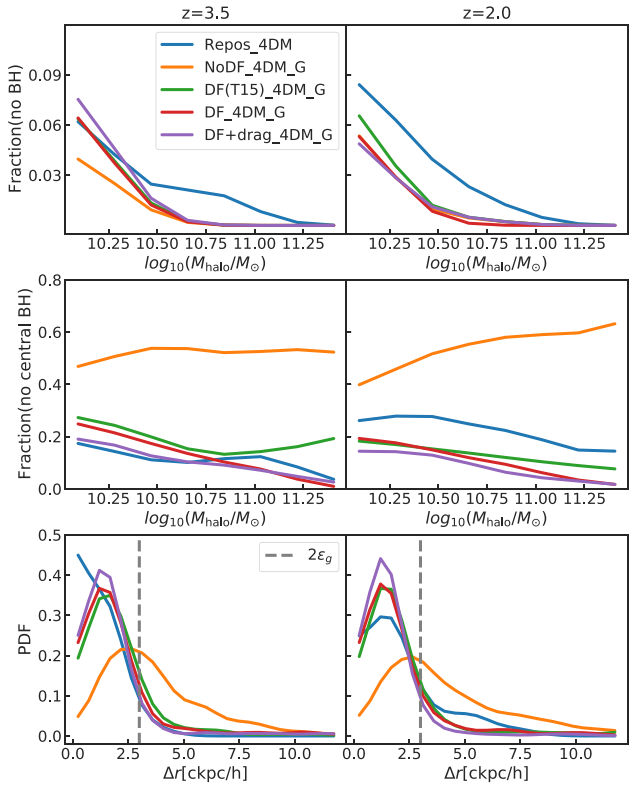

Figure 7. The effect of different BH dynamics modelling on BH position relative to its host. We include the reposition model (blue), no-DF model (orange), DF(T15) model (green), DF(fid) model (red), and the DF+drag model (purple). Top: the fraction of haloes(subhaloes) without a BH for haloes with masses above the BH seeding mass at $M_{\rm halo} = 10^{10}\,{\rm M}_\odot\,h^{-1}$. Middle: the fraction of haloes without a central BH ('central' means within $2\epsilon_{\rm g}$ from the halo centre identified by the halo finder), out of all haloes with BHs. Bottom: distribution of BHs' distance to its host halo centre.

$L_{\rm box} = 15$ Mpc h^{-1} unconstrained simulations. The details of our $L_{\rm box} = 15$ Mpc h^{-1} simulations are shown in Table 2.

5.1 Sinking of the black holes

With the added DF, we hope to assist the orbital decay of the BHs and prevent the stalling due to the smoothed gravity. Hence, we start by looking at the BHs' position relative to the host haloes. Because of the resolution limit of our simulations, we would not expect the BHs to be able to sink to the exact minimum potential. Instead we consider a $<2\epsilon_{\rm g} = 3$ ckpc h^{-1} distance to be 'good sinking'.

In Fig. 7, we show the statistics related to BHs' sinking status. We included the comparison between the reposition model (`L15_Repos_4DM`), the no-DF model (`L15_NoDF_4DM`), the two DF models (`L15_DF_4DM` and `L15_DF(T15)_4DM`), and the DF+drag model (`L15_DF+drag_4DM`). To start with, we simply count the fraction of haloes without a BH when its mass is already above the BH seeding criterion (i.e. $10^{10}\,{\rm M}_\odot\,h^{-1}$). The top panel shows the fraction of large haloes without a BH for different models at $z = 3.5$ and $z = 2$. Surprisingly, the no-DF model ends up with the least haloes without a BH. This is because even though the BHs without dynamical corrections cannot sink effectively, the high dynamical mass still prevents sudden momentum injections from surrounding particles, and therefore most BHs still stay within their host galaxies. The DF models perform equally well, with <10 per cent no-BH haloes at the low-mass end. The reposition model, however, ends up with the most no-BH haloes, even though repositioning is meant to pin the BHs to the halo centre. This happens because under the repositioning model, the central BHs tend to spuriously merge into a larger halo during fly-by encounters, leaving the smaller subhalo BH-less.

Next we look at where the BHs are located within their host galaxies. For all the haloes with at least one BH, we examine whether the BH is located at the centre (i.e. $<2\epsilon_{\rm g} = 3\,{\rm ckpc}\,h^{-1}$ from the halo centre). The middle panel of Fig. 7 shows the fraction of haloes without a central BH. The no-DF model has significantly more haloes without a central BH compared to the other models, with over half of the haloes hosting off-centre BHs. Among the three runs with DF, the DF(T15) and DF(fid) models have a similar fraction of haloes (∼20 per cent) without a central BH, and we can see this fraction dropping from $z = 3.5$ to $z = 2$, meaning that many BHs are still in the process of sinking towards the halo centre. When we further add the gas drag, 10 per cent more haloes host at least one central BH, and the difference between the drag and no-drag central BHs is more prominent at high redshifts.

Our definition of a 'central' BH depends on the resolution of our simulation, but we note that our conclusions above do not change if we shift this criterion by a factor of ∼2. Moreover, even though we consider the smaller fraction of off-centred BHs as an evidence that the DF is taking effect, we note that there are both theoretical and observational evidences that off-centre BHs do exit and merge with each other (e.g. Volonteri & Rees 2005; Bellovary et al. 2010; Kulier et al. 2015). In the case of dual/offset AGN observations, we should also expect to see BHs further than 5 kpc apart (e.g. Barrows, Comerford & Greene 2018; Reines et al. 2020). Therefore, the amount of off-centre BHs should not be treated as an absolute standard for evaluating BH dynamical models.

Interestingly, the repositioning algorithm is not as efficient at sinking the BHs at $z = 2$ as the DF. This is because our repositioning algorithm places the BHs at the minimum potential position within the accretion kernel, instead of within the entire halo. The majority of the offset between the BH positions and the halo centre comes from the offset between the minimum-potential position accessible to the BH (i.e. minimum potential in the accretion kernel) and the minimum-potential position in the halo. Such offset can be especially severe at lower redshift, when the size of the accretion kernel gets smaller and mergers happen more frequently, making it easier for the BHs to get stuck at a local minimum.

