

# Unveiling the hosts of parsec-scale massive black hole binaries: morphology and electromagnetic signatures

David Izquierdo-Villalba,<sup>1,2\*</sup> Alberto Sesana,<sup>1,2</sup> and Monica Colpi<sup>1,2</sup>

<sup>1</sup> *Dipartimento di Fisica “G. Occhialini”, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

<sup>2</sup> *INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Parsec-scale massive black hole binaries (MBHBs) are expected to form in hierarchical models of structure formation. Even though different observational strategies have been designed to detect these systems, a theoretical study is a further guide for their search and identification. In this work, we investigate the hosts properties and the electromagnetic signatures of massive black holes gravitationally bound on parsec-scales with primary mass  $> 10^7 M_\odot$ . For that, we construct a full-sky lightcone by the use of the semi-analytical model L-Galaxies in which physically motivated prescriptions for the formation and evolution of MBHBs have been included. Our predictions show that the large majority of the MBHBs are placed either in spiral galaxies with a classical bulge structure or in elliptical galaxies. Besides, the scaling relations followed by MBHBs are indistinguishable from the ones of single massive black holes. We find that the occupation fraction of parsec-scale MBHBs reaches up to  $\sim 50\%$  in galaxies with  $M_{\text{stellar}} > 10^{11} M_\odot$  and drops below 10% for  $M_{\text{stellar}} < 10^{11} M_\odot$ . Our model anticipates that the majority of parsec-scale MBHBs are unequal mass systems and lie at  $z \sim 0.5$ , with  $\sim 20$  objects per  $\text{deg}^2$  in the sky. However, most of these systems are inactive, and only  $1 - 0.1$  objects per  $\text{deg}^2$  have an electromagnetic counterpart with a bolometric luminosity in excess of  $10^{43}$  erg/s. Very luminous phases of parsec-scale MBHBs are more common at  $z > 1$  but the number of binaries per  $\text{deg}^2$  is  $\lesssim 0.01$  at  $L_{\text{bol}} > 10^{45}$  erg/s.

**Key words:** black hole physics – quasars: supermassive black holes – gravitational waves – black hole binaries

## 1 INTRODUCTION

According to our current hierarchical structure formation paradigm, the assembly of galaxies takes place through mergers with smaller companions and accretion of intergalactic gas funneled into dark matter filaments (White & Rees 1978; White & Frenk 1991; Haehnelt & Rees 1993; Kauffmann et al. 1999). Besides, several observations have demonstrated that massive black holes (MBHs,  $> 10^6 M_\odot$ ) reside at the centres of most of the galaxies (Genzel & Townes 1987; Kormendy 1988; Dressler & Richstone 1988; Kormendy & Richstone 1992; Genzel et al. 1994; O’Dowd et al. 2002; Häring & Rix 2004; Peterson et al. 2004; Vestergaard & Peterson 2006; Kormendy & Ho 2013; Savorgnan et al. 2016). By joining these two pieces of information together, following the frequent galaxy mergers, it is expected that massive black hole binaries (MBHBs) have formed and coalesced throughout cosmic history.

The evolution of MBHBs have been explored in several theoretical studies. Our current framework is based on the pioneering work of Begelman et al. (1980) which argued that after the galaxy merger, the two MBHs are dragged towards the center of the remnant galaxy thanks to the dynamical friction caused by the

background stars and gas (Chandrasekhar 1943). Once the two black holes reach a separation of parsec (pc) scales, they become a gravitational bound system. During this phase, known as hardening phase, interaction with single stars (Quinlan & Hernquist 1997; Sesana et al. 2006; Vasiliev et al. 2014; Sesana & Khan 2015) or torques extracted from a circumbinary gaseous disc (Escala et al. 2004, 2005; Dotti et al. 2007; Cuadra et al. 2009; Franchini et al. 2021) are capable of reducing the binary separation down to sub-pc scales. At these separations, the emission of gravitational waves (GWs) becomes the main mechanism extracting angular momentum from the system, eventually driving the two black holes to coalescence in a few Myrs. During the whole inspiral phase, the MBHB is a loud GW source, whose specific frequency  $< \text{Hz}$  strongly depends on the separation between the two black holes and their masses (see Colpi 2014, for a review). Current and future experiments aim at detecting the GWs generated by these systems. While the lowest frequencies ( $10^{-9} - 10^{-7}$  Hz) are generated by MBHBs above  $10^8 M_\odot$  and are the main target of Pulsar Timing Array experiments (PTA, Sesana et al. 2004, 2008, 2009; Hobbs et al. 2010; Desvignes et al. 2016; Arzoumanian et al. 2015; Reardon et al. 2016; Bailes et al. 2016; Perera et al. 2019; Susobhanan et al. 2021), the frequency range  $\sim 10^{-5} \text{ Hz} - 0.1 \text{ Hz}$ , covered by inspiralling MBHs of  $10^4 - 10^7 M_\odot$ , will be probed by space based interferometers such as ESA’s Laser Interferometer Space Antenna (LISA, Amaro-Seoane et al. 2017) and the Chinese

\* E-mail: david.izquierdovillalba@unimib.it

concepts TianQin and Taiji (Luo et al. 2016; Hu & Wu 2017; Ruan et al. 2018).

