Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **491,** 2301–2317 (2020)
Advance Access publication 2019 November 7

doi:10.1093/mnras/stz3102

Probing massive black hole binary populations with LISA

Michael L. Katz [1,2]★ Luke Zoltan Kelley,[2] Fani Dosopoulou,[3,4] Samantha Berry,[2,5] Laura Blecha[6] and Shane L. Larson[1,2]

[1]*Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA*
[2]*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Evanston, IL 60201, USA*
[3]*Princeton Center for Theoretical Science, Princeton University, Princeton, NJ 08544, USA*
[4]*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[5]*Department of Physical Science, Harry S. Truman College, Chicago, IL 60640, USA*
[6]*Physics Department, University of Florida, Gainesville, FL 32611, USA*

Accepted 2019 October 24. Received 2019 October 15; in original form 2019 August 15

ABSTRACT

ESA and NASA are moving forward with plans to launch Laser Interferometer Space Antenna (LISA) around 2034. With data from the Illustris cosmological simulation, we provide analysis of LISA detection rates accompanied by characterization of the merging massive black hole (MBH) population. MBHs of total mass $\sim 10^5$–$10^{10}\,M_\odot$ are the focus of this study. We evolve Illustris MBH mergers, which form at separations of the order of the simulation resolution (∼kpc scales), through coalescence with two different treatments for the binary MBH evolutionary process. The coalescence times of the population, as well as physical properties of the black holes, form a statistical basis for each evolutionary treatment. From these bases, we Monte Carlo synthesize many realizations of the merging MBH population to build mock LISA detection catalogues. We analyse how our MBH binary evolutionary models affect detection rates and the associated parameter distributions measured by LISA. With our models, we find MBH binary detection rates with LISA of ~ 0.5–$1\ \mathrm{yr}^{-1}$ for MBHs with masses greater than $10^5\,M_\odot$. This should be treated as a lower limit primarily because our MBH hole sample does not include masses below $10^5\,M_\odot$, which may significantly add to the observed rate. We suggest reasons why we predict lower detection rates compared to much of the literature.

Key words: gravitational waves.

1 INTRODUCTION

With the selection of the Laser Interferometer Space Antenna (LISA; Amaro-Seoane et al. 2017) by the European Space Agency (ESA) for its L3 mission, the massive black hole (MBH) community will gain an important tool for understanding the physics and evolutionary history of MBHs. Following a galaxy merger, the two MBHs from the galactic centres eventually form a bound pair and become a binary. This evolving MBH pair radiates gravitational waves at a range of low frequencies ∼1 nHz−1 mHz. At the higher end of this range, when the two MBHs are nearer to coalescing, their gravitational wave emission will make them a prime target for the LISA mission. LISA is sensitive to MBH binaries of $\sim 10^3$–$10^9\,M_\odot$ (Klein et al. 2016; Amaro-Seoane et al. 2017; Katz & Larson 2019) at frequencies from 0.1 to 10 mHz corresponding to separations of less than 10^3 Schwarzschild radii. Signals from these binaries will allow the scientific community to study the origin and evolution of MBHs over cosmic time (eLISA Consortium 2013; Barausse et al. 2015; Amaro-Seoane et al. 2017). Additionally, MBH binary signals have the potential to reach very high signal-to-noise ratios (SNR, also referred to as ρ), unattainable for ground-based gravitational wave detectors (eLISA Consortium 2013; Barausse et al. 2015). This will allow for high precision measurements of cosmological parameters as well as a greater understanding of fundamental physics (Gair et al. 2013; Barausse et al. 2015).

In this paper, we use the Illustris large-scale cosmological simulations (Genel et al. 2014; Vogelsberger et al. 2014a,b; Sijacki et al. 2015) to analyse MBH binary populations, their dynamics, and LISA detection prospects. A few papers have provided similar studies. Blecha et al. (2016) performed an analysis on MBH recoil kicks using Illustris data. Kelley, Blecha & Hernquist (2017a) and Kelley et al. (2017b, 2018) performed a similar analysis to ours in relation to Pulsar Timing Array (PTA) predictions. There have been many papers predicting rates for LISA (e.g. Berti et al. 2016; Klein et al. 2016; Salcido et al. 2016; Bonetti et al. 2019), as well as new predictions for the TianQin gravitational wave observatory, which is similar in construction to LISA (Wang et al. 2019).

★ E-mail: mikekatz04@gmail.com

Many rate prediction papers are built from semi-analytical models (SAM). Most rate predictions with large seeds from SAMs, like those in Klein et al. (2016) and Berti et al. (2016), use seeds of the order of 10^4 $M_\odot$. These masses are unresolved in the Illustris simulations; therefore, our overall rate estimates will be lower than those predicted in these papers. Similarly, as delays between galaxy mergers and their central MBH mergers have been included in SAMs, the predicted rates have lowered a bit from $\sim$20 yr^{-1} (Arun et al. 2009; Sesana et al. 2011) to $\sim$8 yr^{-1} (Berti et al. 2016; Klein et al. 2016). However, the inclusion of triple MBH interactions has increased the rate to $\sim$20 yr^{-1} (Bonetti et al. 2019). In addition to using MBH binaries from Illustris, we use improved prescriptions for the MBH binary evolutionary process (delay prescriptions) from Dosopoulou & Antonini (2017) and Kelley et al. (2017a,b).

Black holes with masses of $\sim10^6$–10^{10} $M_\odot$ are usually considered MBHs. These MBHs are believed to exist in the centres of most galaxies of considerable size (Soltan 1982; Kormendy & Richstone 1995; Magorrian et al. 1998). This stems from observations of dynamics in the centre of other galaxies, as well as our own Milky Way Galaxy. The MBH in the centre of the Milky Way Galaxy has been constrained to a mass of $\sim4.1 \times 10^6$ $M_\odot$ (Boehle et al. 2016). Intermediate mass black holes (IMBH) with masses of $\sim10^2$–10^5 $M_\odot$ have been theorized and observations of IMBHs have been suggested (e.g. Lin et al. 2018; Bellovary et al. 2019), but remain uncertain. However, electromagnetic observations are beginning to find MBHs in between these ranges (10^5–10^6 $M_\odot$) in dwarf galaxies (Reines, Greene & Geha 2013; Moran et al. 2014; Satyapal et al. 2014; Lemons et al. 2015; Sartori et al. 2015; Pardo et al. 2016; Nguyen et al. 2018, 2019).

Within the ΛCDM paradigm (e.g. White & Frenk 1991), galactic haloes merge. If conditions are right, the MBHs in their centres can form a binary (Begelman, Blandford & Rees 1980). The pair of MBHs will inspiral via interactions with surrounding gas and stars until they reach close enough separations to emit detectable gravitational waves (e.g. Sesana et al. 2004; Haiman, Kocsis & Menou 2009; Sesana 2010; Roedig et al. 2011; Dosopoulou & Antonini 2017; Rasskazov & Merritt 2017; Kelley et al. 2017a,b). The only way to study MBHs to date is through electromagnetic observations, which lead to a potential observational bias in the MBH population measurements (Shen et al. 2008; McConnell & Ma 2013; Shankar et al. 2016; Rasskazov & Merritt 2017). LISA will add a strong and independent method for studying these exotic objects. Additionally, the combined measurement of the luminosity distance to a binary with LISA and the electromagnetic measurement of a binary redshift can be used as a 'standard siren' to measure the Hubble parameter (Schutz 1986; Holz & Hughes 2005; Abbott et al. 2017). To perform this measurement, the host galaxy of the binary will generally be needed to get the redshift value. Many groups suggest that some periodic active galactic nuclei could be MBH binaries (e.g. Graham et al. 2015; Charisi et al. 2016; Liu et al. 2016). This type of measurement can help identify the host galaxy and therefore the redshift of the binary. However, even without an electromagnetic counterpart, LISA measurements can help constrain various cosmological parameters (Petiteau, Babak & Sesana 2011). In addition to cosmological parameters, EM counterparts of MBH binaries can help illuminate a variety of astrophysical processes, including accretion physics and galaxy evolution (Burke-Spolaor 2013; Bogdanović 2015).

Understanding formation channels of these large MBHs and how they relate to galaxy formation models is an active area of research. Leading theoretical ideas about MBH formation channels largely differ in their considerations for the mass of MBH seeds at early times in the evolution of the Universe. One such scenario involves the direct collapse of pre-galactic haloes with a seed mass of the order of 10^4–10^6 $M_\odot$ at redshifts of 10–20 (Loeb & Rasio 1994; Begelman, Volonteri & Rees 2006; Latif et al. 2013; Habouzit et al. 2016; Ardaneh et al. 2018; Dunn et al. 2018). Another scenario involves seeds of $\sim10^3$–10^4 $M_\odot$ from runaway cluster collapse (Omukai, Schneider & Haiman 2008; Devecchi & Volonteri 2009; Davies, Miller & Bellovary 2011; Katz, Sijacki & Haehnelt 2015). For smaller seeds, MBH formation channels involve seeds from the collapse of large Population III stars into black holes of the order of 10^2 $M_\odot$. This would occur at earlier times in cosmic history at redshifts of 20–50 (Haiman, Abel & Rees 2000; Fryer, Woosley & Heger 2001; Heger et al. 2003; Volonteri, Madau & Haardt 2003; Alvarez, Wise & Abel 2009; Tanaka & Haiman 2009). Regardless of the formation channel, observations of active galactic nuclei of $\sim10^9$ $M_\odot$ in the centres of galaxies at $z \sim 6$–7 indicate MBHs in the Universe must have formed quickly on cosmological time-scales after the big bang, within 1 billion yr (Fan et al. 2001a,b, 2006; Mortlock et al. 2011).

It may be that reality is a combination of these three theories. Additionally, different accretion types and spin values can affect the growth rate of MBHs (Plowman et al. 2010; Plowman, Hellings & Tsuruta 2011; Sesana et al. 2011). Due to the possible measurement bias of AGNs with EM observations (Lauer et al. 2007; Schulze & Wisotzki 2011), these possibly conflicting scenarios will benefit greatly from LISA detections at higher redshifts across the entire mass spectrum.

The coalescence of two MBHs occurs after a long dynamical process forcing the two MBHs to decay from $\sim$kpc separations down to merger (Begelman et al. 1980; Yu 2002; Merritt & Milosavljević 2005). Once two galaxies have merged, their central MBHs sink to the centre via dynamical friction from interactions with surrounding stars (Chandrasekhar 1943; Quinlan 1996; Quinlan & Hernquist 1997). Once the two MBHs become gravitationally bound, the dynamical friction formalism breaks down, and individual interactions between singular stars and the binary must be considered. These interactions extract angular momentum from the binary, driving them closer to each other (e.g. Merritt 2013; Vasiliev & Merritt 2013). This regime is the 'stellar hardening' or 'loss-cone scattering' regime, which refers to the specific cone in parameter space where stars have to exist in order to extract angular momentum from the binary (Frank & Rees 1976; Lightman & Shapiro 1977). Following these interactions, the binary becomes close enough to interact with a circumbinary gas disc if gas is present. Generally, the torque from the gas disc is expected to bring the binary closer to coalescence (Haiman et al. 2009). However, there is growing evidence that specific binary parameters and gas disc properties can lead to the gas disc forcing the binary outwards to larger separations (Moody, Shi & Stone 2019; Muñoz, Miranda & Lai 2019). After interaction with a gas disc, or if there is little to no gas present, the binary will enter the gravitational wave regime where it will evolve until coalescence. Once the binary enters the gravitational wave regime, its dynamics follow the formalism of Peters & Mathews (1963) at small separations of $\sim$100–1000 Schwarzschild radii (Kelley et al. 2017a). Several studies have predicted some residual eccentricity when MBH binaries enter the LISA frequency band (Amaro-Seoane et al. 2010; Porter & Sesana 2010; Dosopoulou & Antonini 2017; Mirza et al. 2017). However, we will treat all binaries as circular for convenience. This is also conservative as the eccentricity will cause the binaries to merge faster.

We will examine binary lifetime models from Dosopoulou & Antonini (2017) and Kelley et al. (2017a,b). In Section 3.1, we will

discuss our models and their specific mathematical approaches to binary lifetime calculations. For this paper, in order to match with cosmological parameters in the Illustris simulation, we assume a WMAP-9 cosmology with $H_0 = 70.4\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.2726$, $\Omega_\mathrm{b} = 0.0456$, and $\Omega_\mathrm{vac} = 0.7274$ (Hinshaw et al. 2013).

2 ILLUSTRIS SIMULATION

The Illustris cosmological simulations are a suite of simulations evolving gas cells, dark matter (DM), star, and MBH particles from $z = 137$ to $z = 0$ in a cube of side length 106.5 comoving Mpc. Illustris is based on the moving, unstructured-mesh hydrodynamic code AREPO (Springel 2010). Illustris reproduced statistics of large-scale galaxy assembly as well as internal structures of elliptical and spiral galaxies (Vogelsberger et al. 2014b). In particular, we extracted data from Illustris-1, the highest resolution simulation in the suite, with 1820^3 gas cells and DM particles. DM particles have a mass resolution and typical gravitational-softening length of $\sim 6.3 \times 10^6\ \mathrm{M_\odot}$ and 1.5 kpc, respectively. At redshift zero, there are over 3×10^8 star particles, which have a resolution of $\sim 1.3 \times 10^6\ \mathrm{M_\odot}$ and a typical softening length of 700 pc.