In the bottom panels, we show the distributions of the BHs' distance to the halo centres under different models. For the no-DF run, again we see that the BHs fail to move towards the halo centre at lower redshift, resulting in a much flatter distribution compared to all the other models. In comparison, when we add DF to the BHs, for both the DF(fid) and the DF(T15) models the distributions are pushed much closer to the halo centre, with a peak around the gravitational softening length. When we then add the gas drag in addition to DF, the peak at $\epsilon_{\rm g}$ becomes slightly higher than those in the DF-only runs. The combination of DF and gas drag, as we would expect from the case studies, is the most effective in sinking the BHs to the halo centres and stabilizing them. Finally, we plot the repositioning model for reference. It does well in putting the BH close to the minimum potential, and often the BHs can be located at the exact minimum-potential position (the distributions peak at 0 for $z = 3.5$). However, as discussed in the previous paragraph, there are cases where the local minimum potential found by the repositioning algorithm does not coincide with the global minimum potential of the halo, and that

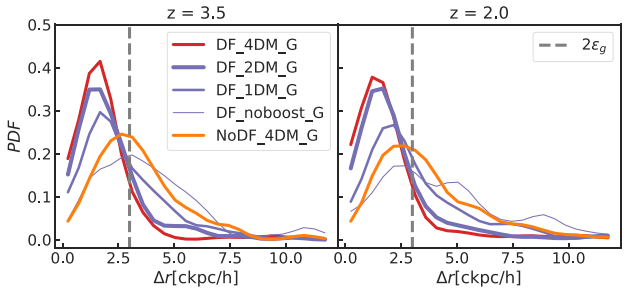

Figure 8. The effect of different choice of dynamical mass on the BHs' sinking status. We compare our fiducial DF model (L15_DF_4DM, red) with $M_{\rm dyn} = 4M_{\rm dm}$ to models with $M_{\rm dyn} = 2M_{\rm dm}$ (thick purple), $M_{\rm dyn} = M_{\rm dm}$ (purple), and $M_{\rm dyn} = M_{\rm BH}$ (thin purple). For $M_{\rm dyn} > M_{\rm dm}$, the majority of the BHs stay within $2\epsilon_{\rm g}$ of the halo centre, while for $M_{\rm dyn} < M_{\rm dm}$, many BHs still stall at a relatively large radius. Noticeably, if we do not boost the dynamical mass of the BHs, the sinking is even worse than if we boost the dynamical mass but do not apply additional DF.

is why we also see non-zero probability density for $\Delta r > 3$ ckpc h^{-1} at $z = 2$.

In Fig. 8, we show the effect of different choice of dynamical mass on the BHs' sinking status. We compare our fiducial DF model (L15_DF_4DM) with $M_{\rm dyn} = 4M_{\rm dm}$ to models with $M_{\rm dyn} = 2M_{\rm dm}$, $M_{\rm dyn} = M_{\rm dm}$, and $M_{\rm dyn} = M_{\rm BH}$. We can see that for $M_{\rm dyn} > M_{\rm dm}$, the majority of the BHs stay within $2\epsilon_{\rm g}$ of the halo centre, while for $M_{\rm dyn} < M_{\rm dm}$, many BHs still stalls at a relatively large radius. Noticeably, if we do not boost the dynamical mass of the BHs, the sinking is even worse than if we boost the dynamical mass but do not apply additional DF.

The statistics we have seen for the models above are consistent with the results from the case studies. This shows that even though for the case studies we have focused mainly on large BHs in one of the biggest halo, a similar trend still applies to other BHs in the cosmological simulations, which are embedded in smaller haloes or subhaloes. Moreover, the fact that we still have off-centre BHs even after adding the DF is in line with predictions from previous simulations (e.g. Governato et al. 1994; Volonteri & Rees 2005; Bellovary et al. 2011; Tremmel et al. 2018), and is not due to the inefficient sinking within our simulation.

5.2 Black hole mass function

Next we look at how different dynamics affect the black hole mass function (BHMF). One problem with the repositioning method is that it places the BHs at the galaxy centre too quickly, which could result in excess accretion and thus a higher mass function. On the other hand, if we do not add any correction to the BH motion, many BHs will not go through efficient accretion and mergers, and we will see a lower mass function. We would expect the BHMF in the DF run to fall between the repositioning case and the no-DF case.

Fig. 9 shows the BHMF from the reposition (L15_Repos_4DM), DF (without gravitational bound check: L15_DF_4DM; with gravitational bound check: L15_DF_4DM_G), and no-DF (without gravitational bound check: L15_NoDF_4DM; with gravitational bound check: L15_NoDF_4DM_G) runs. The reposition model yields the highest mass function, and is the only simulation with more than one 10^8 M$_\odot$ h^{-1} BHs at $z = 2$. This is expected from the overefficient BH mergers and the high-density surroundings in the reposition model. Moreover, it creates increasingly more massive BHs over time, as the increased merger rate produces a stronger effect over time. The no-

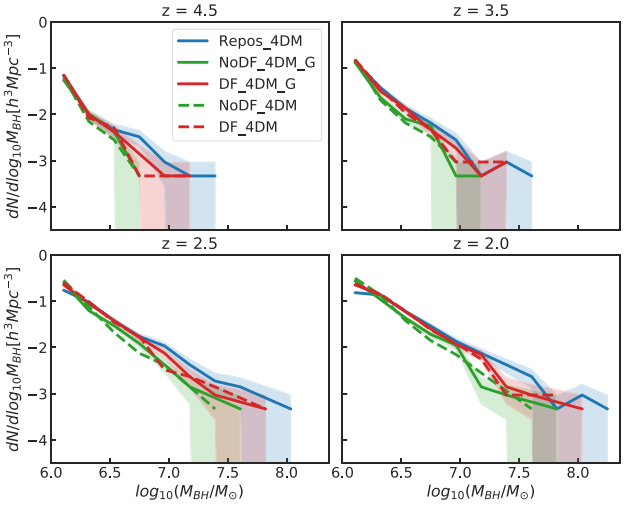

Figure 9. Mass functions for reposition, DF, and no-DF simulations. With reposition (blue), we have the highest mass function and earlier formation of 10^8 M$_\odot$ BHs. The no-DF simulations (green) have lower mass functions, which are expected due to low accretion and merger rates from the BH drifting. The DF model (red) yields a mass function in between.

DF runs produces the lowest mass function due to the off-centring, while the DF mass function falls between the reposition and no-DF case as we expected.

Naively, we would expect the models without gravitational bound checks to produce a higher mass function, because it allows for easier mass accretion via mergers. However, as discussed in Section 4.1.4, this is not the case if we compare the dashed lines and solid lines with the same colours. For example, under the DF model, the L15_DF_4DM_G simulation forms more massive BHs than the L15_DF_4DM simulation, especially at lower redshift. The reason can be traced back to what we have seen in Fig. 2(d): when there is no gravitational bound check, the large momentum injection during a merger kicks the BH out of the halo centre, thus preventing the efficient growth of large BHs.

Considering the relatively large uncertainties due to the limited volume, the difference in the mass function is not very significant. We would expect other factors such as the BH seeding, accretion, and feedback to have a larger effect on the mass function compared to the dynamical models we show here (e.g. Booth & Schaye 2009).

5.3 Dynamical friction time and mergers

Because the reposition method is used in most large-volume cosmological simulations, a post-processing analytical dynamical friction (DF) time is calculated in order to make more accurate merger rate predictions. Now that we have accounted for the DF on-the-fly, we want to study how our numerical mergers with DF compare against the analytical predictions, and how different dynamical models impact the BH merger rate.