Among these GW experiments, PTAs recently reported the detection of a clear common red signal in their data (Arzoumanian et al. 2020; Goncharov et al. 2021; Chen et al. 2021). Although its origin is still unknown, it might be the signature of the long-sought-after GW astrophysical background from inspiraling MBHBs with orbital periods of  $\sim 0.1 - 10$  years and masses above  $10^8 M_\odot$  located at  $z \lesssim 1$  (Sesana et al. 2009, 2012; Kelley et al. 2017; Izquierdo-Villalba et al. 2022). Albeit indirectly, this will inform on the presence of galaxies hosting parsec-scale MBHBs on their way to merge. PTA observations might also detect individual sources if sufficiently loud and nearby to emerge from the background (Rosado et al. 2015; Kelley et al. 2018; Goldstein et al. 2019). Characterizing the properties of the galaxies hosting these parsec-scale MBHBs is compelling and theoretical investigations have been carried on by a number of authors (see e.g. Tanaka et al. 2012; Kelley et al. 2017). In particular, Martínez Palafox et al. (2014), using N-body simulations and semi-empirical approaches to populate galaxies with MBHBs, found that binary systems with  $M_{\text{BH}} > 10^8 M_\odot$  are hosted in halos with masses  $> 10^{12} M_\odot$  and have typically mass ratios  $q > 0.1$ . Similar conclusions were reached by Tanaka et al. (2012) and Rosado & Sesana (2014) which reported that the most massive binaries emitting GW at  $\sim \text{nHz}$  frequencies should be hosted in dense environments and massive galaxies. The possibility of using signatures of recent galaxy interactions as a plausible indicator of the MBHB presence was recently studied in hydrodynamical simulations by Volonteri et al. (2020). The results pointed out that large delays between the galaxy merger and the final MBHB coalescence is expected to wash out distinctive morphological signatures indicating ongoing galaxy interaction.

Besides galaxy properties, direct electromagnetic signatures resulting from gas accretion onto the two holes can be used as a powerful tool to detect the presence of MBHBs (see the recent reviews of De Rosa et al. 2019; Bogdanovic et al. 2021). In fact, given that these systems form in galaxy mergers, strong gas inflows towards the center of the remnant galaxy are expected (Di Matteo et al. 2005; Springel 2005; Hopkins et al. 2009). During such process part of the gas can reach the vicinity of the MBHBs, causing the formation of a circumbinary accretion disc around the MBHB, inducing observational signatures in the form of electromagnetic signals (Escala et al. 2005; Cuadra et al. 2009). To shed light on the properties of these signatures, several work have been done. By performing a simulation of the prototypical circumbinary disc around binaries with  $M_{\text{BH}} > 10^8 M_\odot$ , Sesana et al. (2012) studied the possible counterparts in the X-ray domain. Interestingly, the authors found that double broad  $K\alpha$  iron lines could be a good indicator for the existence of massive binaries. A similar study was performed by Tanaka & Haiman (2009) but using an analytical approach. The authors reported that active binaries with  $M_{\text{BH}} > 10^8 M_\odot$  might be unusually low X-ray sources when compared to single active MBHBs with the same total mass. Another alternative way to detect the presence of a MBHBs is through a variability analysis. As shown in the theoretical work of, e.g., Kelley et al. (2019), the binary period can imprint variations in the luminosity emitted by the MBHBs. The authors reported that the Doppler boosting due to large orbital velocities or hydrodynamic variability caused by the interaction between accretion discs and the MBHBs could lead to a distinctive variability signature. In quantitative terms, Kelley et al. (2019) predicted that hundreds of

MBHBs could be detected by photometric variability studies with 5 years of Large Synoptic Survey Telescope observations (Ivezić et al. 2019) (see also Charisi et al. 2022). On the observational side, while dual active galactic nuclei (AGN) on kpc scales have been observed at different wavelengths (Rodríguez et al. 2006; Koss et al. 2012; Comerford et al. 2012), up to date there are no confirmed AGNs associated to gravitational bound binaries. This is principally caused by the fact that at the early evolutionary stages of a galaxy merger, the two MBHBs can still be spatially resolved. However, when the two MBHBs reach the galactic center and get gravitational bound, their angular separation in the sky ( $< 0.01 \text{ arcsec}$ ) is beyond the resolution capabilities of current instrumentation. To overcome this limitation, searching techniques rely on indirect measurements such as periodic lightcurves of AGNs. While hundred of these periodic systems have been suggested to be the manifestation of MBHBs (see e.g. Valtonen et al. 2008; Graham et al. 2015; Charisi et al. 2016; Liu et al. 2016, 2019; Liao et al. 2021; Witt et al. 2021), it is still unclear what is the number of false positives due to AGN stochastic variability (see e.g. Vaughan et al. 2016). Besides periodicity signatures, the presence of Doppler-shifting of AGN broad emission lines due to the binary's orbital motion has been proposed and searched for a distinctive feature of sub-parsec MBHBs (see e.g. Bogdanović et al. 2009; Tsalmantza et al. 2011; Montuori et al. 2011; Eracleous et al. 2012; Shen et al. 2013). However, one of the main limitations of these studies concerns the difficulty of tracing with high accuracy the variability pattern of the broad lines, fundamental to study the existence of orbital motion related to parsec-scale MBHBs (Eracleous et al. 2012).