MBHs are implemented in Illustris as massive sink particles. When haloes attain a total mass of $7.1 \times 10^{10}\ \mathrm{M_\odot}$, they are seeded with an MBH of mass $1.42 \times 10^5\ \mathrm{M_\odot}$ if it does not already have an MBH in it (Sijacki et al. 2015); the highest density gas cell in the halo is converted to the MBH particle. At this point, the initial dynamical MBH mass will be the same as the gas cell in its previous state. However, the MBH particle is assigned an internal mass of the seed mass, which is tracked from this point (Vogelsberger et al. 2013).

MBH particles grow by Eddington-limited, Bondi–Hoyle accretion from their parent gas cells initially, and then from its nearby gas cells after the dynamical mass of the particle becomes equal to its internal mass. From this point, the dynamical mass and internal MBH mass increase in tandem. The similarity in mass between the MBH particles, gas cells, and star/DM particles would cause the MBH particle to scatter around haloes in an unphysical manner without settling down in the centre of the halo. Therefore, MBH particles are repositioned to the potential minimum of the host halo at every time-step (we refer to this as the 'repositioning algorithm').

For more overview on the Illustris simulations, see Vogelsberger et al. (2013) and Torrey et al. (2014). See Vogelsberger et al. (2014a), Genel et al. (2014), and Sijacki et al. (2015) for detailed simulation results and comparisons of the simulations to observations. The initial data sets used in this study were all obtained at www.illustris-project.org (Nelson et al. 2015). Using this data, we perform more post-processing to create the final data sets used in our analysis. We will further explain this in the following sections.

2.1 Massive black hole merger population

Two MBH particles are merged in the Illustris simulation when they come within a smoothing length of each other (∼kpc). Since these mergers occur at larger scales, we treat this merger event as the formation of the binary, from which we evolve the binary to coalescence with sub-grid models. Following Kelley et al. (2017a), we adopt the term 'merger' to indicate this simulation-only process: the combination of two MBH simulation particles into one, indicating the formation of a binary on ∼kpc scales. We will refer to the final combination of two realistic MBHs into one as the MBH binary 'coalescence.'

Over the course of the simulation, detailed MBH and host galaxy information is saved in a series of 135 snapshots. Higher time-resolution data was saved for each merger, including the time of the merger and the constituent MBH masses (Blecha et al. 2016, Kelley et al. 2017a). We extract detailed properties for all MBHs in the simulation at each snapshot, as well as relevant, global properties for each of their host galaxies from the Illustris 'Group Catalogs' (Nelson et al. 2015). In addition to the higher time-resolution merger data set, we extract information about host galaxies related to the mergers. For each merger, we locate the host galaxies of the constituent MBHs at the snapshot immediately preceding the merger, as well as the host galaxy of the remnant MBH at the snapshot immediately following the merger. For these galaxies, we not only attained global information, we also gather information about their specific distribution of gas, stellar, and DM constituents. The last data needed for our analysis is the Sublink merger trees (Rodriguez-Gomez et al. 2015) to follow galaxies from snapshot to snapshot.

Throughout the cosmic history within Illustris, there are 23 708 MBH merger events. However, a fraction of these mergers are artificial. First, the friends-of-friends (FOF) halo finder will occasionally associate two haloes as one. When this occurs, the aforementioned repositioning algorithm will force the two MBHs in the centres of each galactic halo to the new potential minimum determined during this misstep by the FOF finder. This causes the two MBHs to merge. After this 'fly-by' encounter, the two galaxies may separate into two distinct haloes as seen by the FOF finder. When this occurs, there will be one galaxy without a central MBH. At this point, a new MBH is seeded in this galaxy causing future artificial mergers to inflate the merger catalogue. Similarly, the FOF finder may identify a transient matter overdensity, subsequently seeding a low-mass MBH into the overdensity. This newly seeded MBH is then quickly merged into the MBH in the nearest massive halo due to the repositioning algorithm, once again adding unphysical mergers.

Previously, this was dealt with in Kelley et al. (2017a) and Blecha et al. (2016) using a cut based on mass – only $M_\bullet > 10^6\ \mathrm{M_\odot}$ are kept – to exclude low-mass MBHs that are overwhelmingly the MBHs involved in these numerical issues. As the authors state, this cut had minimal effect on their predictions for the PTA background, which is dominated by high-mass MBH binaries $M \gtrsim 10^8\ \mathrm{M_\odot}$). In a study about LISA, these near-seed mass MBHs play a very important role. Therefore, we designed a post-processing method to avoid removing these small MBHs in the most robust way possible.

To handle these issues we start by requiring that all merger constituent MBHs must exist for at least one snapshot prior to the merger, which is always true for MBHs above $10^6\ \mathrm{M_\odot}$. This removes the MBHs seeded when the FOF finder identifies an overdensity and seeds an unphysical MBH. This is effective in removing these MBHs because the time between the seeding and the merging of the unphysical MBH is less than the duration of one snapshot. We then focus on identifying galaxies that have had their central MBHs removed by the ambiguities related to the FOF finder and the repositioning algorithm. We track the evolution of the galaxy devoid of an MBH as it continues on after its fly-by encounter where it lost its MBH. If this galaxy seeds a new MBH before it merges with another galaxy, we remove this MBH from our catalogue, as well as any of its subsequent mergers.

In addition to filtering the secondary mergers, we analysed the effects of the premature mergers created by the FOF association of two separate haloes. In this process, the removal of the MBH from its galaxy occurs earlier in cosmic time than the actual galactic mergers. This causes the MBH merger to occur at slightly earlier cosmic

times. This effect would increase our rate predictions: mergers at earlier cosmic times inflate the number of mergers because the volume of the observer's past light-cone is larger at higher redshifts. However, we believe this effect to be small for two reasons: the binaries involved in this scenario tend to have largely unequal mass ratios, making them increasingly difficult to detect with LISA; and the delay between the unphysical MBH merger and the subsequent galaxy merger is of the order of 10^8 yr, which represents a small fraction of the lifetime of the Universe as well as a small fraction of the average binary inspiral times predicted from our two binary inspiral models.

We will now discuss the post-processing performed on the merger host galaxies as well as the cuts made based on those galaxies.

2.2 Host galaxy information

To process the host galaxy data, we follow the process of Blecha et al. (2016) and Kelley et al. (2017a). We need density profiles and velocity dispersions of remnant host galaxies for input into our evolution time-scale models. We use the profiles at resolvable scales to extrapolate inward to the centres of the galaxies to infer properties at unresolved scales. Therefore, we confirm each remnant host galaxy is sufficiently resolved for these calculations by requiring the galaxy contain at least 80 DM particles, 80 gas cells, and 300 star particles.

From these remaining galaxies, we construct spherically averaged, radial density profiles for stars, gas, and DM. We calculate these profiles based on the innermost shells of particles (cells for the gas) surrounding the galactic centre. We assume the profile represented by the innermost shells of particles/cells extends inwards to the core of the galaxy. We require at least four particles/cells in each radial bin. The profile is then formed from the innermost 8 bins satisfying the four particle/cell minimum requirement. Binaries were excluded from the final catalogues when fits could not be constructed under these requirements. A graphical representation of this process can be seen in fig. 1 of Kelley et al. (2017a). For our evolution prescription, we constrain the density profile index to be in between 0.5 and 2.5. The distribution function for the models tested becomes unphysical below an index of 0.5. The upper end of 2.5 is determined based on observed stellar cusps of giant elliptical galaxies. After all of our cuts, we are left with 17 535 of the original 23 708 mergers. We compare this number to 9270, which is the amount of mergers remaining after the cuts applied in Kelley et al. (2017a). Therefore, we analyse 8265 more mergers in this work. The mergers that remain form our final merger catalogue. Fig. 1 compares the main binary parameters resulting from our extraction method to the flat mass cut of $10^6\,\mathrm{M}_\odot$. Fig. 2 shows the main properties of mergers in this final catalogue.

3 METHODS

3.1 Binary lifetime models

For the following models, our goal is to calculate the evolutionary time-scale from ∼kpc scales to coalescence, which is usually of order ∼Gyr. We use this time-scale to find the coalescence time by adding the evolution time to the time of binary formation (particle merger in Illustris).

The first model we examine has no evolution of the binaries from their particle mergers in the Illustris simulation. We refer to this model as 'ND' for no delays. In other words, we consider the formation time to equal the coalescence time. This model is our baseline model against which we compare our more detailed models for binary MBH coalescence time-scales: it represents the exact prediction from the simulation if the merger process is not modelled below ∼kpc scales, which is a common assumption in rate prediction papers.

In addition to our ND model, we will examine a subset of our 'no delays' model requiring masses to be greater than $10^6\,\mathrm{M}_\odot$. This allows us to test the difference in our extraction process by comparing the new data set to the old extraction data set similar to the one used in Kelley et al. (2017a). We will refer to this model as 'ND-6.'

In the following, we will describe detailed models that have been constructed and analysed in previous papers. Therefore, we give a quick overview of each model. For more information on the DA17 model (Section 3.1.1), see Dosopoulou & Antonini (2017). For the K17 model (Section 3.1.2), see Kelley et al. (2017a,b).

3.1.1 DA17 model

The equations shown below are taken directly from Dosopoulou & Antonini (2017). In what follows, primary, M, (secondary, m) will refer to the larger (smaller) MBH. For this model, we are assuming the mergers are gas-poor.

The initial time-scale in this model is the large-scale orbital decay from ∼kpc scales to a separation equal to the influence radius, $r_{\rm infl}$, of the primary MBH. $r_{\rm infl}$ is a shorter length scale than the resolution in the Illustris simulation. Therefore, we use an approximation from Merritt, Schnittman & Komossa (2009) given by,

$$r_{\rm infl} = 10.8 \left(\frac{M}{10^8\,\mathrm{M}_\odot}\right)\left(\frac{\sigma}{200\ \mathrm{km\ s^{-1}}}\right)^{-2}\ \mathrm{pc}, \quad (1)$$

where σ is the three-dimensional stellar velocity dispersion of the primary galaxy. Modelling the primary host galaxy as a singular isothermal sphere, Binney & Tremaine (1987) show this decay time-scale is given by,

$$T_\star^{\rm bare} = 17\frac{6.6}{\ln\Lambda}\left(\frac{R_{\rm e}}{10\ \mathrm{kpc}}\right)^2\left(\frac{\sigma}{300\ \mathrm{km\ s^{-1}}}\right)\left(\frac{10^8\,\mathrm{M}_\odot}{m}\right)\mathrm{Gyr}, \quad (2)$$

where $\ln\Lambda$ is the Coulomb logarithm and $R_{\rm e}$ is the effective radius of the primary galaxy. However, Dosopoulou & Antonini (2017) modify this formalism to include the effect of the secondary MBH remaining embedded in a core of stars from the secondary galaxy. This added mass causes the system to sink faster towards the primary galaxy's centre therefore taking less time than is predicted by equation (2). Assuming the mass of stars bound to the secondary to be a constant proportionality of $10^3 m$ (Merritt & Ferrarese 2001), this large-scale decay time-scale becomes,

$$T_{\star,1}^{\rm gx} = 0.06\frac{2}{\ln\Lambda'}\left(\frac{R_{\rm e}}{10\ \mathrm{kpc}}\right)^2\left(\frac{\sigma}{300\ \mathrm{km\ s^{-1}}}\right)\left(\frac{10^8\,\mathrm{M}_\odot}{m}\right)\mathrm{Gyr}, \quad (3)$$

where $\Lambda' = 2^{3/2}\sigma/\sigma_{\rm s}$ with $\sigma_{\rm s}$ representing the stellar velocity dispersion of the secondary galaxy. This equation, however, does not include tidal stripping of stars from the secondary galaxy by the primary galaxy. Including this effect, the large-scale decay time-scale is given by,

$$T_{\star,2}^{\rm gx} = 0.15\frac{2}{\ln\Lambda'}\left(\frac{R_{\rm e}}{10\ \mathrm{kpc}}\right)\left(\frac{\sigma}{300\ \mathrm{km\ s^{-1}}}\right)^2\left(\frac{100\ \mathrm{km\ s^{-1}}}{\sigma_{\rm s}}\right)\mathrm{Gyr}. \quad (4)$$

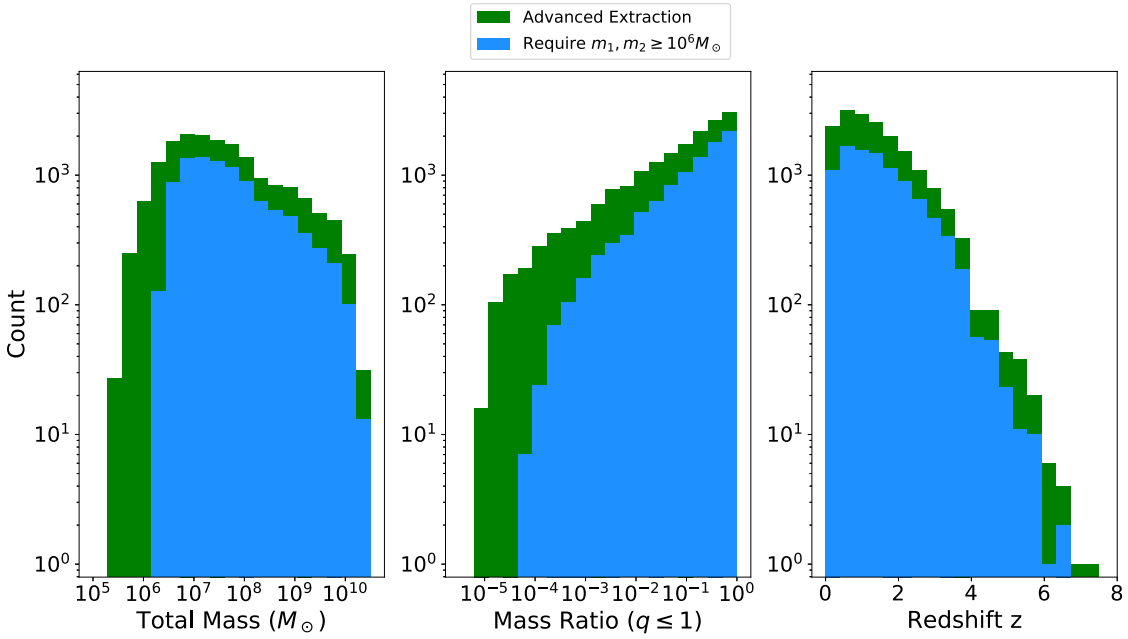

Figure 1. Histograms for the main extraction parameters (M_T, q, and z) are shown here. We compare our new advanced extraction (green) to the extraction used previously in Blecha et al. (2016) and Kelley et al. (2017a) requiring $m_1, m_2 \geq 10^6$ M$_\odot$ (blue). These counts are given after we apply the cuts described in Section 2.