In Section 4.2.1, we compared the numerical merging time to the analytical predictions for two merger cases. Now we use the same method to calculate an analytical DF time for all BH mergers in our L15_DF_4DM_G simulation. For each pair, we begin the calculation at the time $t_{\rm beg}$ when the BH pair first comes within 3 ckpch^{-1} of each other, as this mimics the merging time without the gravitational bound check, and is also close to the merging criterion under the reposition model. The numerical DF time $t_{\rm num}$ is the time between the

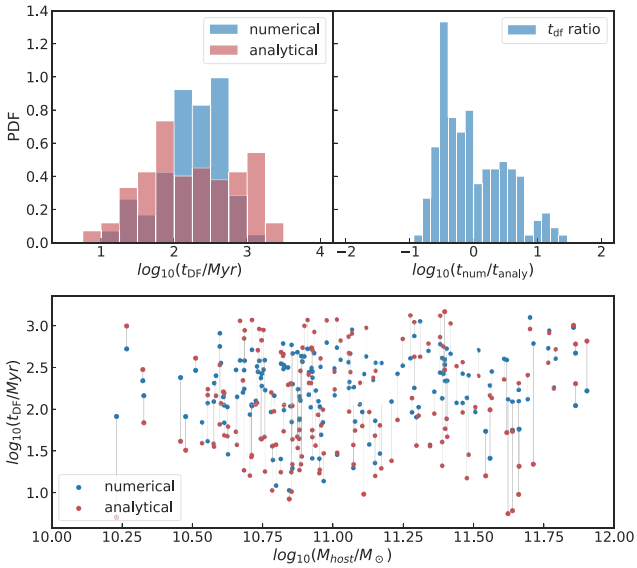

Figure 10. The delay of mergers due to the DF time. Here we compare the numerical DF time, t_{num}, to the analytically calculated time (following equation 17) t_{analy}. Top left: distribution of the DF time from numerical merger (blue) and analytical predictions (red). Top right: ratio between the numerical and analytical t_{df}. Their difference is less than one order of magnitude in all merger cases. Bottom: DF time as a function of the virial mass of the host halo for the numerical (blue) merger and analytical predictions (red). The same merger event is linked by a grey line.

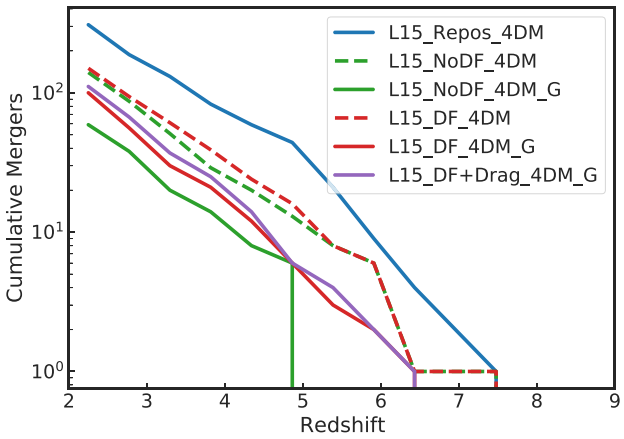

Figure 11. The cumulative mergers for different BH dynamics and merging models. The reposition model (blue solid) predicts more than two times the total mergers compared with the other models. Without the gravitational bound check, the DF model (red dashed) and the no-DF model (green dashed) predict similar numbers of mergers, indicating that the first encounters of the BH pairs are similar under the two models. However, if we add the gravitational bound check, the DF model (red solid) yields ∼50 per cent more mergers compared to the no-correction model. Adding the gas drag in addition to DF (purple solid) raises the mergers by a few.

numerical merger and t_{beg}. The analytical DF time t_{analy} is calculated using the host halo information in the snapshot just before t_{beg} and the BH information at the exact time-step of t_{beg}.

Fig. 10 shows the comparison between the numerical and analytical DF times. In the top panel, we show the distribution of the two times and the distribution of their ratio. We note that for all the mergers happening numerically, t_{analy} does not exceed 2 Gyr, and most have t_{analy} less than 1 Gyr. This means that we do not have many fake mergers that should not merge until much later (or never). Also, the ratio plot shows that the numerical and analytical times are always within an order of magnitude of each other, with most of the numerical mergers earlier than the analytical mergers. The numerical merger time is peaked between 100 Myr and 1 Gyr, whereas the analytical calculation yields a flatter distribution. We would expect t_{analy} to be longer than t_{num}, both because we have a selection bias on t_{DF} by ending the simulation at $z = 2$, and because we numerically merge the BHs when their orbit is still larger than 3 ckpc h^{-1}. However, this does not explain why t_{analy} has a higher probability between 10 and 100 Myr.

To see the individual merger cases in the distribution more clearly, in the lower panel of Fig. 10, we plot all the numerical and analytical DF times as a function of the host halo's virial mass. From this figure we do not see a clear dependence of either DF times on the host halo's virial mass. There is also no strong correlation between the t_{num}/t_{analy} ratio and the halo mass. We do not further investigate the discrepancies between the numerical and analytical results, as these results can vary significantly from system to system.

We note that although the numerical model has free parameters (such as b_{max}, $M_{dyn,\,seed}$) that can impact the merging time (but see Appendix C), it can account for the immediate environment around BH and adjust the DF on-the-fly. More importantly, it also accounts for the interaction between the satellite BH and its own host galaxy, which could reduce the sinking time significantly (e.g. Dosopoulou & Antonini 2017). The analytical model, though verified by *N*-body simulations, does not react to the environment of the merging galaxies by always assuming an NFW profile. Moreover, it only models the sinking of a single BH without embedding it in its host galaxy. Therefore, we expect the numerical result to be a more realistic modelling of the binary sinking process.

After comparing the DF model against the analytical prediction, next we compare different numerical models in terms of the BH merger rate. Fig. 11 shows the cumulative mergers from $z = 8$ to $z = 2$. We have included comparisons between the reposition, DF, and no-DF models, both with and without the gravitational bound check. The reposition model predicts more than twice the total number of mergers compared to the other models. Without the gravitational bound check, the DF and the no-DF models predict similar numbers of mergers, indicating that the first encounters of the BH pairs are similar under the two models. However, if we add the gravitational bound check, the DF model yields ∼50 per cent more mergers compared to the no-DF model, because the addition of DF assists energy loss of the binaries and leads to earlier bound pairs. Finally, the merger rate is not very sensitive to adding the gas drag: the merger rate in the DF-only model is similar to that of the DF+drag model. This can be foreseen in the comparison shown in Fig. 3, where the gas drag is subdominant in magnitude.