The purpose of this paper is characterize the morphological properties of the galaxies hosting  $> 10^7 M_\odot$  parsec-scale MBHBs and their electromagnetic signatures in further depth, by using a simulated lightcone detailed enough to account for the cosmological evolution of galaxies and their embedded single and dual MBHBs, over a wide range of masses and binary separations. With the mock universe, we aim at providing first quantitative estimates on the occupation fraction of parsec-scale MBHBs, the number of these objects per  $\text{deg}^2$ , and how many of them can be seen in an active phase. For this purpose, we make use of the L-Galaxies semi-analytical model (SAM, Henriques et al. 2015) applied on the Millennium merger trees (Springel 2005). The version of the SAM used in this work is the one presented in Izquierdo-Villalba et al. (2022) which includes physically-motivated prescriptions to treat the physical processes driving MBHBs. Indeed, the model has been tested to be consistent with the PTA limits on the GW background. On top of this, the version of Izquierdo-Villalba et al. (2022) has been modified to include the procedure presented in Izquierdo-Villalba et al. (2019b) which let us transform the simulated boxes provided by L-Galaxies into a lightcone.

The paper is organised as follows: In Section 2 we describe the main characteristics of the Millennium simulation and L-Galaxies, and the procedure used to create a lightcone where the physics of galaxies, MBHBs and the dynamics of MBHBs have been taken into account. In Section 3 we explore two different physically motivated models for tracking the mass growth of MBHBs and MBHBs. In Section 4 and 5 we present the properties of the galaxies hosting MBHBs and explore the electromagnetic counterparts of these binaries, respectively. Finally, in Section 6 we summarize our main findings. A Lambda Cold Dark Matter ( $\Lambda\text{CDM}$ ) cosmology with parameters  $\Omega_m = 0.315$ ,  $\Omega_\Lambda = 0.685$ ,  $\Omega_b = 0.045$ ,  $\sigma_8 = 0.9$  and

$H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$  is adopted throughout the paper (Planck Collaboration et al. 2014).

## 2 A LIGHTCONE FOR THE STUDY OF MASSIVE BLACK HOLE BINARIES

In this section we present the galaxy formation model used to generate a lightcone tailored for the study of pasrec-scale MBHBs. Specifically, we make use of the L-Galaxies semi-analytical model, in the version of Henriques et al. (2015) with the modifications presented in Izquierdo-Villalba et al. (2019a, 2020, 2022). These changes were included to improve the galaxy morphology, extend the physics of MBHs and introduce the formation and evolution of MBHBs. In the following, we summarize the main features of the model, and we refer the reader to the papers cited above for a detailed description of the baryonic physics included.

### 2.1 Generating a galaxy population: Dark matter and baryons

The backbone of L-Galaxies are the subhalo (in the following sections just halo) merger trees of the Millennium (MS, Springel 2005) and Millennium-II (MSII, Boylan-Kolchin et al. 2006) dark matter (DM) only simulations. In this work we use the MS, whose mass resolution allows us to study the evolution of MBHs and MBHBs hosted in galaxies with stellar mass  $> 10^9 M_\odot$ . In brief, the MS follows the cosmological evolution of  $2160^3$  DM particles with mass  $8.6 \times 10^8 M_\odot/h$  within a periodic cube of  $500^3 \text{ Mpc}^3 h^{-3}$  comoving volume. Subhalos were extracted using the SUBFIND algorithm and arranged according to their evolutionary path in the so-called merger trees (Springel et al. 2001; Springel 2005). These trees contain the information of all MS subhalos at 63 different epochs or *snapshots*. To improve the tracing of the baryonic physics, L-Galaxies does an internal time interpolation between two consecutive snapshots with approximately  $\sim 5\text{--}20 \text{ Myr}$  of time resolution depending on redshift. Finally, by using the methodology of Angulo & White (2010), L-Galaxies re-scales the original cosmological parameters of the MS to match the ones provided by the Planck first-year data release (Planck Collaboration et al. 2014).

Regarding the baryonic component, L-Galaxies assumes that galaxies form in the center of the DM subhalos. The birth of a galaxy starts with the formation of a diffuse hot gas component which gradually cools-down and migrates towards the center of the subhalo, forming a disc-like structure. Thanks to the cold gas, the protogalaxy is capable to assemble a stellar disc component through star formation processes which, at the same time, self-regulate the formation of new stars through supernovae explosions. Moreover, the stellar disc can be prone to non-symmetric instabilities (or just disc instabilities) which ultimately lead to the formation of a central ellipsoidal component, called *pseudobulge*. Finally, galaxies interact through mergers, triggering burst of star formation that change the galaxy morphology and promote the formation of a *classical bulge* or *elliptical* structure. As we will discuss in the next section, all these processes have an impact not only on the assembly of galaxies but also on the evolution of MBHs and MBHBs hosted their center.

### 2.2 The population of massive black holes and massive black hole binaries

Each halo is seeded with a MBH of  $10^4 M_\odot$  and random spin in the range  $0 < \chi < 0.998$  (further improvements on the seeding

paradigm will be done by including the model of Spinoso et al. 2022). Once the black hole seed is placed in its host galaxy, it can grow through three different channels: *cold gas accretion*, *hot gas accretion*, and *mergers* with other black holes. Specifically, the first channel is the main driver of the black hole growth and it is triggered by both galaxy mergers and disc instability events. During these processes, the cold gas available for accretion<sup>1</sup> is progressively consumed according to a two phases model, described in Section 3. During any of the MBH growth events, L-Galaxies follows the evolution of the black hole spin modulus in a self-consistent way. During gas accretion events, the model links the number of accretion events that spin-up or spin-down the MBH with the degree of coherent motion in the bulge (following Sesana et al. 2014). In particular, we assume that disc instabilities (mergers) increase (decrease) the coherence of the bulge kinematics. On the other hand, after a MBH coalescence the final spin is determined by the expression of Barausse & Rezzolla (2009).