Using equations (3) and (4), we can approximate the large-scale decay time-scale as

$$T_\star = \max\left(T^{\rm gx}_{\star,1}, T^{\rm gx}_{\star,2}\right). \tag{5}$$

During the large-scale decay, the stellar velocity distribution is treated as Maxwellian. Once the secondary MBH reaches $r_{\rm infl}$ of the larger MBH, this assumption no longer holds because the potential is dominated by the central MBH. Therefore, the stellar velocity distribution is treated according to equation 20 in Dosopoulou & Antonini (2017). We refer to this next regime as the 'dynamical friction regime' to match the conventions of the original paper. The time-scale for the binary to decay to a shorter separation $r = \chi r_{\rm infl}$ ($\chi < <1$) is given by,

$$T^{\rm bare}_\bullet = 1.5\times10^7 \frac{[\ln\Lambda\alpha+\beta+\delta]^{-1}}{(3/2-\gamma)(3-\gamma)}\left(\chi^{\gamma-3/2}-1\right) \times\left(\frac{M}{3\times10^9\,{\rm M}_\odot}\right)^{1/2}\left(\frac{m}{10^8\,{\rm M}\odot}\right)^{-1}\left(\frac{r_{\rm infl}}{300\,{\rm pc}}\right)^{3/2}{\rm yr}, \tag{6}$$

where γ is the power-law exponent in the stellar density profile, $\rho(r) = \rho_0(r/r_{\rm infl})^{-\gamma}$; α, β, and δ are calculated from equations 17 to 19 in Dosopoulou & Antonini (2017).[1] If we include the stars bound to the secondary, as in equation (4), this time-scale becomes,

$$T^{\rm gx}_\bullet = 1.2\times10^7 \frac{[\ln\Lambda\alpha+\beta+\delta]^{-1}}{(3-\gamma)^2}\left(\chi^{\gamma-3}-1\right) \times\left(\frac{M}{3\times10^9\,{\rm M}_\odot}\right)\left(\frac{100\,{\rm km\,s^{-1}}}{\sigma_{\rm s}}\right)^3{\rm yr}. \tag{7}$$

[1]In these equations, we assume a circular orbit setting $\xi = 1$ (see the paper for more details). However, the dynamical friction decay time-scale is not greatly affected by the orbital eccentricity.

Similar to equation (5), we find $T_\bullet$ with,

$$T_\bullet = \min\left(T^{\rm bare}_\bullet, T^{\rm gx}_\bullet\right). \tag{8}$$

We use equation (8) to evolve the binary down to the hardening radius, $a_{\rm h}$, given by (Merritt 2013),

$$a_{\rm h} \approx 36\frac{q}{(1+q)^2}\frac{M+m}{3\times10^9\,{\rm M}_\odot}\left(\frac{\sigma}{300\,{\rm km\,s^{-1}}}\right)^{-2}{\rm pc}, \tag{9}$$

where q is the mass ratio ($q \leq 1$). Therefore, we set $\chi = a_{\rm h}/r_{\rm infl}$.

The final phase in the DA17 model is the 'hardening phase', as it includes the effect of gravitational radiation. With dry mergers, remnant galaxies after a merger are expected to be triaxial. In this configuration, an efficient hardening of the binary is exposed to a full and consistently refilling loss-cone (Khan et al. 2011, Vasiliev et al. 2014). From $a_{\rm h}$ until coalescence, including the gravitational wave regime, the time-scale is given by (Vasiliev, Antonini & Merritt 2015),

$$T_{\rm h,GW} \approx 1.2\times10^9\left(\frac{r_{\rm infl}}{300\,{\rm pc}}\right)^{\frac{10+4\psi}{5+\psi}}\left(\frac{M+m}{3\times10^9\,{\rm M}_\odot}\right)^{\frac{-5-3\psi}{5+\psi}} \times\phi^{-\frac{4}{5+\psi}}\left(\frac{4q}{(1+q)^2}\right)^{\frac{3\psi-1}{5+\psi}}{\rm yr}, \tag{10}$$

where $\phi = 0.4$ and $\psi = 0.3$ are triaxial parameters estimated from Monte Carlo simulations in Vasiliev et al. (2015). In equation (10), we left out the eccentricity factor as it is unity because we are assuming circularity. Gravitational radiation takes over at $a_{\rm GW}$, determined from the ratio (Vasiliev et al. 2015),

$$\frac{a_{\rm h}}{a_{\rm GW}} \approx 55\left(\frac{r_{\rm infl}}{30\,{\rm pc}}\right)^{5/10}\left(\frac{M+m}{10^8\,{\rm M}_\odot}\right)^{-5/10}\left(\frac{4q}{(1+q)^2}\right)^{4/5}, \tag{11}$$

where we have once again left out the eccentricity factor as it is equal to unity. For $q \approx 10^{-3}$, $a_{\rm h}$ is less than $a_{\rm GW}$. For these binaries, we use the inspiral time for a circular binary due to only gravitational

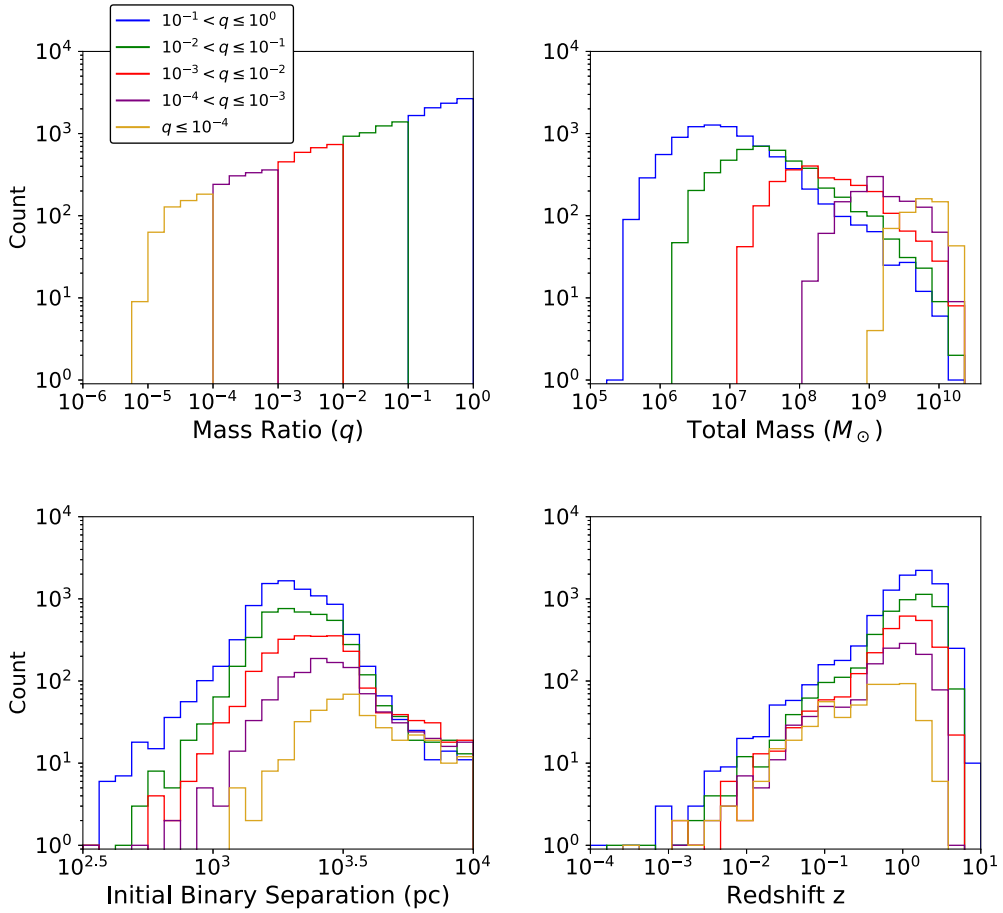

Figure 2. Histograms are shown for the binaries that make up our catalogue after all of our cuts to the MBH binary population. We group the histograms by mass ratio. The initial separation shown represents the upper limit on the MBH binary separation at binary formation. This is determined from the MBH simulation smoothing length when the MBH particles are merged in the Illustris simulation. Similarly, the redshift here is the redshift at binary formation.

radiation according to (Peters & Mathews 1963),

$$T_{\rm GW} = 2.3 \times 10^6 \left(\frac{a_h}{10^{-3}{\rm pc}}\right)\left(\frac{10^8\,{\rm M}_\odot}{M}\right)^2\left(\frac{10^5\,{\rm M}_\odot}{m}\right) \times \left(\frac{1}{1+q}\right)\,{\rm yr}. \quad (12)$$

Therefore, our final time-scale, $T_{\rm final}$ is given by,

$$T_{\rm final} = \begin{cases} T_{\rm h,GW}, & \text{if } q \geq 10^{-3} \\ T_{\rm GW}, & \text{if } q < 10^{-3}. \end{cases} \quad (13)$$

The DA17 model's final coalescence time-scale, $t_{\rm coal}$, is therefore given by

$$t_{\rm coal} = T_\star + T_\bullet + T_{\rm final}. \quad (14)$$

3.1.2 *K17 model*

This section introduces the K17 model, described in detail in Kelley et al. (2017a,b). Binaries are numerically integrated from their formation at large separations until their eventual coalescence. Dynamical friction is implemented following Chandrasekhar (1943), where the deceleration is given as,

$$\left.\frac{{\rm d}v}{{\rm d}t}\right|_{\rm DF} = -\frac{2\pi G^2(M+m_{\rm DF})\rho}{v^2}\ln\Lambda_{\rm c}, \quad (15)$$

where the relative velocity is taken to be the maximum of the orbital velocity and stellar velocity dispersion, i.e. $v = \max(v_{\rm orb}, \sigma)$, ρ is the total mass-density, and the Coulomb logarithm is set to $\ln\Lambda_{\rm c} = 15$. The effective mass of the secondary, $m_{\rm DF}$, assumes that the mass of the secondary host galaxy is stripped over the course of a dynamical time, i.e.

$$m_{\rm DF} = m\left(\frac{m+m_{\rm host}}{m}\right)^{1-t/\tau_{\rm dyn}}. \quad (16)$$

The initial host galaxy mass is measured from Illustris in the snapshot preceding the merger event. Once the binary shrinks below the 'loss-cone radius' (Begelman et al. 1980), the hardening rate is calculated following the stellar-scattering prescription from Sesana, Haardt & Madau (2006),

$$\left.\frac{{\rm d}a}{{\rm d}t}\right|_{\rm SS} = -\frac{G\rho}{\sigma}a^2 H, \quad (17)$$

where a is the semimajor axis of the binary which is being integrated, and H is a dimensionless coefficient calculated from numerical scattering experiments.

The accretion rate calculated in Illustris provides an estimate of the presence of circumbinary gas. To model the energy extraction from this material, we assume the gas settles into a geometrically thin alpha-disc (Shakura & Sunyaev 1973) with different regions corresponding to the dominant components of pressure (radiation

versus thermal) and opacity (Thomson versus free–free), following Shapiro & Teukolsky (1986). The hardening rates in each regime are calculated in Haiman et al. (2009), as a function of disc surface density and binary mass ratio. Numerical simulations have found these analytical prescriptions to be quite accurate over the parameter ranges studied (Tang, MacFadyen & Haiman 2017; Fontecilla, Haiman & Cuadra 2019). We assume these discs extend out to the radius at which they become Toomre unstable. GW energy extraction is implemented at all radii, following Peters & Mathews (1963).

3.2 Determining detectability

3.2.1 Characteristic strain

MBH binaries provide a variety of signals measurable by LISA since their chirp evolution in the frequency domain occurs near the low-frequency band edge of the LISA sensitivity curve. Binaries with $\sim 10^5$–$10^7\,\mathrm{M}_\odot$ total mass will provide a measurable inspiral, merger, and ringdown leading to very loud signals even out to the cosmic horizon (Amaro-Seoane et al. 2017).

The binary inspiral is the initial stage of binary black hole coalescence when the two MBHs orbit one-another at separations greater than the innermost stable circular orbit ($R = 6GM/c^2$). At these separations, the orbit is usually treated with a post-Newtonian formalism.

The merger stage follows the binary inspiral with a highly non-linear relativistic process. This process continues until the MBHs have contacted each other to form a single event horizon, leading to ringdown. The dominant mode of the ringdown spectrum is expected to be the $l = m = 2$ quasi-normal mode. Deviations from general relativity can be measured if LISA can detect subdominant modes in the ringdown spectrum. This process is referred to as the so-called black hole spectroscopy (Berti, Cardoso & Will 2006; Berti et al. 2016; Baibhav et al. 2018; Baibhav & Berti 2019).