6 MERGER RATES IN THE 35 MPC h^{-1} SIMULATIONS

Based on all the previous test of BH dynamics modelling, we have reached the conclusion that the DF+drag model with $M_{dyn} = 4M_{DM}$ is most capable of sinking the BH to the halo centre. Hence, we choose to use this model to run our larger volume simulation `L35_DF+drag_4DM_G` for the prediction of the BH coalescence rate. Besides this model, we also perform a same-size run without the DF, `L35_NoDF_4DM_G`, as a lower limit for the predicted rate. Our `L35` simulations are run down to $z = 1.1$. The BH seed mass is $5 \times$

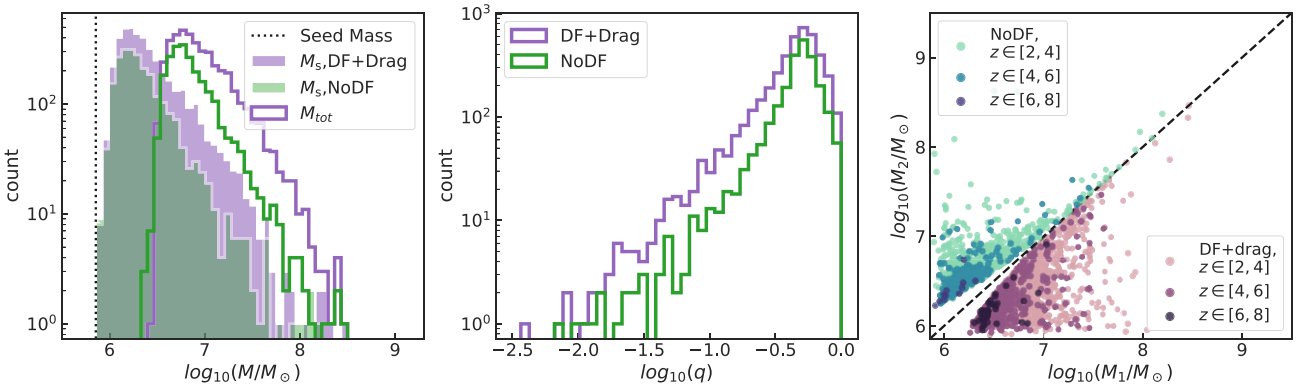

Figure 12. Left: distribution of the mass of the smaller BH (M_s), and distribution of the total mass of the binary (M_{tot}). For both simulations, the mergers in which at least one of the BHs is slightly above the seed mass dominate. The most massive binary has a total mass of $3 \times 10^8\,M_\odot$. Middle: the mass ratio q between the two BHs in the binary. We see a peak at $\log(q) = -0.5$, corresponding to pairs in which one BH is about three times larger than the other. Right: scatter of the two BH masses in the binaries, binned by redshift. To separate the scatter in the two simulations, for the DF+drag run we take M_1 to be the mass of the larger BH, while for the No-DF run M_2 is the larger BH.

$10^5\,M_\odot\,h^{-1}$ and the minimum halo mass for seeding is $10^{10}\,M_\odot\,h^{-1}$. The details of these two simulations are shown in Table 2.

6.1 The binary population

Because this work mainly focuses on model verification and is not intended for accurate merger rate predictions, we do not account for the various post-numerical-merger time delays. These delays can be caused by physical processes such as sub-ckpc scale DF, scattering with stars, gravitational-wave-driven inspiral, and triple MBH systems (e.g. Quinlan 1996; Sesana et al. 2007b; Vasiliev et al. 2015; Dosopoulou & Antonini 2017; Bonetti et al. 2018). We consider all the numerical mergers as true BH merger events. Without any post-process selection, there are 25 224 BHs and 4237 mergers in the `L35_DF+drag_4DM_G` run, and 27 693 BHs and 2349 mergers in the `L35_NoDF_4DM_G` run down to $z = 1.1$.

Fig. 12 shows the distribution of the binary parameters for the mergers in our simulations. For both simulations, there is at least one BH around the seed mass for most mergers, but the peak does not lie at the exact seed mass. The most massive binary has a total mass of $3 \times 10^8\,M_\odot$. For the mass ratio q between the two BHs in the binary, we see a peak at $\log(q) = -0.3$, corresponding to pairs in which one BH is about two times larger than the other. Finally, we show the scatter of the two progenitor masses. The low-mass end of the population deviates more from $q = 1$, while the majority of same-mass mergers come from the 5×10^6–$5 \times 10^7\,M_\odot$ mass range.

Comparing with previous simulations such as Salcido et al. (2016) and Katz et al. (2020), we do not see as many cases of seed–seed mergers, but our distribution in q is similar to that shown in Weinberger et al. (2017) where the larger progenitor is a few times larger than the small progenitor. This is due to our larger BH seed mass of $5 \times 10^5\,M_\odot$ ($10^6\,M_\odot$ in Weinberger et al. 2017): the mass accretion in the early stage is proportional to $M_{\rm BH}^2$, and so during the time before the BH mergers, our BHs accrete more mass compared to the simulations with smaller seeds. This explains why both of our BHs in the binaries are not peaked at the exact seed mass.

6.2 Merger rate predictions

We use the binary population shown in the previous section to predict the merger rate observed per year per unit redshift. The merger rate per unit redshift per year is calculated as

$$\frac{{\rm d}N}{{\rm d}z\,{\rm d}t} = \frac{N(z)}{\Delta z\,V_{\rm c,sim}}\frac{{\rm d}z}{{\rm d}t}\frac{{\rm d}V_{\rm c}(z)}{{\rm d}z}\frac{1}{1+z}, \tag{19}$$

where $N(z)$ is the total number of mergers in the redshift bin z, Δz is the width of the redshift bin, $V_{\rm c,sim}$ is the comoving volume of our simulation box, and ${\rm d}V_{\rm c}(z)$ is the comoving volume of the spherical shell corresponding to the z bin.

We compare our results against recent predictions from hydrodynamical simulations of similar resolution (Salcido et al. 2016; Katz et al. 2020; Volonteri et al. 2020). Here we briefly summarize relevant information about their merger catalogues. The Ref-L100N1504 simulation in the EAGLE suite used in Salcido et al. (2016) has a 2^3 times larger simulation box and slightly higher resolution than our simulations. They seed $1.4 \times 10^5\,M_\odot$ BHs in $1.4 \times 10^{10}\,M_\odot$ haloes. They adopt the reposition algorithm for BH dynamics, but set a distance and relative speed upper limit on the repositioning to prevent BHs from jumping to satellites during fly-by encounters. We compare with their no-delay rate during the inspiral phase. The Illustris simulation used in Katz et al. (2020) has a similar box size, resolution and BH dynamics to the Ref-L100N1504 simulation in EAGLE, except that their halo mass threshold for seeding BHs is $7 \times 10^{10}\,M_\odot$. We compare against their ND model, in which mergers are also taken to occur at the numerical merger time without any delay processes. The Horizon-AGN simulation in Volonteri et al. (2020) is 4^3 times larger than our simulation box, with ∼5 times coarser mass resolution and a BH seed mass of $10^5\,M_\odot$. Instead of seeding BHs in haloes above certain mass threshold, the seeding in Volonteri et al. (2020) is based on the local gas density and velocity dispersion, and seeding is stopped at $z = 1.5$. For BH dynamics, they apply DF from gas, but not from collisionless particles.