On top of the evolution of single MBHs, L-Galaxies deals with the formation and evolution of MBHBs (Izquierdo-Villalba et al. 2022). Following the picture of Begelman et al. (1980), we divide the evolution of these systems in three different stages: *pairing*, *hardening* and *gravitational wave* phase. The first one starts after a galaxy merger. During this phase, dynamical friction caused by the stars of the remnant galaxy reduces the initial separation ( $\sim \text{kpc}$ ) between the nuclear and the satellite MBH, sinking the latter towards the galactic center. To determine the time needed for the satellite MBH to reach the nuclear part of the galaxy,  $t_{\text{dyn}}^{\text{BH}}$ , L-Galaxies uses the expression (Binney & Tremaine 2008):

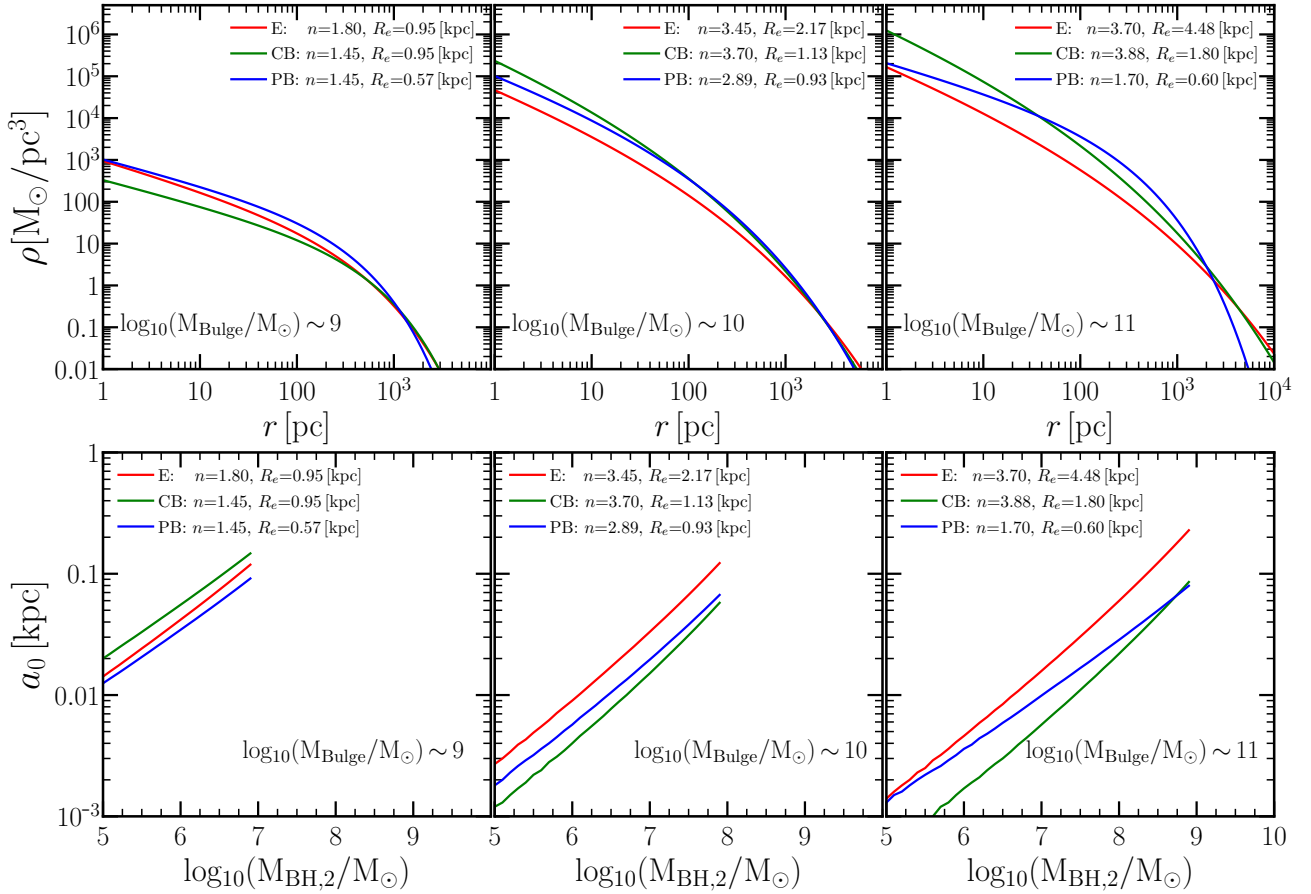
$$t_{\text{dyn}}^{\text{BH}} = 19 f(\epsilon) \left( \frac{r_0}{5 \text{ kpc}} \right)^2 \left( \frac{\sigma}{200 \text{ km/s}} \right) \left( \frac{10^8 M_\odot}{M_{\text{BH}}} \right) \frac{1}{\Lambda} [\text{Gyr}], \quad (1)$$

where  $f(\epsilon)$  is a function which depends on the orbital circularity of the black hole  $\epsilon$  (Colpi et al. 1999),  $r_0$  is the initial position of the black hole deposited by the satellite galaxy after the merger,  $\sigma$  is the velocity dispersion of the remnant galaxy,  $M_{\text{BH}}$  is the mass of the satellite (lighter) black hole and  $\Lambda = \ln(1 + M_{\text{stellar}}/M_{\text{BH}})$  is the Coulomb logarithm (Mo et al. 2010). We refer the reader to Izquierdo-Villalba et al. (2022) for further information about the calculations of these quantities.