We use the characteristic strain, h_c, to model the binary signal which accounts for the time the binary spends in each frequency bin (Finn & Thorne 2000). The characteristic strain is given by (Moore, Cole & Berry 2015),

$$h_\mathrm{c}^2 = 4f^2 \left|\tilde{h}(f)\right|^2, \tag{18}$$

where $\tilde{h}(f)$ represents the Fourier transform of a time domain signal. To find $\tilde{h}(f)$, we use the phenomenological waveform *PhenomD* (Husa et al. 2016; Khan et al. 2016). *PhenomD* is based on fitting analytical templates to numerical relativity waveforms. For a detailed description of its constructions, see Husa et al. (2016) and Khan et al. (2016). Here, we focus on how the waveform is determined based on the parameters of the MBHs in our population.

To generate the waveforms, we use the `gwsnrcalc` PYTHON package from the BOWIE analysis tool (Katz & Larson 2019). `gwsnrcalc` takes as inputs the masses of the MBHs, M, m; the dimensionless spin of each MBH a_1, a_2; the redshift of the binary, z; and the start and end times of the binary's orbit, in relation to the merger of the binary, t_st and t_end.

The dimensionless spin of each MBH is $a_i = J_i/m_i^2$, where J is the magnitude of the spin angular momentum. a ranges from -1.0 (anti-aligned to the orbital angular momentum) to 1.0 (aligned to the orbital angular momentum). For convenient use of *PhenomD*, we treat the spins as aligned. Measurements of MBH spins have shown spins near maximal (Miller 2007; Reynolds 2013). For this reason, we choose to model spins of $a_1 = a_2 = a = 0.8$. As the spin magnitude is raised, the waveform will gain more signal. For near-equal mass systems, which represent a majority of systems in our catalogue, the difference in the spin does not change the signal significantly. For systems of mass ratio farther from unity, the spin can have a significant impact on their detectability because the signal peak can increase by an order of magnitude from the spin-down ($a = -1$) to the spin-up case ($a = 1$). Therefore, applying this spin configuration ($a = 0.8$) represents the optimistic case for these systems. The choice to use the same spin for both MBHs is made because *PhenomD* was calibrated in mostly equal-spin configurations. Within its calibration range, *PhenomD* performs accurately matching waveforms to better than $\sim$1 per cent error. Outside of its calibration range, it produces physically reasonable results, indicating it can be useful for basic studies (Khan et al. 2016). See Fig. 3 for examples of characteristic strain curves.

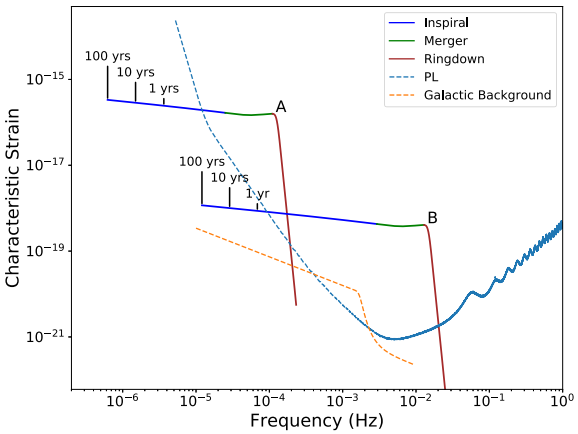

Figure 3. Two examples of the characteristic strain, h_c, curves are shown here with solid lines. The blue, green, and red portions of the binary signals represent the construction we use for the inspiral, merger, and ringdown, respectively. Both examples show $a = 0.8$ and $q = 0.2$ for a signal beginning 100 yr before merger. To plot these curves, we use $t_\mathrm{st} = 100$ yr and $t_\mathrm{end} = 0$ so that we encapsulate 100 yr of inspiral as well as the merger and ringdown. The times before merger are labelled above the strain curve for 100, 10, and 1 yr before merger. Example A shows a binary of $M_\mathrm{T} = 10^8\,\mathrm{M}_\odot$ and $z = 0.75$. Example B shows $M_\mathrm{T} = 5 \times 10^5\,\mathrm{M}_\odot$ and $z = 2$. In addition to binary signals, the sensitivity curve tested in this work (PL) is shown in characteristic strain of the noise, h_N (Amaro-Seoane et al. 2017). Additionally, the Galactic background noise we use is shown with a dashed orange line.

3.2.2 Start and end times

The start times, t_st, and end times, t_end, both represent the time until merger for a specific binary at which LISA begins and ends its observation of the binary's signal. These times are effectively a map to the frequency bounds of the *PhenomD* waveform model. These times will be of the order of years. Since the merger and ringdown time-scale is of the order of minutes to hours for MBHs, we do not include this time-scale in t_st and t_end. The main reason for this construction is it allows us to refrain from assuming a specific observation time for each binary: it allows us to test, within our Monte Carlo sample, binaries that merge at early and late times in the LISA observation window, as well as binaries that merge after the LISA observation window. Early merging binaries will have less time where their inspiral signal can be observed, compared to later mergers. Similarly, LISA will not be able to detect the merger or ringdown for binaries that merge after the LISA observing window, leading to inspiral-only signals (if the inspiral is detectable over

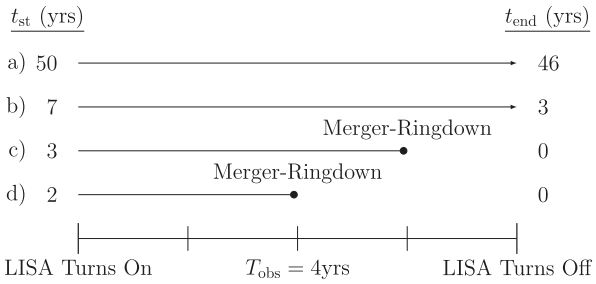

Figure 4. Our construction for t_{st} and t_{end}, given in Section 3.2.2, is illuminated with this diagram. We show two binaries (a, b) for which LISA will only measure the inspiral signal because the binary remains far from merger when LISA is turned off. For these binaries the difference between t_{st} and t_{end} will be T_{obs}. With $T_{obs} = 4$ yr, we will accumulate signal for 4 yr as the binary inspirals towards each other. Cases (c) and (d) represent binaries that merge during LISA observation. For these binaries, t_{end} is always zero; t_{st} determines the duration of time that LISA observes these sources.

the observation time). Fig. 4 displays a diagram showing how t_{st}, t_{end}, and T_{obs} are related for various sources. See Section A1 in the Appendix for our analysis related to inspiral-only signals.

As a point of reference, the merger frequency separating inspiral from merger, is given by

$$f_{mrg} = \frac{c^3}{G}\frac{1}{6^{3/2}\pi M_T\,(1+z)}, \tag{19}$$

where $M_T = M + m$ is the total mass of the binary in its source frame. The $(1 + z)$ term redshifts this mass to the detector frame.

The frequency at a time before merger to 1PN order is given by (Blanchet 2014),

$$f(t) = \frac{c^3}{G}\frac{1}{8\pi M_T(1+z)\tau^{3/8}}\left(1+\left(\frac{11}{32}\eta+\frac{743}{2688}\right)\tau^{-1/4}\right), \tag{20}$$

where

$$\tau = \frac{c^3}{G}\frac{\eta t}{5M_T(1+z)}. \tag{21}$$

t is given in the detector frame. The start frequency of the waveform, f_{st}, is therefore $f(t_{st})$. For signals that exhibit inspiral, merger, and ringdown, t_{end} is zero. In this case, the end frequency, f_{end}, is the highest frequency used in the *PhenomD* model, representing the end of the ringdown, given by (Khan et al. 2016),

$$f_{end} = \frac{G}{c^3}\frac{0.2}{M_T(1+z)}. \tag{22}$$

If t_{end} is not zero, indicating that the source is only detected in the inspiral stage, f_{end} is $f(t_{end})$.

3.2.3 *LISA sensitivity*

The LISA sensitivity configuration used is from the LISA Mission Proposal (Amaro-Seoane et al. 2017). We refer to this sensitivity as 'PL' in Katz & Larson (2019). This sensitivity is based on a 3-arm triangular configuration with 2.5 million km armlengths, 30 cm diameter telescopes, and 2 W end-of-life laser power. In Section A1 of the Appendix, we perform an analysis comparing PL to an older iteration of the LISA configuration (Larson, Hiscock & Hellings 2000) to show how the LISA configuration changes will affect various aspects of LISA MBH analysis.

We show our sensitivity curve in Fig. 3 in terms of the sky-averaged characteristic strain, h_N. The sky-averaging factor is 3/20 (Robson, Cornish & Liu 2019). Sensitivity curves are generally presented in terms of the power spectral density of the noise, S_N. To convert from S_N to h_N, we use $h_N = \sqrt{f S_N}$ (Moore et al. 2015).

We also include the effect of the Galactic background noise in addition to the instrumental noise. We use the analytical approximation of Hiscock et al. (2000) to the Galactic background noise suggested in Bender & Hils (1997). This is shown in Fig. 3. Compared to recent predictions from Robson & Cornish (2017), this is a conservative estimate of this noise contribution. The contribution of this background can be decreased with proper global fitting methods and a longer observation window (Robson & Cornish 2017).

3.2.4 *Signal-to-noise ratio*

We use the SNR to determine the detectability of the sources in our catalogue. The SNR is estimated by integrating the ratio of the signal to noise in the frequency domain. The sky, orientation, and polarization averaged SNR is given by (Robson et al. 2019),

$$\langle\rho^2\rangle = \frac{16}{5}\int_0^\infty \frac{h_c^2}{h_N^2}\frac{1}{f}\mathrm{d}f. \tag{23}$$

The SNR is then multiplied by a factor of $\sqrt{2}$ because we consider a 2 channel interferometer.

An additional question we analyse is how many sources exist in our models with a high enough SNR to perform black hole spectroscopy. To do this, we use the General Likelihood Ratio Test (GLRT) formalism suggested in Berti et al. (2016). Using the GLRT, the SNR of the $l = m = 2$ ringdown mode is used as a proxy for determining the detectability of the $l = m = 3$ or $l = m = 4$ modes. Sources can be spectroscopically measured if $\rho_{l=m=2} > \rho_{GLRT} \equiv \min\left(\rho^{2,3}_{GLRT}, \rho^{2,4}_{GLRT}\right)$, where ρ_{GLRT} for each mode is is given by (Berti et al. 2016),

$$\rho^{2,3}_{GLRT} = 17.687 + \frac{15.4597}{q-1} - \frac{1.65242}{q}, \tag{24}$$

$$\rho^{2,4}_{GLRT} = 37.9181 + \frac{83.5778}{q} + \frac{44.1125}{q^2} + \frac{50.1316}{q^3}. \tag{25}$$

3.3 Monte Carlo analysis

We use a Monte Carlo analysis technique based on Poisson statistics to characterize the range of possibilities resulting from the Illustris output. We do this for multiple reasons. The primary reason is the Illustris output is one iteration of the evolution of a fractional volume within the Universe. We want to understand how the detection rate and source characteristics will vary with the Illustris output as the statistical backdrop. Additionally, we wanted to create a catalogue generator for LISA MBH binary signals, which requires a Monte Carlo draw of a new sample each time. As we will discuss in Section 3.3.2, the Monte Carlo sampling allows us to refrain from assuming an observable duration of the waveform for each of the binaries. Most detection rate predictions assume an observable time for each MBH of 1 yr. This does not account for binaries that will merge before 1 yr of LISA observation. It also does not include the longer measurement of an inspiral signal if the binary signal is observable at times longer than 1 yr before merger. In other words, our method provides a more realistic basis for assessing binary detectability during the LISA mission.

3.3.1 Merger rate prediction

The first parameter in the Monte Carlo sampling process is the coalescence rate of MBH binaries in the Illustris simulation. We calculate this parameter one time for each evolutionary prescription. At this stage of the sampling process, we are not considering detectability; we consider any coalescence that occurs prior to $z = 0$. For the ND and ND-6 models, all binaries in our sample coalesce before $z = 0$ because these two models assume no inspiral time for all MBH binaries. For the binary inspiral models of DA17 and K17, we find that 84 per cent and 66 per cent of binaries coalesced before $z = 0$, respectively. See Section 4.1 and Fig. 6 for more information on how the inspiral models affected the population of binaries coalescing before $z = 0$. We determine the number of coalescences, N, in a given redshift interval $z + \Delta z$. We then compute the number of coalescence events across redshift intervals per comoving volume element,

$$\frac{\mathrm{d}^2 n(z)}{\mathrm{d}z \mathrm{d}V_\mathrm{c}} \approx \frac{N(z)}{\Delta z V_\mathrm{c}}, \tag{26}$$

where V_c, the term on the right-hand side, is the comoving volume of the Illustris simulation, $(106.5\ \mathrm{Mpc})^3$. We can then calculate the number of coalescences per observing time interval given by,

$$\frac{\mathrm{d}N_\mathrm{coal}}{\mathrm{d}t_\mathrm{obs}} = \int_0^\infty \frac{\mathrm{d}^2 \bar{n}(z)}{\mathrm{d}z \mathrm{d}V_\mathrm{c}} \frac{\mathrm{d}z}{\mathrm{d}t} \frac{\mathrm{d}V_\mathrm{c}}{\mathrm{d}z} \frac{\mathrm{d}z}{1+z}, \tag{27}$$

where the $1 + z$ redshifts the infinitesimal time element in $\mathrm{d}z/\mathrm{d}t$ to the observer frame time interval. When we refer to the 'integral rate calculation,' we are referring to equation (27). For our Monte Carlo catalogues, this quantity becomes our input into our Poisson rate calculator.