Fig. 13 shows our merger rate prediction in the `L35_DF+drag_4DM_G` and `L35_NoDF_4DM_G` simulations. The `L35_DF+drag_4DM_G` run predicts ∼2 mergers per year of observation down to $z = 1.1$, while the `L35_NoDF_4DM_G` run predicts ∼1. The merger rates from both simulations peak at $z \sim 2$. This factor of 2 difference between the two simulations is consistent with what we predicted in the $L_{\rm box} = 15$ Mpc h^{-1} runs in Fig. 11. Although we did not run a $L_{\rm box} = 35$ Mpc h^{-1} simulation with the repositioning model, we expect such a run to predict 5–6 mergers per year down to $z = 1.1$ according to Fig. 11.

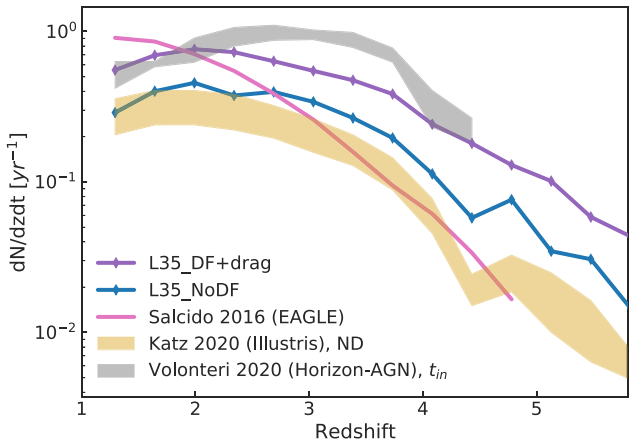

Figure 13. Merger rate per year of observation per unit redshift predicted from our `L35_DF+drag_4DM_G` (purple) and `L35_NoDF_4DM_G` (blue) simulations. For comparison, we also show the prediction from recent hydrodynamical simulations. We include three simulations of similar mass resolution: Volonteri et al. (2020) from the Horizon-AGN simulation (grey), Katz et al. (2020) (yellow) from the Illustris simulation, and Salcido et al. (2016) from the EAGLE simulations (pink). Since we do not apply any post-processing delays after the numerical mergers, we only compare to results without delays.

Generally speaking, our simulations yield similar merger rates as the raw predictions from the previous works of comparable resolution. However, we still note some differences both in the overall rates and in the peak of the rates. We will now elaborate on the reasons for those discrepancies.

First, both of our simulations predict more mergers compared with the Katz et al. (2020) ND model prediction. This is surprising given that in the 15 Mpc h^{-1} runs we saw 2–3 times more mergers when we used the reposition method like Katz et al. (2020) and Salcido et al. (2016) did, comparing to our DF+drag model. Although Katz et al. (2020) cut out ∼25 per cent secondary seed mergers and binaries with extreme density profiles, their rate is still lower after adding the cut-out population. One major reason for the higher rate from our simulation compared to Katz et al. (2020) is the different seeding parameters we use: our minimum halo mass for seeding a BH is $10^{10}\,\mathrm{M}_\odot\,h^{-1}$, which is five times smaller compared with Katz et al. (2020). Moreover, our seeds are a factor of 5 larger. Hence, we have a denser population of BHs in less massive galaxies, which is likely to result in a higher merger rate even compared to the reposition model used in Illustris.

Second, although the rates from EAGLE, Horizon-AGN, and our `L35_DF+drag_4DM_G` simulation cross over at $z \sim 2$, the slope of our merger rate is very different. Volonteri et al. (2020) predict most mergers at $z \sim 3$, whereas the Salcido et al. (2016) rate peaks at $z \sim 1$. This difference can also be traced to the different seeding rate in the three simulations: in Salcido et al. (2016), the seeding rate keeps increasing until $z \sim 0.1$, while we observe a drop in seeding rate at $z = 3$ in our simulations. In Volonteri et al. (2020), due to the different seeding mechanism, BH seeds form significantly earlier, leading to a peak in merger rate at a higher redshift. Hence the peak in the BH merger rate is strongly correlated with the peak in the BH seeding rate.

Finally, besides the effect due to different BH seed models on the merger rate, higher resolution can significantly increase the BH merger rates in the simulations. As was shown in previous work (e.g. Barausse et al. 2020; Volonteri et al. 2020), dwarf galaxies in low-mass haloes can have large numbers of (small mass) BH mergers, and so resolving such haloes and galaxies can increase the BH merger rate significantly. The merger rate differences between high and low resolution and the associate choice for the seed models can lead to large differences in the predictions of merger rates than taking account DF in the BH dynamics.

7 CONCLUSIONS

In this work, we have tested methods for implementing DF from collisionless particles (i.e. stars and dark matter) and gas in low-resolution cosmological simulations (with mass resolution $M_{\rm DM} \sim 10^7\,\mathrm{M}_\odot$ and spatial resolution of $\epsilon_{\rm g} \sim 1\,\mathrm{kpc}\,h^{-1}$), both for single BH evolution/mergers using constrained simulations and for the BH population using unconstrained simulations.

We showed that DF from collisionless particles can effectively assist the BH orbit to decay to within $2\epsilon_{\rm g}$ of the galaxy centre, representing a marked improvement over models that do not include any dynamical correction. Importantly, we find that for our prescription to work well, the dynamical mass of the BHs must be at least twice the mass of the dark matter particles. This is in agreement with results from Tremmel et al. (2015). The DF implementation from Tremmel et al. (2015) (DF(T15)) and our implementation adapted to lower resolution simulations (DF(fid)) result in DF of a similar magnitude, and have comparable effects on the BHs' dynamics. However, we find that our fiducial model is marginally more suitable for low-resolution simulations, as the nature of the calculation results in less noisy force corrections.

After applying the DF and performing the gravitational bound check on the BH pairs, the DF time of the BHs is consistent with analytical predictions, although the variances can be large for individual BHs due to their varied environments. We note that checking whether the two BHs are gravitationally bound at the time of the merger is necessary both for preventing sudden momentum injection on to the BHs and for allowing a more realistic orbital decay time.

By direct comparison of the force magnitudes throughout the simulation, we find that DF from collisionless particles dominates in the majority of cases. The influence of gas drag is highest at the high redshifts, but even then it is typically similar to or less than the contribution from stars and dark matter. This is in broad agreement with the results from Pfister et al. (2019), though we stress that our simulations cannot resolve the structure of gas on the smallest scales. It is possible that interactions with gas are still important, such as migration within circumbinary discs (e.g. Haiman et al. 2009).

Using our fiducial DF+drag model, we calculate the cumulative merger rate down to $z = 1.1$ using a $L_{\rm box} = 35$ Mpc h^{-1} simulation. Without considering any post-merger delays, we predict ∼2 mergers per year for $z > 1.1$, and we lower bound our prediction by a no-dynamical-friction run that predicts ∼1 merger per year. Compared with existing predictions from hydrodynamical simulations (Salcido et al. 2016; Katz et al. 2020; Volonteri et al. 2020), our rates are consistent with the raw merger rates (rates before post-processing delays are added) from previous works of similar resolution. While the dynamics modelling has significant effects (factor of a few according to our experiments) on the BH merger rate, we also found that the different BH seeding criteria and mechanisms account also play a big role in the merger rate predictions.