Once the satellite MBH reaches the galaxy nucleus, it binds with the central MBH ( $\sim \text{pc}$  separation) and the so-called *hardening phase* begins. From hereafter, we assume that at this moment the two MBHs form a binary. We refer as *primary black hole* (with mass  $M_{\text{BH},1}$ ) the most massive black hole in the system whereas the less massive one is tagged as *secondary black hole* (with mass  $M_{\text{BH},2}$ ). The initial eccentricity of the binary,  $e_0$ , is selected randomly between  $[0 - 1]$  while the initial separation,  $a_0$ , is set to the scale in which  $M_{\text{Bulge}}(< a_0) = 2 M_{\text{BH},2}$ , where  $M_{\text{Bulge}}(< a_0)$  corresponds to the mass in stars of the hosting bulge within  $a_0$ . To determine  $a_0$ , L-Galaxies assumes that the bulge mass density profile,  $\rho(r)$ , follows a Sérsic model (Sérsic 1968):

$$\rho(r) = \rho_0 \left( \frac{r}{R_e} \right)^{-p} e^{-b \left( \frac{r}{R_e} \right)^{1/n}}, \quad (2)$$

<sup>1</sup> The adjustable parameters which control the efficiency of cold gas accretion are similar to the ones presented in Izquierdo-Villalba et al. (2022). These parameters were originally tuned to give the best agreement between the observations and model predictions for the  $z=0$  black hole-bulge correlation.



**Figure 1. Upper panel:** Bulge mass density profile,  $\rho(r)$ . Left, middle and right panels show the results for a bulge mass of  $10^9 M_\odot$ ,  $10^{10} M_\odot$  and  $10^{11} M_\odot$ , respectively. Red, green and blue lines represent the predictions for elliptical galaxies, classical bulges and pseudobulges, respectively. The median structural properties ( $n$ , Sérsic index and  $R_e$  bulge effective radius) of the bulge type in each bin of mass have been extracted from Gadotti (2009). **Lower panel:** The initial separation of the binary at the beginning of the hardening phase,  $a_0$ , as a function of the mass of the secondary black hole,  $M_{\text{BH},2}$ . Line colors represent the same as in the upper panel. Motivated by the bulge-black hole mass correlation, in each panel we have only presented the results for  $M_{\text{BH},2} < 0.01 \times M_{\text{Bulge}}$ .

where  $R_e$  is the bulge effective radius,  $\rho_0$  is the central bulge density and  $n$  its Sérsic index. Finally,  $p$  and  $b$  are two different quantities that depend on the Sérsic index of the bulge:  $p = 1 - 0.6097n^{-1} + 0.05563n^{-2}$  and  $b = 2n - 0.33 + 0.009876n^{-1}$  (Márquez et al. 2000). Following Terzić & Graham (2005) we can transform the density of Eq. 2 into mass and determine  $a_0$  by solving:

$$\gamma\left(n(3-p), b \frac{a_0}{R_e}\right) = \frac{M_{\text{BH},2}}{2\pi \rho_0 R_e^3 n b^{n(p-3)}}, \quad (3)$$

where  $\gamma$  is the incomplete gamma function. Given that Eq. 3 needs to be solved numerically, in Fig. 1 we present the typical values of  $a_0$  and  $\rho(r)$  for three different bulge masses,  $M_{\text{Bulge}}$ . For each mass, we have explored three different parameter combinations to mimic the properties of three different types of galactic bulges: elliptical galaxies, classical bulges, and pseudobulges. Instead of employing arbitrary values, we have used as a reference the median parameters estimated by Gadotti (2009) who explored the structural properties of  $z=0$  bulges by using images of *Sloan Digital Sky Survey*. As we can see, due to the different values of  $n$  and  $R_e$ , pseudobulges, classical bulges and elliptical structures display a different behaviour in  $\rho(r)$  and  $a_0$ . For instance, massive ellipticals display the largest values of  $a_0$  given that they are less centrally concentrated, more extended and have lower densities. Concerning

the typical values of  $a_0$  for  $M_{\text{BH},2} \sim 10^5 - 10^6 M_\odot$  we get  $a_0 \lesssim 30 \text{ pc}$  whereas for  $M_{\text{BH},2} \sim 10^7 - 10^9 M_\odot$  the values can increase up to  $\sim 300 \text{ pc}$ .