3.3.2 Poisson sampling

The two parameters needed to perform the desired Poisson sampling is rate of coalescences determined from equation (27) and the duration for which we want to draw potential sources, t_dur. If we only wanted to draw sources for the observation window, we would set $t_\mathrm{dur} = T_\mathrm{obs}$. However, this would only focus on sources coalescing within the observation window. We also want to test for inspiraling sources that would coalesce some time after the LISA observing window. Therefore, we choose $t_\mathrm{dur} > T_\mathrm{obs}$. We tested a variety of values for t_dur. We found $t_\mathrm{dur} = 10^2$ yr encompassed all of the observable systems in our catalogues, while maintaining computational efficiency.

Our final Poisson parameter, λ, is given by,

$$\lambda = \frac{\mathrm{d}N_\mathrm{coal}}{\mathrm{d}t_\mathrm{obs}} t_\mathrm{dur}. \tag{28}$$

In other words, this is the expected number of coalescence events over 10^2 yr. For each catalogue, we draw the number of sources occurring within our 100 yr window from this Poisson distribution.

3.3.3 Event times

When we have the number of sources drawn, we assign each a random coalescence time, t_ev, between zero and 10^2 yr. At the time-scales we are considering ($\sim 10^2$ yr), the distribution of sources over time will not be affected by the evolution of the Universe. $t_\mathrm{ev} = 0$ indicates an event occurring at the moment LISA begins observations. $t_\mathrm{ev} = T_\mathrm{obs}$ represents an event occurring at the moment the LISA observation window ends.

When considering the waveform described in Section 3.2.1, we use t_ev and T_obs to determine the start (t_st) and end (t_end) times related to waveform creation. $t_\mathrm{st} = t_\mathrm{ev}$ because the event time represents the time before merger at the start of LISA observation. Therefore, $t_\mathrm{end} = t_\mathrm{ev} - T_\mathrm{obs}$ if $t_\mathrm{ev} > T_\mathrm{obs}$. If $t_\mathrm{ev} \leq T_\mathrm{obs}$, $t_\mathrm{end} = 0$ (see Fig. 4).

3.3.4 Resampling binary parameters

After sampling the number of binaries and the event times for each event, we need to sample binary parameters of M, m, and z. To do this, we use kernel density estimation methods. However, there is a key distinction that needs to be made for sampling these binary parameters: we must incorporate the volume and time redshifting factors implicit in the expansion of the Universe as weights, w, in the density estimation. If you assume an infinitesimal redshift bin width for equation (27) ($\Delta z \rightarrow \mathrm{d}z$), this weighting factor as a function of redshift is given by,

$$W(z) = \frac{\mathrm{d}z}{\mathrm{d}t}(z) \frac{\mathrm{d}V_\mathrm{c}}{\mathrm{d}z}(z) \frac{1}{1+z}. \tag{29}$$

The weight applied to the ith binary is then $w_i = W_i / \sum_{i=1}^{N} W_i$. We also include the covariance across these parameters in the KDE, so as to sample the population accurately.

4 RESULTS

4.1 Binary lifetimes

We first test and compare our evolutionary prescriptions to understand how the initial population of binaries will change when evolved to coalescence with different sub-grid models. In the ND and ND-6 models, all binaries are considered to be coalesced prior to $z = 0$. By modelling the sub-grid physics as in the DA17 and K17 models, some binaries will no longer merge before $z = 0$ and will therefore deflate the merger rate. Additionally, the mass distributions of the coalesced binaries may change because the prescriptions have different dependencies on the masses. Fig. 5 shows the evolutionary time-scales calculated for all binaries binned by total mass and mass ratio. The DA17 model generally results in a more peaked distribution, while K17 shows a flatter profile across all plots. However, at higher total masses, the two prescriptions become very similar in their predictions. Fig. 6 shows the effect of these evolutionary time-scales on the coalescence fractions of our population. This figure illustrates the global differences, in terms of binary parameters, between the two prescriptions. DA17 favours near-equal mass and low total mass systems, while K17 favours near-equal mass systems with total masses towards the higher end. With the K17 model, larger masses are favoured because they are embedded in higher density, more centrally concentrated stellar cores. On the other hand, the specific dynamical friction prescription used in the DA17 model (see equations 6, 7 and 8) causes high-mass systems to exist for longer times in the dynamical friction stage between the influence radius of the larger MBH (equation 1) and the hardening radius (equation 9) (see the original papers for more details). It is also clear the overall coalescence fractions are higher with DA17 than with K17. The overall coalescence fraction for DA17 was 84 per cent. With the K17 model, only 66 per cent of all binaries coalesced before $z = 0$.

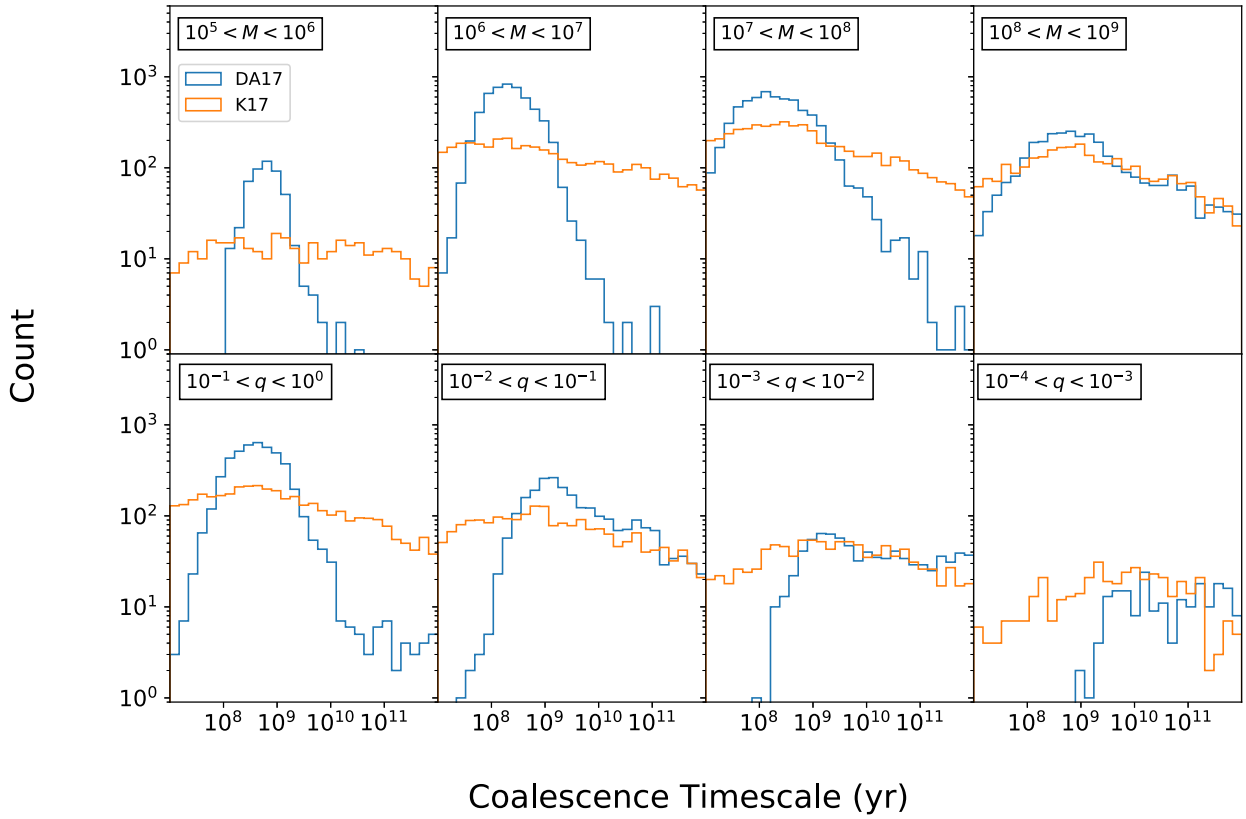

Figure 5. Coalescence time-scales are shown for the DA17 and K17 models in blue and orange, respectively. The top row shows binaries grouped by decades in total mass, M_T. The bottom row shows binaries grouped by decades in mass ratio, q.

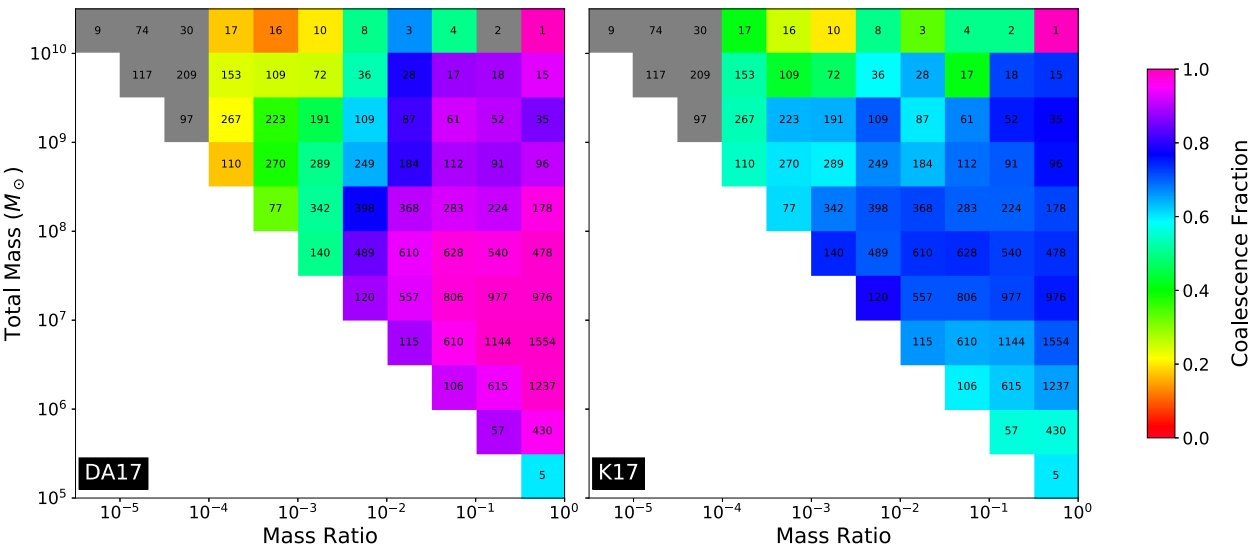

Figure 6. Coalescence fractions are compared for the DA17 (left) and K17 (right) models. These fractions are binned in total mass and mass ratio. The number in each bin represents the total number of binaries residing in that bin. Therefore, this is the same for both models. The colour represents the coalescence fraction based on the colour bar on the right. None of the binaries shown in grey coalesce before $z = 0$. The white space represents binaries not analysed here due to the Illustris resolution limit.

4.2 Rate predictions

We calculate merger rates with two methods: integrating the redshift distributions with equation (27) and Monte Carlo sampling (Section 3.3). The results are similar as expected; however, the Monte Carlo aspect allowed for more freedom in terms of not setting specific values for observation duration and starting times before merger. It also ensures we are examining different realizations of the merging MBH population. Table 1 shows integral merger rate results for each prescription. It also shows detection rates for each stage of binary black hole coalescence as well as the rate of sources where

Table 1. Merger and detection rate calculations per year are shown using the integral calculation (equation 27). The evolution prescriptions are listed in the first column. The top row shows the 'no delays' model: ND. The second row shows the ND-6 model, displayed in italics because it represents a subset of the ND model with $m_1, m_2 \geq 10^6\,M_\odot$. The final two rows show the DA17 (Dosopoulou & Antonini 2017) and K17 (Kelley et al. 2017a,b) models, respectively. The merger rate gives the rate of coalescences without considering LISA detectability. The remaining columns show the detection rate ($\rho > 8$). Additionally, detection rates are separated into signal types: All, Ins, MR, and BHS. 'All' indicates reaching detection threshold using the entire signal. 'Ins' and 'MR' represent detection rates of inspiral signals and signals from the merger and ringdown, respectively. These categories are not independent: a single binary can add to the rate in both categories. 'BHS' is the detection rate of MBH binaries where black hole spectroscopy is possible (see equation 24). We estimate our errors in these predictions to all fall below 0.01 according to equation 17 in Salcido et al. (2016).

Prescription	Merger rate	All	Ins	MR	BHS
ND	0.98	0.75	0.45	0.77	0.74
ND-6	*0.57*	*0.44*	*0.23*	*0.45*	*0.43*
DA17	0.80	0.70	0.42	0.70	0.67
K17	0.55	0.44	0.28	0.45	0.43

spectroscopic measurements (equation 24) are possible. Similar results for the Monte Carlo method are shown in Table A1. The redshift distributions used in these calculations are shown in Fig. 7. For the main integral merger rate and detection rate results, the rate is quoted as per year. For the following integral rate calculations, we estimate our standard deviation in our predictions to be less than 0.01 (based on our chosen redshift bin width), according to equation 17 in Salcido et al. (2016).

First, comparing ND with ND-6 ('ND-6' represents the subset of binaries in 'ND' which have each constituent mass above $10^6\,M_\odot$), we see the advanced extraction was important for LISA-related analysis since the ND merger rate was almost two times the merger rate of ND-6. This holds true for the detection rates as well. An interesting statistic here is the inspiral detection rate, 0.23 yr^{-1} with ND-6 versus 0.45 yr^{-1} with ND, as this demonstrates the importance of the lower mass systems retained by our advanced extraction technique.