Our work has demonstrated the feasibility of recovering sub-kpc-scale BH dynamics in low-resolution cosmological simulations by adding the unresolved DF. This is the first step in improving upon the widely adopted reposition model and in tracking the BH dynamics

directly down to the resolution limit. Beyond the resolution limit, we still need to account for several smaller scale binary processes before we can make realistic merger rate predictions (e.g. Quinlan 1996; Sesana et al. 2007b; Haiman et al. 2009; Vasiliev et al. 2015; Dosopoulou & Antonini 2017; Kelley et al. 2017; Bonetti et al. 2018; Katz et al. 2020). Nevertheless, having access to the full dynamical information of the binary at the time of the numerical merger also helps us to better model these small-scale processes. We will leave the analysis of post-merger delays for future works.

There are still several aspects of the DF model that remain somewhat uncertain. Most importantly, the parameters (e.g. $b_{\rm max}$, $M_{\rm dyn,seed}$) in the current DF model can induce uncertainties in the sinking time-scale and the merger rate predictions. For example, reducing $M_{\rm dyn,seed}$ to a value similar to or below the dark matter particle mass will reduce the merger rate by a factor of 2 or more. Our current choice is well tested in our simulations, but it is still subject to the limitations of our spatial and mass resolution. The limit in the $M_{\rm BH}/M_{\rm DM}$ ratio also hinders comprehensive studies of BH seeding scenarios in the cosmological context. We would need insights from high-resolution simulations (e.g. Dosopoulou & Antonini 2017; Pfister et al. 2019) to better model the dynamics of low-mass BHs within cosmological simulations.

ACKNOWLEDGEMENTS

We thank Marta Volonteri for discussions on the merger rate comparisons. We thank the anonymous reviewer for reading the paper carefully and provide helpful comments. The simulations were performed on the Bridges and Vera clusters at the Pittsburgh Supercomputing Center (PSC). TDM acknowledges funding from NSF ACI-1614853, NSF AST-1616168, NASA ATP 19-ATP19-0084, NASA ATP 80NSSC18K101, NASA ATP NNX17AK56G, and 80NSSC20K0519. SB was supported by NSF grant AST-1817256.

DATA AVAILABILITY

The data underlying this paper will be shared on reasonable request to the corresponding author.

REFERENCES

Abbott B. P. et al., 2016, Phys. Rev. Lett., 116, 061102
Amaro-Seoane P. et al., 2012, Classical Quantum Gravity, 29, 124016
Amaro-Seoane P. et al., 2017, preprint (arXiv:1702.00786)
Antonini F., Merritt D., 2012, ApJ, 745, 83
Barausse E., 2012, MNRAS, 423, 2533
Barausse E., Dvorkin I., Tremmel M., Volonteri M., Bonetti M., 2020, ApJ, 904, 16
Barrows R. S., Comerford J. M., Greene J. E., 2018, ApJ, 869, 154
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nature, 287, 307
Bellovary J. M., Governato F., Quinn T. R., Wadsley J., Shen S., Volonteri M., 2010, ApJ, 721, L148
Bellovary J., Volonteri M., Governato F., Shen S., Quinn T., Wadsley J., 2011, ApJ, 742, 13
Bhowmick A. K., Di Matteo T., Feng Y., Lanusse F., 2018, MNRAS, 474, 5393
Biernacki P., Teyssier R., Bleuler A., 2017, MNRAS, 469, 295
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ
Bonetti M., Haardt F., Sesana A., Barausse E., 2018, MNRAS, 477, 3910
Bonetti M., Sesana A., Haardt F., Barausse E., Colpi M., 2019, MNRAS, 486, 4044
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Chandrasekhar S., 1943, ApJ, 97, 255
Chapon D., Mayer L., Teyssier R., 2013, MNRAS, 429, 3114
Crain R. A. et al., 2015, MNRAS, 450, 1937
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
DeGraf C., Sijacki D., 2020, MNRAS, 491, 4973
Derdzinski A. M., D'Orazio D., Duffell P., Haiman Z., MacFadyen A., 2019, MNRAS, 486, 2754
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, MNRAS, 467, 4243
Dosopoulou F., Antonini F., 2017, ApJ, 840, 31
Dubois Y., Pichon C., Devriendt J., Silk J., Haehnelt M., Kimm T., Slyz A., 2013, MNRAS, 428, 2885
Dubois Y. et al., 2014, MNRAS, 444, 1453
Escala A., Larson R. B., Coppi P. S., Mardones D., 2004, ApJ, 607, 765
Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, MNRAS, 455, 2778
Governato F., Colpi M., Maraschi L., 1994, MNRAS, 271, 317
Haiman Z., Kocsis B., Menou K., 2009, ApJ, 700, 1952
Hinshaw G. et al., 2013, ApJS, 208, 19
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hopkins P. F., 2013, MNRAS, 428, 2840
Huang K.-W., Di Matteo T., Bhowmick A. K., Feng Y., Ma C.-P., 2018, MNRAS, 478, 5063
Jenet F. A., Lommen A., Larson S. L., Wen L., 2004, ApJ, 606, 799
Jenet F. A., Hobbs G. B., Lee K. J., Manchester R. N., 2005, ApJ, 625, L123
Katz N., Hernquist L., Weinberg D. H., 1999, ApJ, 523, 463
Katz M. L., Kelley L. Z., Dosopoulou F., Berry S., Blecha L., Larson S. L., 2020, MNRAS, 491, 2301
Kazantzidis S. et al., 2005, ApJ, 623, L67
Kelley L. Z., Blecha L., Hernquist L., 2017, MNRAS, 464, 3131
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36
Kulier A., Ostriker J. P., Natarajan P., Lackner C. N., Cen R., 2015, ApJ, 799, 178
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, ApJ, 742, 103
Magorrian J. et al., 1998, AJ, 115, 2285
Mangiagli A., Bonetti M., Sesana A., Colpi M., 2019, ApJ, 883, L27
Marshall M. A., Ni Y., Di Matteo T., Wyithe J. S. B., Wilkins S., Croft R. A. C., Kuusisto J. K., 2020, MNRAS, 499, 3819
Marshall M. A., Wyithe J. S. B., Windhorst R. A., Di Matteo T., Ni Y., Wilkins S., Croft R. A. C., Mechtley M., 2021, MNRAS, 506, 1209
Mingarelli C. M. F. et al., 2017, Nat. Astron., 1, 886
Natarajan P., Pacucci F., Ferrara A., Agarwal B., Ricarte A., Zackrisson E., Cappelluti N., 2017, ApJ, 838, 117
Ni Y., Di Matteo T., Feng Y., Croft R. A. C., Tenneti A., 2018, MNRAS, 481, 4877
Ni Y., Di Matteo T., Feng Y., 2021, MNRAS, 509, 3043
Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208
Ostriker E. C., 1999, ApJ, 513, 252
Pelupessy F. I., Papadopoulos P. P., van der Werf P., 2006, ApJ, 645, 1024
Pfister H., Volonteri M., Dubois Y., Dotti M., Colpi M., 2019, MNRAS, 486, 101
Pillepich A. et al., 2018, MNRAS, 473, 4077
Quinlan G. D., 1996, New Astron., 1, 35
Reines A. E., Condon J. J., Darling J., Greene J. E., 2020, ApJ, 888, 36
Ricarte A., Natarajan P., 2018, MNRAS, 481, 3278
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Salcido J., Bower R. G., Theuns T., McAlpine S., Schaller M., Crain R. A., Schaye J., Regan J., 2016, MNRAS, 463, 870
Schaye J. et al., 2015, MNRAS, 446, 521