Once the hardening phase is started, we assume that the binary system can reduce its semi-major axis,  $a_{\text{BH}}$ , through two different processes. If the system is surrounded by a gas reservoir with a mass larger than the mass of the binary, the interaction with a massive gaseous circumbinary disc is the principal mechanism of shrinking the binary separation (*gas hardening*). In this case, L-Galaxies tracks the evolution of  $a_{\text{BH}}$  following the formalism of Dotti et al. (2015) and Bonetti et al. (2019):

$$\left(\frac{da_{\text{BH}}}{dt}\right)_{\text{Gas}} = -\frac{2\dot{M}_{\text{Bin}}}{\mu} \sqrt{\frac{\delta}{1-e_{\text{BH}}^2}} a_{\text{BH}}, \quad (4)$$

where  $G$  is the gravitational constant,  $c$  the light speed,  $\delta = (1+q)(1+e_{\text{BH}})$ ,  $q = M_{\text{BH},2}/M_{\text{BH},1}$  the mass ratio of the binary,  $\dot{M}_{\text{Bin}}$  the sum of the accretion rate of both MBHs in the system (see Section 3) and  $\mu$  is the reduced mass of the binary. During this process, the eccentricity of the binary,  $e_{\text{BH}}$ , is kept constant to a value of 0.6 (Roedig et al. 2011). On the other hand, if the binary is placed in a gas poor environment the shrinking of  $a_{\text{BH}}$  is caused by three-body interactions with single stars (*stellar hardening*, Quinlan & Hern-



quist 1997). In this case, the values of  $a_{\text{BH}}$  and  $e_{\text{BH}}$  are tracked by integrating numerically the equations presented in Sesana & Khan (2015):

$$\left(\frac{da_{\text{BH}}}{dt}\right)_{\text{Stars}} = -\frac{GH\rho_{\text{inf}}}{\sigma_{\text{inf}}}a_{\text{BH}}^2, \quad (5)$$

$$\left(\frac{de_{\text{BH}}}{dt}\right)_{\text{Stars}} = \frac{G\rho_{\text{inf}}HK}{\sigma_{\text{inf}}}a_{\text{BH}}, \quad (6)$$

where  $H \approx 15 - 20$  and  $K$  are respectively the hardening and eccentricity growth rate extracted from the tabulated values of Sesana et al. (2006). The values of  $\rho_{\text{inf}}$  and  $\sigma_{\text{inf}}$  correspond respectively to the density and velocity dispersion of stars at the MBHB sphere influence, computed according to a Sérsic model.

Once the binary reaches sub-parsec scales, it enters in the *gravitational wave inspiral* phase which drives the system to the final coalescence. In this phase, L-Galaxies follows the evolution of  $a_{\text{BH}}$  and  $e_{\text{BH}}$  by using the model of Sesana & Khan (2015):

$$\left(\frac{da_{\text{BH}}}{dt}\right)_{\text{GW}} = \frac{64G^3(M_{\text{BH}_1} + M_{\text{BH}_2})^3 F(e_{\text{BH}})}{5c^5(1+q)^2 a_{\text{BH}}^3}, \quad (7)$$

$$\left(\frac{de_{\text{BH}}}{dt}\right)_{\text{GW}} = -\frac{304}{15} \frac{G^3 q (M_{\text{BH}_1} + M_{\text{BH}_2})^3}{c^5 (1+q)^2 a_{\text{BH}}^4 (1-e_{\text{BH}}^2)^{5/2}} \left(e_{\text{BH}} + \frac{121}{304} e_{\text{BH}}^3\right), \quad (8)$$

where  $F(e_{\text{BH}})$  is a function which depends on the binary eccentricity (Peters & Mathews 1963):

$$F(e_{\text{BH}}) = (1 - e_{\text{BH}})^{-7/2} \left[1 + \left(\frac{73}{24}\right) e_{\text{BH}}^2 + \left(\frac{37}{96}\right) e_{\text{BH}}^4\right]. \quad (9)$$

Finally, L-Galaxies treats the triple black hole interaction by using the model of Bonetti et al. (2018). In particular, Bonetti et al. (2018) tabulated values are used to select those triple interactions which lead to the merger of a pair of MBHs and those causing the ejection of the lighter MBH from the system. In this latter case the separation of the leftover MBHB is computed following Volonteri et al. (2003) and the resulting  $e_{\text{BH}}$  is select as a random value between  $[0 - 1]$ .

### 2.3 Lightcone construction

One of the main objectives of this work is generating a realistic population of galaxies, MBHs and MBHBs that can be compared directly with observations, i.e taking into account similar selection effects, flux limiting observations, line-of-sight contamination or cosmic variance effects. Therefore, we need to transform the discrete number of comoving boxes provided by Millennium into a lightcone, i.e a mock Universe in which only galaxies, whose light has just enough time to reach the observer, are included. For that, we implemented in our version of L-Galaxies the method presented in Izquierdo-Villalba et al. (2019b). In the following we briefly describe the procedure.

The side-length of the Millennium simulation only allow to construct a lightcone with  $z \sim 0.1$  depth. To extend this limit and reach larger redshifts, we take advantage of the periodic boundary conditions of Millennium and we replicate the box 15 times in each Cartesian coordinate. In this way we can generate a lightcone with a depth of  $z \sim 4$ , regardless of the line-of-sight. Once complete, we

place the observer at the origin of the first replication and we determine the moment when galaxies cross the observer past lightcone. For that, we make use of the galaxy merger trees provided by L-Galaxies which accurately follow in time the evolution of individual galaxies (and their corresponding MBHs and MBHBs) between DM snapshots with a time step resolution  $\sim 5 - 20$  Myr (see Izquierdo-Villalba et al. 2019b for further details). With this procedure we generate a full-sky lightcone with depth  $z \sim 4$  in which all the properties of galaxies, MBHs and MBHBs are stored. In Fig. 2 we present a slice of the full-sky lightcone used in this work, showing that galaxies distributed in filaments and groups. To improve the readability we just plot the galaxies with  $M_{\text{stellar}} > 10^{10} M_{\odot}$  at  $z < 0.15$ . In the figure we highlight the galaxies hosting MBHBs in an active phase, finding that active binaries tend to reside in overdense regions of the Universe (as we will see in the following sections). We also point out three galaxies where we can find MBHBs in different phases of their evolution: pairing phase (top panel), stellar hardening phase (top right panel) and gas hardening phase (lower right panel).