Comparing all four models, we see some interesting results. The hierarchy predicted in Section 3.1 is apparent. The DA17 rates resemble much more strongly the ND rates compared to K17 due to the inability of K17 to coalesce the large number of low-mass systems. Interestingly, the K17 rates strongly resemble the ND-6 rates, indicating the loss of the low-mass systems, as well as the general loss in coalescing systems due to adding sub-grid modelling, caused the rate to decrease to similar levels as without the low-mass systems entirely. These aspects can be seen clearly in Fig. 7. After calculating the SNR and making our cut at $\rho = 8$, we see that the DA17 curve tracks the ND curve, while the K17 curve mirrors the ND-6 curve.

When analysing the mass distributions, we find all models produce similar results. Mass distributions from our Monte Carlo analysis can be seen in Fig. 8. The black solid line shows the limit imposed by the Illustris simulation seed mass at $\sim 10^5\,M_\odot$. Binaries with total mass and mass ratio values above this line cannot exist. Similarly, the black dashed line shows the limit imposed by the $10^6\,M_\odot$ cut. This once again highlights the effect of the more advanced extraction. ND, DA17, and K17 exhibit roughly the same structure. Specifically, their mean values and higher order moments about the mean values are within a smallpercentage of each other; however, DA17 has a slightly smaller kurtosis in the mass ratio. This general similarity does indicate the K17 model has a relatively flat effect across the parameter space, where sources are detectable by LISA, suppressing each mass and mass ratio regime in an equivalent manner.

5 DISCUSSION

The merger rates and LISA detection rates of MBH binaries from the Illustris simulation are low compared to the majority of the literature on this subject. Table 2 summarizes predicted rates in the literature, including whether their base population is from a SAM or a hydrodynamic simulation. Additionally, the type of delay prescription employed is mentioned. For this summary, we focused on predictions related to high-mass seeds since this seeding prescription better matches our setup in this paper. Even with our ND model, which represents the rates directly from the simulation without any sub-grid modelling, our predicted rate is low. LISA will be sensitive to binaries $\leq 10^5\,M_\odot$. Therefore, any study that cannot resolve the galaxies and haloes where these low masses evolve (see section D in Klein et al. 2016) will underpredict the total detection rate. Examining Fig. 8, we see LISA primarily observes binaries of $1 \leq q \leq 10^2$. Therefore, if we consider our population to be complete when we can resolve all constituent MBHs at mass ratios up to 10^2, then our study is only complete above $10^7\,M_\odot$ leading to a further underestimation of the overall detection rate. Another way to consider this effect is by understanding how the seeding prescription will affect the base population. Without seeds below $10^5\,M_\odot$, the detection rate for masses above $10^5\,M_\odot$ will also be deflated: seeds below this mass may grow to a mass above $10^5\,M_\odot$ before undergoing a merger. This aspect significantly affects our merger rate prediction as it removes MBHs from our population at the most common and detectable masses in our sample.

The literature references for SAMs shown in Table 2 employ seeds of $\sim 10^4\,M_\odot$. Within the subset of these SAMs predictions, the higher rate predictions occur with no binary inspiral model employed (similar to our ND model). However, early SAMs predictions did analyse various delay prescriptions, but for smaller seeds ($\sim 150\,M_\odot$) (e.g. Volonteri et al. 2003; Sesana et al. 2004; Sesana et al. 2005). Historically, SAMs have predicted rates about an order of magnitude larger than our findings; however, our results are within an order of magnitude of those predicted in Salcido et al. (2016) for the EAGLE simulations (McAlpine et al. 2016), indicating similarity between two hydrodynamic-based populations. While hydrodynamic simulations strive to produce populations from simple physics, they are restricted to a much more limited parameter space than SAMs (discussed further below). This means that rate predictions from SAMs and hydrodynamic-based models fundamentally differ.

One clear difference is the ability of SAMs to explore an arbitrarily large range of masses, while hydrodynamic simulations are resolution limited, typically to above $\sim 10^5\,M_\odot$. If we consider the increased prevalence of MBHs as we move towards smaller masses, we would expect that models with access to binaries below $10^5\,M_\odot$ would greatly inflate the merger rate in a similar fashion to what we have seen when including the lower masses from Illustris through our advanced extraction. Along the same lines, Illustris, and simulations like it, do not access the scales needed to examine

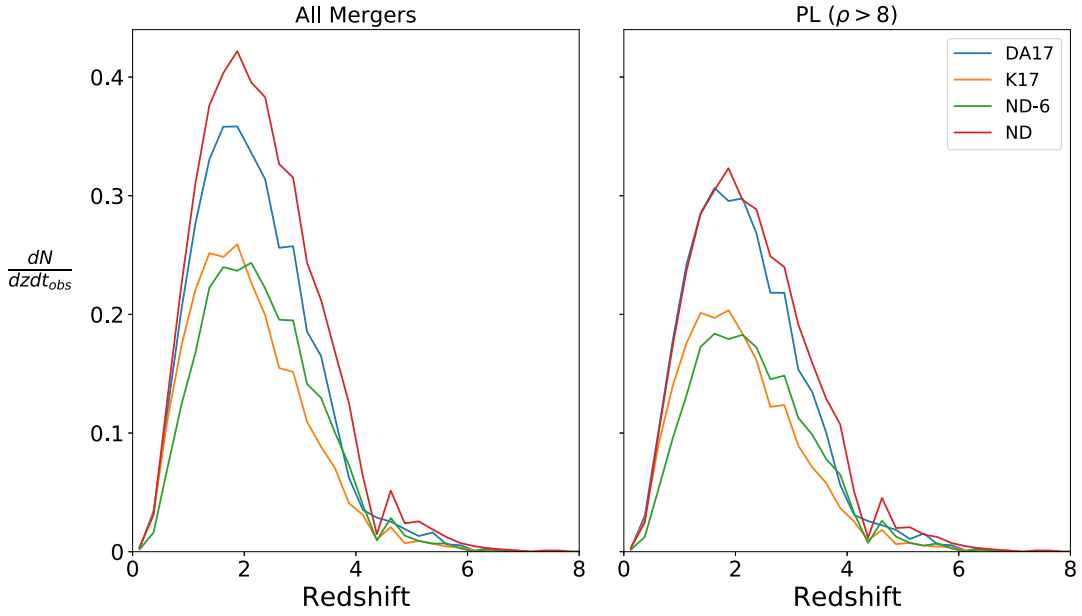

Figure 7. Merger rates per year per unit redshift are shown above. ND (no delays), ND-6 (subset of ND model with $m_1, m_2 \geq 10^6$ M$_\odot$), K17 (Kelley et al. 2017a,b), and DA17 (Dosopoulou & Antonini 2017) models are shown in red, green, orange, and blue, respectively. The left plot shows all mergers from the simulation with each prescription. The right shows the detection rate per year per redshift for each model assuming an SNR cut of $\rho = 8$.

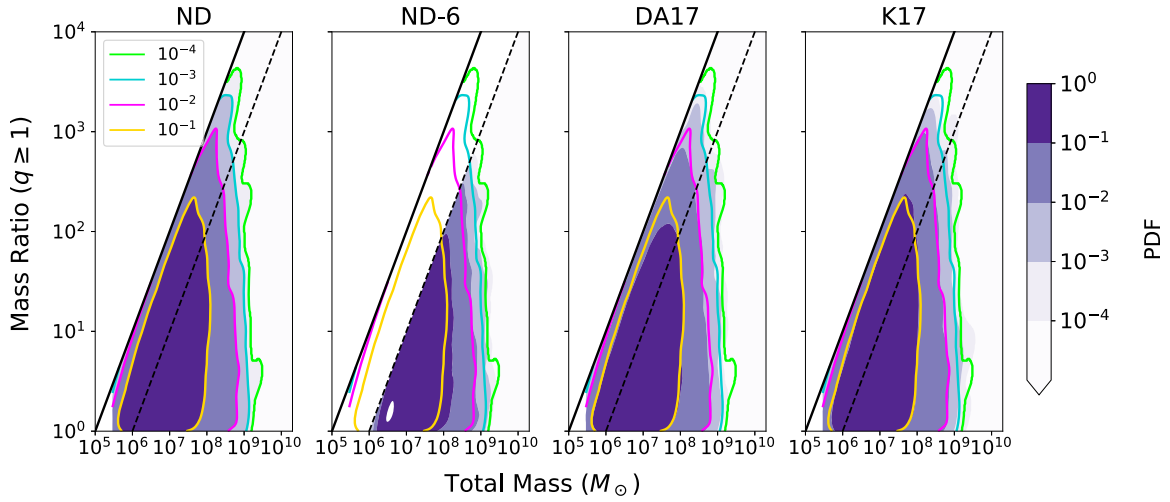

Figure 8. Probability density functions (PDF) are shown for mass ratios and total masses of observed binaries ($\rho \geq 8$) from our 10 000 Monte Carlo catalogues. The coloured, filled contours show the PDF for the model given in the title of each plot. The coloured line contours represent the PDF of the ND model, which we overplot in each panel. The left plot shows the ND ('no delays'). In the centre-left, we compare ND-6 with the ND model. The ND-6 model is a subset of the ND model with $m_1, m_2 \geq 10^6$ M$_\odot$. We then show the DA17 and K17 models as the centre-right and right plots, respectively. When comparing DA17 and K17 to the ND model, there is minimal discernible difference between the three. Empirically, their means and higher order moments in the total mass and mass ratio are within a smallpercentage of one another. Slightly higher mass ratios are more prevalent in the ND and K17 models compared to the DA17 model. The solid and dashed black lines show the effect of a mass cut-off at 10^5 and 10^6 M$_\odot$, respectively.

populations of dwarf galaxies. As evidence mounts that dwarf galaxies house MBHs in their centre in both observations (Reines et al. 2013; Moran et al. 2014; Satyapal et al. 2014; Lemons et al. 2015; Sartori et al. 2015; Pardo et al. 2016; Nguyen et al. 2018, 2019) and simulations (Volonteri, Lodato & Natarajan 2008; van Wassenhove et al. 2010; Bellovary et al. 2019), we must improve models since these dwarf galaxy sources are enitrely missing from our analysis. Due to the prevalence of dwarf galaxies as well as the general understanding that dwarf galaxies consistently merge into larger 'host' galaxies over time, missing the dwarf galaxy MBHs could deflate our rate calculations significantly. This is especially true because these dwarf galaxies will house smaller MBHs that will produce strong signals in the LISA frequency band. Similarly, a number of potential systems may still be missed as we approach the resolution limit of $\sim 10^5$ M$_\odot$.

Similar to this issue with dwarf galaxies, the MBH seeding mechanism plays a large role in rate calculations. The seeding mechanism in Illustris (and EAGLE) is ad hoc. The seeding model chosen for Illustris produces seeds at later times than those seen in other simulations like Tremmel et al. (2018). Additionally,

Table 2. A collection of papers quoting rates for the detection of MBH binaries by LISA is shown above. We focus here on papers analysing the large-seed (10^4–10^6 $M_\odot$) formation channels since this is similar to the seeding mechanism in Illustris. The base population tells if the models are based on SAMs or hydrodynamic simulations. Short descriptions of the MBH binary evolutionary prescriptions are also given. 'None' indicates that no delays between galactic and MBH binary mergers were included. 'Constant' indicates that a constant delay was used for all binaries. 'DF' is a dynamical friction prescription. 'LC' indicates inclusion of a stellar hardening or loss-cone scattering model. A prescription involving torque from a gas disc is indicated with 'VD.' 'GW' indicates a gravitational wave driven inspiral. A prescription containing triple MBH systems is expressed with 'Tri'.

Reference	Base population	MBHB evolution prescription	Merger rate (yr^{-1})
Arun et al. (2009)	SAM	None	∼22
Sesana et al. (2011)	SAM	None	∼25
Klein et al. (2016)	SAM	DF, LC, VD, GW, Tri	∼8
Berti et al. (2016)[a]	SAM	DF, LC, VD, GW, Tri	∼8
Salcido et al. (2016)	Hydrodynamic	Constant[b]	∼2
Bonetti et al. (2019)	SAM	DF, LC, VD, GW, Tri	∼23
This paper	Hydrodynamic	DF, LC, VD, GW	∼0.5–1

[a]For detectable ringdown signals. For the mass ranges we consider in this paper, ringdown signals will be measurable for all detectable sources.
[b]A constant delay was chosen based on gas-heavy versus gas-poor host galaxies.

the Illustris seeding prescriptions will produce seeds at much later times when compared to SAMs. Later seeding means a smaller volume accessible to LISA observations. Therefore, merger rates from SAMs may also be intrinsically higher at higher redshits.

The sub-grid models also affect our rates given that they cause coalescences to occur after $z = 0$. The particular model chosen has the potential to deflate the rate significantly. As previously mentioned, SAM predictions decreased as more 'delay' models were incorporated. In Tremmel et al. (2018), their simulation works to numerically resolve smaller scales in order to track MBH pair formation. They show that the expected number of close MBH binary pairs is lower than expected, which would lead to a depletion in the LISA rate. However, as seen in Bonetti et al. (2019), including triple interactions can remedy this issue quite significantly, producing rates similar to those found when not considering a delay prescription (∼23). Additionally, including prescriptions for gas-driven migration can also increase our rates with our DA17 model (it is included in the K17 model). Without the inclusion of gas-driven migration or triple interactions, our rate is truly a lower limit. However, it can be seen that inclusion of these prescriptions would not significantly alter our results since the rate from our ND model is only slightly larger than those found with the DA17 and K17 prescriptions.