Sesana A., Volonteri M., Haardt F., 2007a, MNRAS, 377, 1711
Sesana A., Haardt F., Madau P., 2007b, ApJ, 660, 546
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Soltan A., 1982, MNRAS, 200, 115
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Taffoni G., Mayer L., Colpi M., Governato F., 2003, MNRAS, 341, 434
Tenneti A., Di Matteo T., Croft R., Garcia T., Feng Y., 2018, MNRAS, 474, 597
Tremmel M., Governato F., Volonteri M., Quinn T. R., 2015, MNRAS, 451, 1868
Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121
Tremmel M., Governato F., Volonteri M., Quinn T. R., Pontzen A., 2018, MNRAS, 475, 4967
van de Weygaert R., Bertschinger E., 1996, MNRAS, 281, 84
Vasiliev E., Antonini F., Merritt D., 2015, ApJ, 810, 49
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Volonteri M., Rees M. J., 2005, ApJ, 633, 624
Volonteri M., Silk J., Dubus G., 2015, ApJ, 804, 148
Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, MNRAS, 460, 2979
Volonteri M. et al., 2020, MNRAS, 498, 2219
Waters D., Wilkins S. M., Di Matteo T., Feng Y., Croft R., Nagai D., 2016, MNRAS, 461, L51
Weinberger R. et al., 2017, MNRAS, 465, 3291
Wilkins S. M., Feng Y., Di Matteo T., Croft R., Lovell C. C., Waters D., 2017, MNRAS, 469, 2517
Wurster J., Thacker R. J., 2013, MNRAS, 431, 2513

APPENDIX A: DYNAMICAL MASS AND RESOLUTION EFFECT

A1 Varying dynamical mass

One major difference between our model and previous modelling of the DF is that we boost the mass term during the early stage of BH growth by a factor of $k_{\rm dyn} = M_{\rm dyn,seed}/M_{\rm BH}$. This is to prevent the drifting of the BHs due to dynamical heating when the BH mass is below the dark matter particle mass in the context of large and low-resolution cosmological simulations.

Here we show the effect of setting different $k_{\rm dyn}$ by running three simulations with the same resolution and DF models, but various $k_{\rm dyn}$ ratios. They are listed in Table 1 as `DF_4DM_G`, `DF_2DM_G`, and `DF_1DM_G`, with $k_{\rm dyn} = 4$, 2, and 1, respectively.

Fig. A1(a) shows the evolution of the same BH for different $k_{\rm dyn}$. By comparing the three cases, we can see that the BH's behaviour is very similar for all the physical quantities we have plotted. However, we also note that the similar behaviour of different $M_{\rm dyn}$ is case dependent. The case we present here is a BH within a large density peak where the BH is subject to a deep potential and can sink more easily, but the sinking of BHs in shallower potentials can be more sensitive to the seed dynamical mass. Nevertheless, $k_{\rm df} = 2$ is generally sufficient to assist the sinking of most BHs and produces similar merger rates to $k_{\rm df} = 4$ (see Appendix C). The convergence at $k_{\rm df} = 2$ is consistent with the $M_{\rm BH}/M_{\rm DM} = 3$ ratio used in Tremmel et al. (2018), and relaxes the ratio used in previous works (e.g. Tremmel et al. 2015; Pfister et al. 2019) of $M_{\rm BH}$–$10M_{\rm DM}$.

A2 Resolution effect

Here we show how our model performs under different resolutions. For this experiment we use our fiducial resolution run `DF_4DM_G`, a higher resolution run `DF_HR_4DM_G` with the same $k_{\rm df}$, but a factor of 3 difference in the mass resolution, and a high-resolution run `DF_HR_12DM_G` with the same $M_{\rm dyn,seed}$ as the fiducial resolution run. We would want the BHs to behave similarly independent of resolution if the $M_{\rm dyn,seed}/M_{\rm DM}$ is kept constant.

Fig. A1(b) shows the same BH in the simulations with different resolution. In the high-resolution run `DF_HR_4DM_G`, even though the seeding dynamical mass is three times smaller than the low-resolution run, the sinking time remains the same. Furthermore, if we keep the absolute seeding dynamical mass the same in the low-resolution and high-resolution runs (by comparing `DF_HR_12DM_G` with `DF_4DM_G`), the BHs still shows similar evolution. This indi-

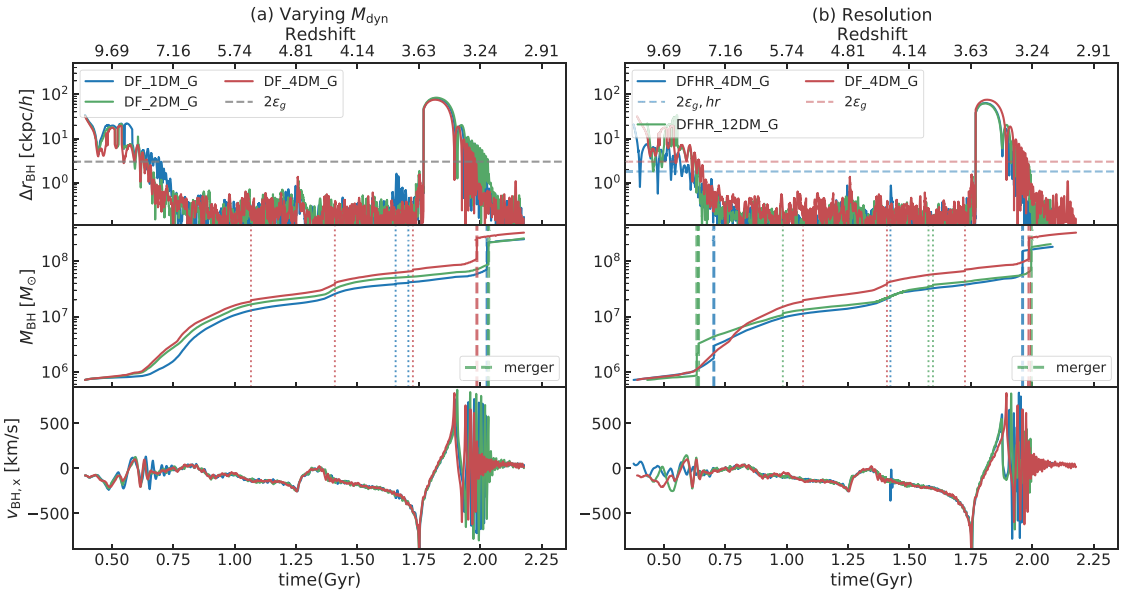

Figure A1. (a) Comparisons of different BH seed dynamical mass. The effect of varying $M_{\rm dyn,\ seed}$ is small in this case. But this is partially due to the large BH we pick. (b) Comparison with higher resolution run with the same $M_{\rm dyn}/M_{\rm DM}$ ratio.

cates that a constant $k_{df} = M_{dyn,seed}/M_{DM}$ is robust under different resolutions, and our model of dynamical mass does converges to the true BH mass if we go to higher resolutions.