### 3 MODELS FOR THE GROWTH FOR MASSIVE BLACK HOLES AND MASSIVE BLACK HOLE BINARIES

In this work, we explore two different models for cold gas accretion onto the MBHs. The first model (*fiducial model*) assumes that MBHs exhaust their gas reservoir in a *fast* way. On contrary, the second one (*delayed model*) relaxes that assumption allowing longer periods of gas accretion. These two different models allow us to explore how uncertainties in the gas consumption rate affects our predictions. In the following, we describe the two models:

(i) *Fiducial model*: This model has been extensively used in Izquierdo-Villalba et al. (2020) and Izquierdo-Villalba et al. (2022), providing a good match between predicted and observed quasar luminosity functions. As discussed in Section 2, the main channel of MBH growth is the consumption of cold gas starting right after a galaxy merger or a disc instability episode. After these events, the cold gas available for accretion is assumed to settle in a reservoir around the black hole which is progressively exhausted according to a two phases model. The first phase corresponds to an Eddington-limited growth, which lasts until the MBH consumes a fraction  $\mathcal{F}$  of the available gas reservoir.  $\mathcal{F}$  is a free parameter of the model and is set to 0.7 in order to match the faint end of the low- $z$  AGN LFs (see e.g. Marulli et al. 2008; Bonoli et al. 2009). Once this phase ends, the BH enters in a self-regulated or quiescent growth regime characterized by progressively smaller accretion rates:

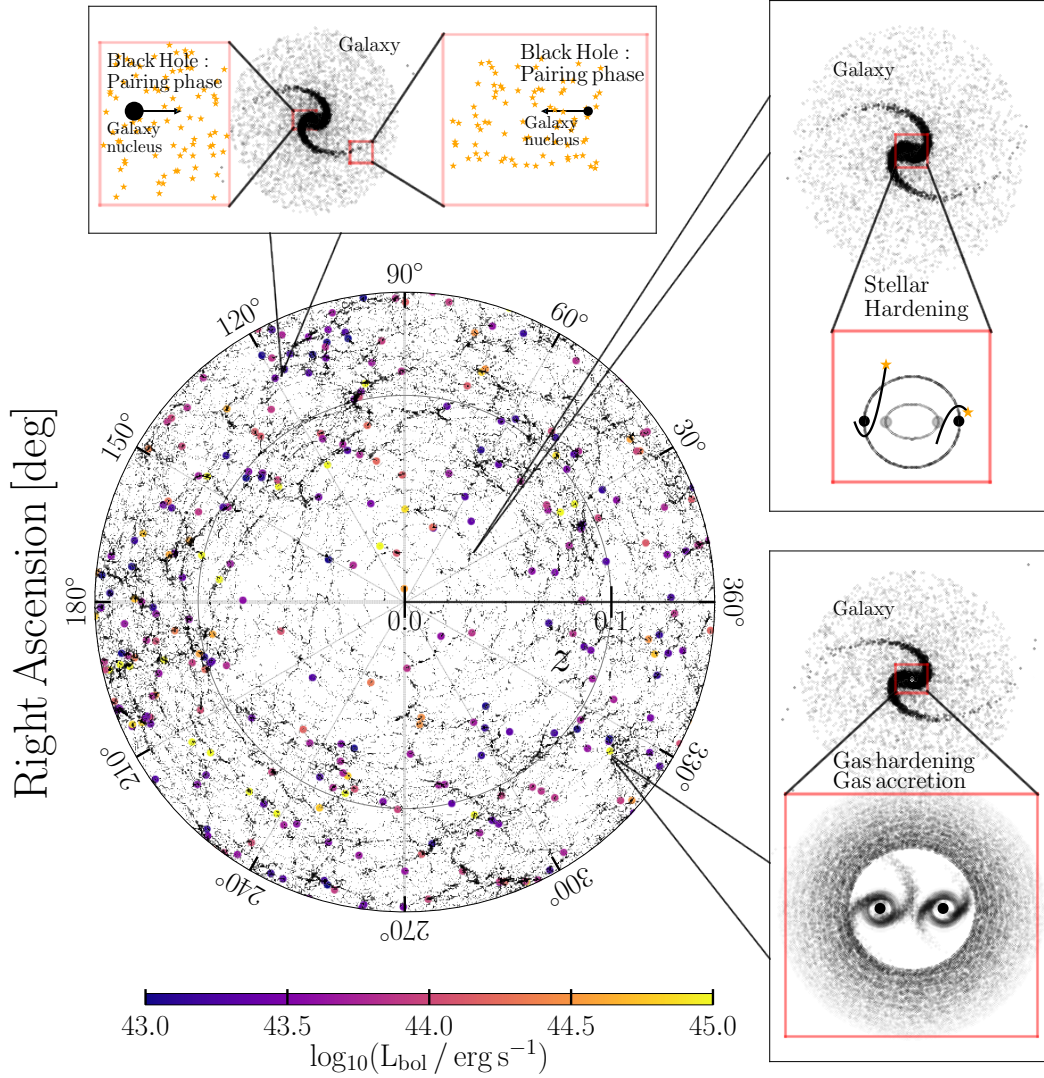
$$f_{\text{Edd}} = \left[1 + (t/t_Q)^{1/2}\right]^{-2/\beta}, \quad (10)$$