For the EAGLE simulations, Salcido et al. (2016) test seeds of $\sim 10^5$ $M_\odot$, similar to the seed model in Illustris. Therefore, we believe two reasons other than the seed mass are the main factors in our predicted rate difference. First, Salcido et al. (2016) perform a different extraction analysis, compared to the analysis done for this paper, to deal with numerical issues similar to those suggested in Section 2.1. This may inflate the true number of mergers seen in the EAGLE simulations, especially at masses near the seed mass. Secondly, we use more detailed binary inspiral models than the model used in Salcido et al. (2016). In their paper, they choose an inspiral time for all binaries as flat values based on if a galaxy merger was gas-rich (0.1 Gyr) or gas-poor (5 Gyr). When we compare these inspiral times with those from our more detailed models (see Fig. 5), we see that our models will generally predict longer inspiral times, as well as inspiral times that are longer than the age of the Universe. Therefore, the fraction of coalescing binaries before $z = 0$ is larger with their choice of inspiral model, which would lead to a higher predicted detection rate.

If we consider our rates in a relative way, even though our predicted rates are low, the comparison of our binary inspiral models can have an impact on LISA MBH science. We have shown the difference between merger models is relatively small, but still varies by a factor of 2. Since most detections are from low-mass systems (see Fig. 6), much of this factor of two difference will be lost in the low-mass regime. This means the K17 binary inspiral prescription will lead to more difficulty characterizing the low-mass population, as well as seeding models which will predominantly be constrained by this lower mass regime.

6 CONCLUSIONS

We have presented new MBH binary LISA rate calculations based on the Illustris cosmological simulations. Our MBH catalogue is determined from the Illustris output using a new advanced extraction method allowing us to probe masses down to the simulation seed mass of $\sim 10^5$ $M_\odot$. Previous extraction methods made a mass cut requiring $m_1, m_2 \geq 10^6$ $M_\odot$. By strictly following the interaction between MBH mergers and the MBH host galaxies, we were able to refine the analysis and retain an additional 8265 mergers in our sample (∼50 per cent of mergers analysed). This doubled the rate of predicted detections for LISA. Binaries containing an MBH of $\sim 10^5$ $M_\odot$, especially near equal-mass binaries, are prime targets of the LISA mission due to their time spent evolving in the LISA band, their high SNR potential, and their ability to better probe MBH seeding models.

With this MBH merger catalogue from Illustris, we tested four evolutionary prescriptions to understand how binary inspiral models affect our rate predictions, as well as the binary properties of the detected population. Our base model was the ND model, which represented the exact prediction from the Illustris simulation without any delays between the mergers in the simulation (at ∼kpc scales) and true coalescence of the MBH pairs. In order to understand the effect of our new extraction method, we also tested a model without delays while requiring $m_1, m_2 \geq 10^6$ $M_\odot$. We found the detection rate was diminished by a factor of 2 when we included this mass cut.

In addition to our base models without delays, we tested two recent sub-grid models proposed in Dosopoulou & Antonini (2017) (DA17; Section 3.1.1) and Kelley et al. (2017a,b) (K17; Section 3.1.2). When comparing the evolution time-scales predicted by these two models, we find that DA17 produces a smaller spread in orders of magnitude of this time-scale when looking at different masses and mass ratios. K17 produces a relatively flat profile of time-scales when looking at all masses and mass ratios in our catalogue. When examining coalescence fractions in Fig. 6, we find that DA17 favours lower total masses while K17 favours higher total masses. They both favour more equal mass ratio binaries. Additionally, it was clear from our analysis that DA17 predicts that more binaries from our catalogue will merge prior to $z = 0$ (84 per cent) compared to K17 (66 per cent).

Due to the ability to coalesce lower mass binaries, the DA17 model resembled the ND model in terms of detection rate at $\sim$0.7 yr^{-1} (integral rate calculation). The K17 model, with its lower overall coalescence fraction and inability to retain the low-mass binaries, led to a detection rate prediction similar to the ND-6 model at $\sim$0.4 yr^{-1}. The rates predicted for black hole spectroscopy were similar in magnitude to the overall detection rates. These rates represent lower limits for similar rate predictions (see Section 5).

We also examined the probability density functions of the total masses and mass ratios of the detected binaries using Monte Carlo generated catalogues. We found that all models with low-mass support (ND, DA17, and K17) produced similar detectable populations in terms of these parameters. This indicates that the surpression of sources by K17 is effectively equivalent across mass regimes, leading to a lower detection rate, while maintaining a similar total mass and mass ratio probability density function to the DA17 and ND models. The ND-6 model cannot match this because it has an entirely inaccessible region where at least one constituent MBH of less than 10^6 M$_\odot$ is required. Overall, we show that these two detailed models for the evolution of MBH binaries from Dosopoulou & Antonini (2017) and Kelley et al. (2017a), Kelley et al. (2017b) lead to differences in the detection rate and the observable population by LISA.

ACKNOWLEDGEMENTS

We would like to thank Marta Volonteri and the referee for helpful suggestions to improve our paper. MLK acknowledges support from the National Science Foundation under grant DGE-0948017. This research was supported in part through the computational resources and staff contributions provided for the Quest/Grail high performance computing facility at Northwestern University. FD acknowledges support from PCTS and Lyman Spitzer Jr fellowships. LB acknowledges support from National Science Foundation grant AST-1715413. ASTROPY, a community-developed core PYTHON package for Astronomy, was used in this research (Astropy Collaboration 2013). This paper also employed use of SCIPY (Jones et al. 2001), NUMPY (Walt, Colbert & Varoquaux 2011), and MATPLOTLIB (Hunter 2007).

REFERENCES

Abbott B. P. et al., 2017, Nature, 551, 85
Alvarez M. A., Wise J. H., Abel T., 2009, ApJ, 701, L133
Amaro-Seoane P., Eichhorn C., Porter E. K., Spurzem R., 2010, MNRAS, 401, 2268
Amaro-Seoane P. et al., 2017, preprint (arXiv:1702.00786)
Ardaneh K., Luo Y., Shlosman I., Nagamine K., Wise J. H., Begelman M. C., 2018, MNRAS, 479, 2277
Arun K. G. et al., 2009, Class. Quantum Gravity, 26, 94027
Astropy Collaboration, 2013, A&A, 558, A33
Baibhav V., Berti E., 2019, Phys. Rev. D, 99, 024005
Baibhav V., Berti E., Cardoso V., Khanna G., 2018, Phys. Rev. D, 97, 44048
Barausse E., Bellovary J., Berti E., Holley-Bockelmann K. Farris B., Sathyaprakash B., Sesana A., 2015, J. Phys.: Conf. Ser., 610, 12001
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nature, 287, 307
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bellovary J. M., Cleary C. E., Munshi F., Tremmel M., Christensen C. R., Brooks A., Quinn T. R., 2019, MNRAS, 482, 2913
Bender P. L., Hils D., 1997, Class. Quantum Gravity, 14, 1439
Berti E., Cardoso V., Will C. M., 2006, Phys. Rev. D, 73, 64030
Berti E., Sesana A., Barausse E., Cardoso V., Belczynski K., 2016, Phys. Rev. Lett., 117, 101102
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, NJ
Blanchet L., 2014, Living Rev. Relativ., 17, 2
Blecha L. et al., 2016, MNRAS, 456, 961
Boehle A. et al., 2016, ApJ, 830, 17
Bogdanović T., 2015, in Sopuerta C. F., ed., Astrophysics and Space Science Proceedings, Vol. 40, Gravitational Wave Astrophysics. Springer, Berlin, p. 103
Bonetti M., Sesana A., Haardt F., Barausse E., Colpi M., 2019, MNRAS, 486, 4044
Burke-Spolaor S., 2013, Class. Quantum Gravity, 30, 224013
Chandrasekhar S., 1943, ApJ, 97, 255
Charisi M., Bartos I., Haiman Z., Price-Whelan A. M., Graham M. J., Bellm E. C., Laher R. R., Márka S., 2016, MNRAS, 463, 2145
Davies M. B., Miller M. C., Bellovary J. M., 2011, ApJ, 740, L42
Devecchi B., Volonteri M., 2009, ApJ, 694, 302
Dosopoulou F., Antonini F., 2017, ApJ, 840, 31
Dunn G., Bellovary J., Holley-Bockelmann K., Christensen C., Quinn T., 2018, ApJ, 861, 39
eLISA Consortium, 2013, preprint (arXiv:1305.5720)
Fan X. et al., 2001a, AJ, 121, 54
Fan X. et al., 2001b, AJ, 122, 2833
Fan X. et al., 2006, AJ, 131, 1203
Finn L. S., Thorne K. S., 2000, Phys. Rev. D, 62, 124021
Fontecilla C., Haiman Z., Cuadra J., 2019, MNRAS, 482, 4383
Frank J., Rees M. J., 1976, MNRAS, 176, 633
Fryer C. L., Woosley S. E., Heger A., 2001, ApJ, 550, 372
Gair J. R., Vallisneri M., Larson S. L., Baker J. G., 2013, Living Rev. Relativ., 16, 7
Genel S. et al., 2014, MNRAS, 445, 175
Graham M. J. et al., 2015, MNRAS, 453, 1562
Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529
Haiman Z., Abel T., Rees M. J., 2000, ApJ, 534, 11
Haiman Z., Kocsis B., Menou K., 2009, ApJ, 700, 1952
Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288
Hinshaw G. et al., 2013, ApJS, 208, 19
Hiscock W. A., Larson S. L., Routzahn J. R., Kulick B., 2000, ApJ, 540, L5
Holz D. E., Hughes S. A., 2005, ApJ, 629, 15
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Husa S., Khan S., Hannam M., Pürrer M., Ohme F., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 44006
Jones E., Oliphant T., Peterson P. et al., 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Katz M. L., Larson S. L., 2019, MNRAS, 483, 3108
Katz H., Sijacki D., Haehnelt M. G., 2015, MNRAS, 451, 2352
Kelley L. Z., Blecha L., Hernquist L., 2017a, MNRAS, 464, 3131
Kelley L. Z., Blecha L., Hernquist L., Sesana A., Taylor S. R., 2017b, MNRAS, 471, 4508
Kelley L. Z., Blecha L., Hernquist L., Sesana A., Taylor S. R., 2018, MNRAS, 477, 964

Khan F. M., Just A., Merritt D., 2011, ApJ, 732, 89
Khan S., Husa S., Hannam M., Ohme F., Pürrer M., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 44007
Klein A. et al., 2016, Phys. Rev. D, 93, 24003
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Larson S. L., Hiscock W. A., Hellings R. W., 2000, Phys. Rev. D, 62, 62001
Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J. C., 2013, MNRAS, 436, 2989
Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, ApJ, 670, 249
Lemons S. M., Reines A. E., Plotkin R. M., Gallo E., Greene J. E., 2015, ApJ, 805, 12
Lightman A. P., Shapiro S. L., 1977, ApJ, 211, 244
Lin D. et al., 2018, Nat. Astron., 2, 656
Liu T. et al., 2016, ApJ, 833, 6
Loeb A., Rasio F. A., 1994, ApJ, 432, 52
McAlpine S. et al., 2016, Astron. Comput., 15, 72
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
Magorrian J. et al., 1998, AJ, 115, 2285
Merritt D., 2013, Dynamics and Evolution of Galactic Nuclei. Princeton Univ. Press, Princeton, NJ
Merritt D., Ferrarese L., 2001, MNRAS, 320, L30
Merritt D., Milosavljević M., 2005, Living Rev. Relativ., 8, 8
Merritt D., Schnittman J. D., Komossa S., 2009, ApJ, 699, 1690
Miller J. M., 2007, ARA&A, 45, 441
Mirza M. A., Tahir A., Khan F. M., Holley-Bockelmann H., Baig A. M., Berczik P., Chishtie F., 2017, MNRAS, 470, 940
Moody M. S. L., Shi J.-M., Stone J. M., 2019, ApJ, 875, 66
Moore C. J., Cole R. H., Berry C. P. L., 2015, Class. Quantum Gravity, 32, 15014
Moran E. C., Shahinyan K., Sugarman H. R., Vélez D. O., Eracleous M., 2014, AJ, 148, 136
Mortlock D. J. et al., 2011, Nature, 474, 616
Muñoz D. J., Miranda R., Lai D., 2019, ApJ, 871, 84
Nelson D. et al., 2015, Astron. Comput., 13, 12
Nguyen D. D. et al., 2018, ApJ, 858, 118
Nguyen D. D. et al., 2019, ApJ, 872, 104
Omukai K., Schneider R., Haiman Z., 2008, ApJ, 686, 801
Pardo K. et al., 2016, ApJ, 831, 203
Peters P. C., Mathews J., 1963, Phys. Rev., 131, 435
Petiteau A., Babak S., Sesana A., 2011, ApJ, 732, 82
Plowman J. E., Jacobs D. C., Hellings R. W., Larson S. L., Tsuruta S., 2010, MNRAS, 401, 2706
Plowman J. E., Hellings R. W., Tsuruta S., 2011, MNRAS, 415, 333
Porter E. K., Sesana A., 2010, preprint (arXiv:1005.5296)
Quinlan G. D., 1996, New Astron., 1, 255
Quinlan G. D., Hernquist L., 1997, New Astron., 2, 533
Rasskazov A., Merritt D., 2017, ApJ, 837, 135
Reines A. E., Greene J. E., Geha M., 2013, ApJ, 775, 116
Reynolds C. S., 2013, Class. Quantum Gravity, 30, 244004
Robson T., Cornish N., 2017, Class. Quantum Gravity, 34, 244002
Robson T., Cornish N. J., Liu C., 2019, Class. Quantum Gravity, 36, 105011
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Roedig C., Dotti M., Sesana A., Cuadra J., Colpi M., 2011, MNRAS, 415, 3033
Salcido J., Bower R. G., Theuns T., McAlpine S., Schaller M., Crain R. A., Schaye J., Regan J., 2016, MNRAS, 463, 870
Sartori L. F., Schawinski K., Treister E., Trakhtenbrot B., Koss M., Shirazi M., Oh K., 2015, MNRAS, 454, 3722
Satyapal S., Secrest N. J., McAlpine W., Ellison S. L., Fischer J., Rosenberg J. L., 2014, ApJ, 784, 113
Schulze A., Wisotzki L., 2011, A&A, 535, A87
Schutz B. F., 1986, Nature, 323, 310
Sesana A., 2010, ApJ, 719, 851
Sesana A., Haardt F., Madau P., Volonteri M., 2004, ApJ, 611, 623
Sesana A., Haardt F., Madau P., Volonteri M., 2005, ApJ, 623, 23
Sesana A., Haardt F., Madau P., 2006, ApJ, 651, 392
Sesana A., Gair J., Berti E., Volonteri M., 2011, Phys. Rev. D, 83, 44036
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Shankar F. et al., 2016, MNRAS, 460, 3119
Shapiro S. L., Teukolsky S. A., 1986, Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects, Wiley, New York
Shen Y., Greene J. E., Strauss M. A., Richards G. T., Schneider D. P., 2008, ApJ, 680, 169
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Soltan A., 1982, MNRAS, 200, 115
Springel V., 2010, MNRAS, 401, 791
Tanaka T., Haiman Z., 2009, ApJ, 696, 1798
Tang Y., MacFadyen A., Haiman Z., 2017, MNRAS, 469, 4258
Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985
Tremmel M., Governato F., Volonteri M., Quinn T. R., Pontzen A., 2018, MNRAS, 475, 4967
van Wassenhove S., Volonteri M., Walker M. G., Gair J. R., 2010, MNRAS, 408, 1139
Vasiliev E., Merritt D., 2013, ApJ, 774, 87
Vasiliev E., Antonini F., Merritt D., 2014, ApJ, 785, 163
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., 2014a, MNRAS, 444, 1518
Vogelsberger M. et al., 2014b, Nature, 509, 177
Volonteri M., Madau P., Haardt F., 2003, ApJ, 593, 661
Volonteri M., Lodato G., Natarajan P., 2008, MNRAS, 383, 1079
Walt S. v. d., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Wang H.-T. et al., 2019, Phys. Rev. D, 100, 043003
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Yu Q., 2002, MNRAS, 331, 935