APPENDIX B: DF(FID) VERSUS DF(T15): CASES OF SMALLER BLACK HOLES EVOLUTION

In Section 4.1, we compared the two DF models by showing the example of an early forming BH located at the centre of the largest halo in the simulation. However, that BH might not be representative of the entire BH population due to its early seeding and large mass. Now we pick more cases of smaller BHs to demonstrate the differences/similarities between the models. In particular, we will look at how the smaller BHs are affected by the DF(fid)/DF(T15) implementation.

Fig. B1 shows the evolution of three small BHs in the `DF_4DM_G` and the `DF(T15)_4DM_G` simulations. We plot three $M_{BH} < 5 \times 10^6$ M$_\odot$ BHs. In these cases, the number of particles within the SPH kernel is still at least an order of magnitude more than 100 at lower redshift, and so the density calculated in DF(T15) still tends to be larger but more noisy. The value of the Coulomb logarithm is now mainly affected by b_{max}, because we do not see as much noise in the velocity of the surrounding particles as in the case of a very large BH. The density and the Coulomb logarithm counteract each other, and the magnitude of the DF is similar in the two models.

These cases again verify that the two models are consistent with each other, with DF(T15) a more localized implementation than DF(fid). The choice of DF(fid) as our fiducial model is mainly due to our resolution limit.

APPENDIX C: EFFECT OF MODEL PARAMETERS ON THE MERGER RATE

For the merger rate predictions in Section 6, we use the DF+drag model with $b_{max} = 10$ ckpch^{-1} and $M_{dyn,seed} = 4M_{DM}$. In this section, we will show that the merger rate prediction is not sensitive to the choice of these two parameters, and hence our prediction is relatively robust against parameter variations within a reasonable range.

Fig. C1 shows the cumulative merger rates for different values of b_{max} in the $L_{box} = 15$ Mpc h^{-1} simulations. We tested b_{max} values of 3 ckpc h^{-1},10 ckpc h^{-1}, and 30 kpc, and the difference in the cumulative merger rate is less than 10 per cent. The difference between the DF(T15) model and the DF(fid) model with $b_{max} =$ 1.5 ckpc is also very small. Hence, although different choices of b_{max} change the magnitude of the DF, it does not affect the merger rate predictions significantly.

We also test a lower value of $M_{dyn,seed} = 2M_{DM}$ using the L_{box} = 15 Mpc h^{-1} simulation. The resulting cumulative merger rate prediction is also shown in Fig. C1. Compared with the similar run with $M_{dyn,seed} = 4M_{DM}$, the earliest merger is slightly postponed, but the cumulative rate at $z \sim 2$ has very little difference. Therefore, even though for the predictions in Section 6 we have chosen a particular set of parameter values, changing those parameters would not affect the result significantly given the larger effects of other factors such as the resolution and seeding.

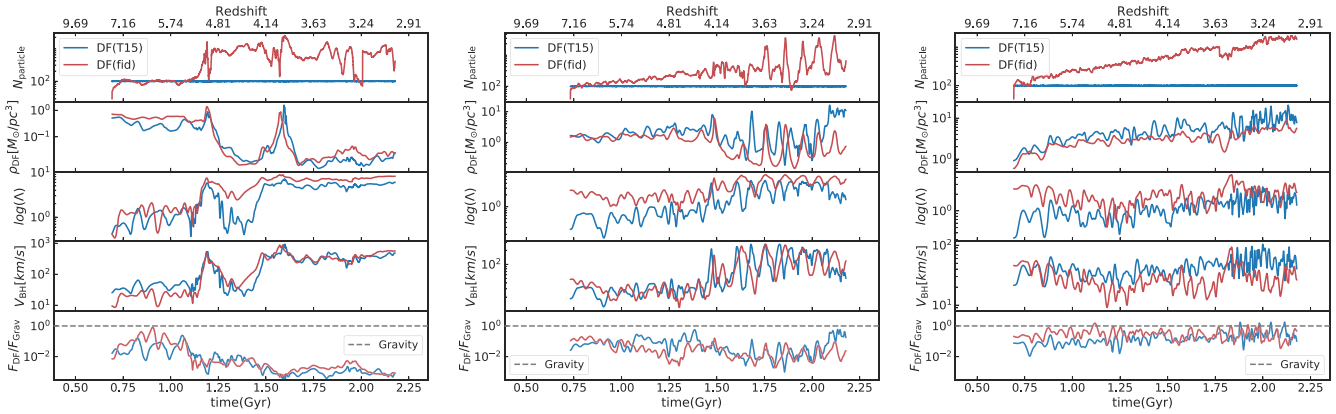

Figure B1. Components of the DF in the `DF(fid)_4DM_G` (red) and the `DF(T15)_4DM_G` (blue) simulations, for three $M < 5 \times 10^6$ M$_\odot$ BHs. In these cases, the number of particles within the SPH kernel is still at least an order of magnitude more than 100 at lower redshift. The value of the Coulomb logarithm is now mainly affected by b_{max}, because we do not see as much noise in the velocity of the surrounding particles as in the case of a very large BH. In all three cases shown, the magnitude of the DF is similar in the two models.

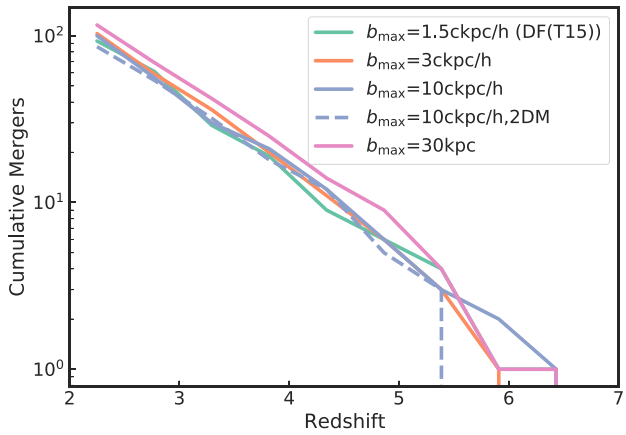

Figure C1. The cumulative merger rates for different values of b_{max}, in the $L_{box} = 15$ Mpc h^{-1} simulations. We tested b_{max} values of 3 ckpc h^{-1}, 10 ckpc h^{-1}, and 30 kpc, and the difference in the cumulative merger rate is less than 10 per cent. The difference between the DF(fid) model and the DF(T15) model with $b_{max} = 1.5$ ckpc is also very small. Hence, although different choices of b_{max} change the magnitude of the DF, it does not affect the merger rate predictions significantly.

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