APPENDIX A: EFFECT OF LISA CONFIGURATION ON RATE PREDICTIONS

Here, we analyse the difference between the classic LISA sensitivity (Larson et al. 2000) and the proposed LISA sensitivity (Amaro-Seoane et al. 2017). We do this purely to draw comparisons and understand how the measurement of MBH binaries has changed with the change in mission design. For this section, we refer to the Proposal sensitivity curve (Amaro-Seoane et al. 2017) as 'PL'. The classic LISA curve has an unrealistic low-frequency behaviour following an f^{-2} power law to infinitely low frequencies. To correct for this, we copy the low-frequency band edge behaviour of PL, move it to lower strains, and spline it together with the classic LISA curve. We refer to this curve as 'CLLF'. See Katz & Larson (2019) for more information about this construction. Both PL and CLLF are shown in Fig. A1. The basic difference between these two curves is better low-frequency performance exhibited by CLLF due to a longer armlength (5 million km).

The CLLF LISA configuration allows for observation of sources earlier in their inspiral as well as stronger overall measurements of this signal in the low-frequency regime. Additionally, the low-frequency difference can allow for detection of larger total mass sources (Katz & Larson 2019). The difference between these two curves can be seen visually in Fig. A1 as the space between the two curves on the low-frequency end. For this paper, we study MBHs of masses greater than $\sim 10^5\,M_\odot$. This mass regime radiates gravitational waves observable at frequencies below 10^{-3} Hz. Therefore, we want to focus our sensitivity analysis on the low-frequency band edge. Katz & Larson (2019) show that the high-mass range observable by LISA is strongly dependent on the low-frequency band edge behaviour and can reach masses of $\sim 10^9\,M_\odot$.

We also tested the ability of each detector configuration to observe partial signals as well as binaries earlier in their evolution during

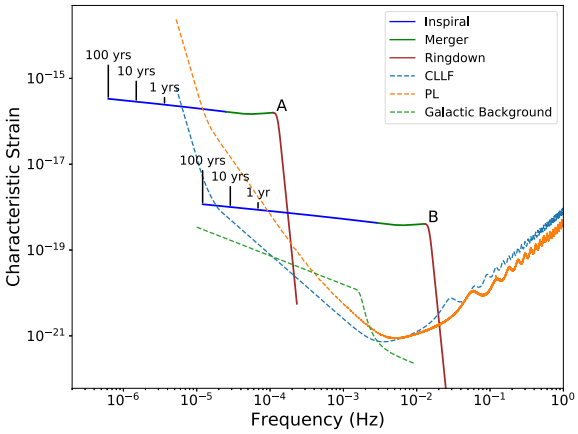

Figure A1. This figure is the same as Fig. 3, but with the addition of the modified classic LISA curve (see below). Two examples of the characteristic strain, h_c, curves are shown here with solid lines. The blue, green, and red portions of the binary signals represent the construction we use for the inspiral, merger, and ringdown, respectively. Both examples show $a = 0.8$ and $q = 0.2$ for a signal beginning 100 yr before merger. To plot these curves, we use $t_{st} = 100$ yr and $t_{end} = 0$ so that we encapsulate 100 yr of inspiral as well as the merger and ringdown. The times before merger are labelled above the strain curve for 100, 10, and 1 yr before merger. Example A shows a binary of $M_T = 10^8$ M$_\odot$ and $z = 0.75$. Example B shows $M_T = 5 \times 10^5$ M$_\odot$ and $z = 2$. In addition to binary signals, the two sensitivity curves tested in this section are shown in characteristic strain of the noise, h_N. PL (dashed orange) is the curve proposed in Amaro-Seoane et al. (2017). CLLF (dashed blue) is a modified version of the classic LISA curve (Larson et al. 2000). At low frequencies, the classic LISA curve has an unphysical constant slope. To correct for this, we move the PL low-frequency behaviour to lower strains and spline it together with the original classic LISA curve. Additionally, the Galactic background noise we use is shown with a dashed green line.

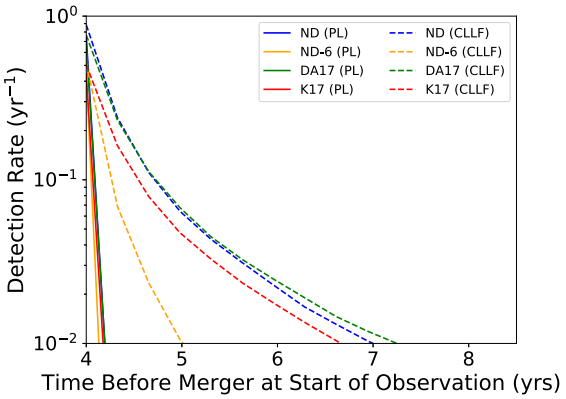

Figure A2. The integral detection rate calculation for different values of t_{st} is shown above. t_{st} is set to the value on the horizontal axis. T_{obs} is set to 4 yr. Each binary inspiral model is then tested with both LISA configurations. The 'no delay' (ND) model is shown in blue. ND-6, a subset of the ND model with $m_1, m_2 \geq 10^6$ M$_\odot$, is shown in orange. DA17 (Dosopoulou & Antonini 2017) and K17 (Kelley et al. 2017a,b) models are shown in green and red, respectively. The PL (CLLF) LISA configuration is shown with solid (dashed) lines.

only their inspiral phase. Fig. A2 shows the detection rate of binaries versus their time-to-merger at the start of LISA observation. The binaries shown in this figure will not reach their coalescence by the time LISA observing terminates. Therefore, the rates shown represent detections from only the inspiral portion of coalescence. Immediately to the right of the observation time, the rate drops off significantly with the loss of the merger and ringdown signals. Additionally, the inspiral signal is not observable for larger binaries, meaning LISA will not detect any of the larger mass systems if they do not merge in the observing window. At longer times before merger, the rate is further decreased as signals from lower mass systems fall below the noise. Therefore, in order for these detections to occur, they have to be lower in mass, sufficiently close in luminosity distance, and close enough in time-to-merger to be detectable. Here, we see a stark difference between the two detector configurations. The PL configuration almost drops off entirely to the right of the observation time. CLLF displays a different behaviour: it shows a more gradual decrease in the detection rate with increasing start time. Detections at start times between 4 and 7 yr occur at rates between $\sim 10^{-2}$ and 1 yr^{-1}. Therefore, while our magnitude of the overall rates is low, the detection rate of these inspiral-only sources can enhance the overall detection rate by $\sim$20 per cent if the detector's low-frequency performance is closer to the classic LISA configuration. If lower mass binaries were included, this percentage would increase because at masses lower than those tested here, we enter a regime where the inspiral stage can be observed, but the merger and ringdown are no longer detectable.

All of our delay models have the same general behaviour for both LISA configurations. Since ND-6 bottoms out at 10^6 M$_\odot$, the difference between ND-6 and the other models can be see as the low-mass ($< 10^6$ M$_\odot$) contribution added by the advanced extraction. Consequently, ND-6 does have a much steeper drop for CLLF because it does not have the low-mass support to boost detections of inpiral-only sources. Once again, the advanced extraction helps to establish a more complete LISA analysis. Similarly, due to the inability of K17 to maintain the low-mass systems, the K17 curve is below the DA17 and ND model curves at all start times.

A1 Monte Carlo rate results

Table A1 shows our Monte Carlo results for 10 000 sampled catalogues. We estimate the errors in our Monte Carlo results to be approximately $\sqrt{1/N} = \sqrt{1/10\,000} = 1$ per cent. The rates for the measurement of the entire signal roughly match our integral calculations (Table A2). Testing the inspiral-only results was useful in this setting because we can get a rate without assuming specific start times. We see that PL measures a negligble rate of inspiral-only sources for all of our time-scale models. CLLF, on the other hand, shows that its low-frequency performance constitutes $\sim$20 per cent of all detections for models that include the lower mass binaries (9 per cent of detections for ND-6). These systems had the lowest masses in our catalogues.

Table A1. Monte Carlo results for the observed detection rate per year from our 10 000 catalogues are shown above (see Section 3.3). We focused our Monte Carlo calculation on the overall signal detection rate as well as sources only detected in their inspiral stage. The coalescence time-scale prescriptions are listed in the first column: ND is the 'no delay' model; ND-6 is a subset of the ND model with all constituent masses below $10^6\,M_\odot$ eliminated from consideration; DA17 is the binary inspiral model from Dosopoulou & Antonini (2017); and K17 is the inspiral model from Kelley et al. (2017a, 2017b). The full signal detection rate with PL and CLLF are shown in the second and third columns, respectivley. Similary, inspiral-only detection rates are shown for PL and CLLF in the fourth and fifth columns, respectively. The last two columns show the fraction of the total detection rate contributed by inspiral-only signals using PL and CLLF. For our Monte Carlo results, the error in the predictions is approximately $\sqrt{1/N} = \sqrt{1/10\,000} = 1$ per cent.

Prescription	PL All signal	CLLF All signal	PL Ins only	CLLF Ins only	PL Fraction of Ins only	CLLF Fraction of Ins only
ND	0.79	1.08	<0.01	0.19	<1 per cent	18 per cent
ND-6	*0.47*	*0.57*	*<0.01*	*0.05*	*<1 per cent*	*9 per cent*
DA17	0.72	0.97	<0.01	0.20	<1 per cent	21 per cent
K17	0.47	0.64	<0.01	0.13	<1 per cent	20 per cent

Table A2. This table is the same as Table 1, with rates added for the CLLF LISA configuration. Merger and detection rate calculations per year are shown using the integral calculation (equation 27). The evolution prescriptions are listed in the first column. The top row shows the 'no delays' model: ND. The second row shows the ND-6 model, displayed in italics because it represents a subset of the ND model with $m_1, m_2 \geq 10^6\,M_\odot$. The final two rows show the DA17 (Dosopoulou & Antonini 2017) and K17 (Kelley et al. 2017a,b) models, respectively. The merger rate gives the rate of coalescences without considering LISA detectability. The remaining columns are labelled with the sensitivity curve used (PL or CLLF) to determine detectability. Additionally, detection rates are separated into signal types: All, Ins, MR, and BHS. 'All' indicates reaching detection threshold using the entire signal. 'Ins' and 'MR' represent detection rates of inspiral signals and signals from the merger and ringdown, respectively. These categories are not independent. A single binary can add to the rate in both categories. 'BHS' is the detection rate of MBH binaries where black hole spectroscopy is possible (see equation 24). We estimate our errors in these predictions to all fall below 0.01 according to equation 17 in Salcido et al. (2016).

Prescription	Merger rate	PL All	CLLF All	PL Ins	CLLF Ins	PL MR	CLLF MR	PL BHS	CLLF BHS
ND	0.98	0.75	0.89	0.45	0.75	0.77	0.89	0.74	0.88
ND-6	*0.57*	*0.44*	*0.51*	*0.23*	*0.43*	*0.45*	*0.52*	*0.43*	*0.51*
DA17	0.80	0.70	0.77	0.42	0.68	0.70	0.77	0.67	0.77
K17	0.55	0.44	0.50	0.28	0.44	0.45	0.51	0.43	0.50

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