where  $f_{\text{Edd}}$  is the accretion rate in Eddington units,  $t$  is the time referred to the starting time of the second phase and  $t_Q = t_0 \xi^\beta / (\beta \ln 10)$ , being  $t_0 = 1.26 \times 10^8$  yr,  $\beta = 0.4$  and  $\xi = 0.3$  (the choice of these values is based on the discussion presented in Hopkins et al. 2006).

Regarding the growth of the binary, we follow Izquierdo-Villalba et al. (2022) which assumes that the accretion rates of the two MBHs of the binary are correlated, as proposed by Duffell et al. (2020):

$$\dot{M}_{\text{BH}_1} = \dot{M}_{\text{BH}_2} (0.1 + 0.9q), \quad (11)$$

where  $q = M_{\text{BH}_2}/M_{\text{BH}_1} \leq 1$  is the binary merger ratio,  $\dot{M}_{\text{BH}_1}$



**Figure 2.** The central plot displays the right ascension-redshift distribution of galaxies at  $z < 0.15$  with  $M_{\text{stellar}} > 10^{10} M_{\odot}$  in a thin angular slice (1 deg) inside our lightcone. The colored points highlights the position of active MBHBs (the color encodes the bolometric luminosity of the binary,  $L_{\text{bol}}$ ). The cartoons in the upper and left panels illustrate the different stages of binary evolution found in our lightcone. The upper panel represents the pairing phase, where the right panels display the stellar (top) and gas (bottom) hardening phase.

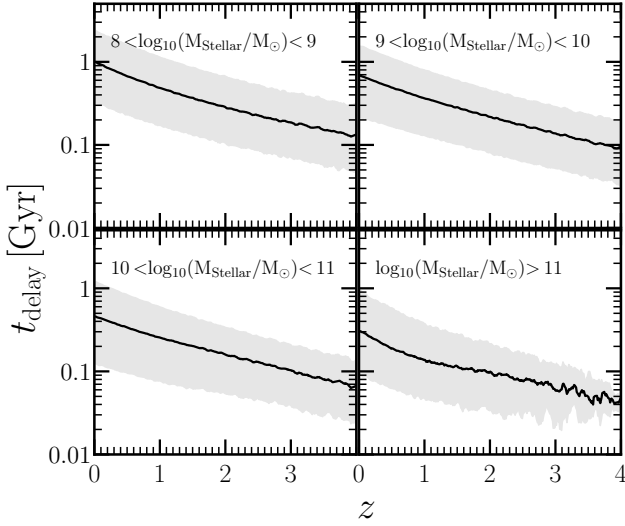
and  $\dot{M}_{\text{BH}_2}$  are respectively the accretion rate of the primary and secondary MBH. As done in Izquierdo-Villalba et al. (2022), we fix the accretion rate of the secondary MBH to Eddington limit.

(ii) *Delayed model:* In this model we relax the fact that the growth of MBHs starts right after the galaxy merger. Instead, motivated by the theoretical work of Hopkins & Quataert (2010), we assume that the gas needs several galaxy dynamical times before reaching the vicinity of the MBH and triggering the AGN activity. Therefore, we consider that the growth of the MBH is delayed with respect to the galaxy merger by a time,  $t_{\text{delay}}$ , given by:

$$t_{\text{delay}} = n t_{\text{dyn}}^{\text{gas}} = n \frac{R_{\text{gas}}}{V_{\text{max}}} \quad (12)$$

where  $R_{\text{gas}}$  is the cold gas disc radius (see Henriques et al. 2015 for further details),  $V_{\text{max}}$  is the maximum circular velocity of the DM halo and  $n$  is a free parameter set to 3. We have checked that

the results presented here do not dramatically change when we double the value of  $n$ . Even though  $n$  might vary with the galaxy merger ratio (see e.g. Hernquist 1989; Naab & Burkert 2001), we prefer to keep the model simple and we assume the same value of  $n$  for all types of mergers. We highlight that we do not impose any  $t_{\text{delay}}$  after a disc instability given that the model assumes that the stars that build-up the bulge and the gas which fuels the MBHs belong to the innermost part of the disc (see Guo et al. 2011 and Henriques et al. 2015 for further details). To give an idea about the typical gas delay included in the model, in Fig. 3 we present the evolution of  $t_{\text{delay}}$  as a function of redshift and stellar mass. As shown, regardless of galaxy mass, at  $z > 2$  the values rarely exceed 0.1 Gyr while at  $z < 2$ ,  $t_{\text{delay}}$  increases up to 1 Gyr. As expected, when dividing the population in different stellar masses we can see that the smaller is the mass of the galaxy the larger is  $t_{\text{delay}}$ . This is principally caused by the small maximum circu-

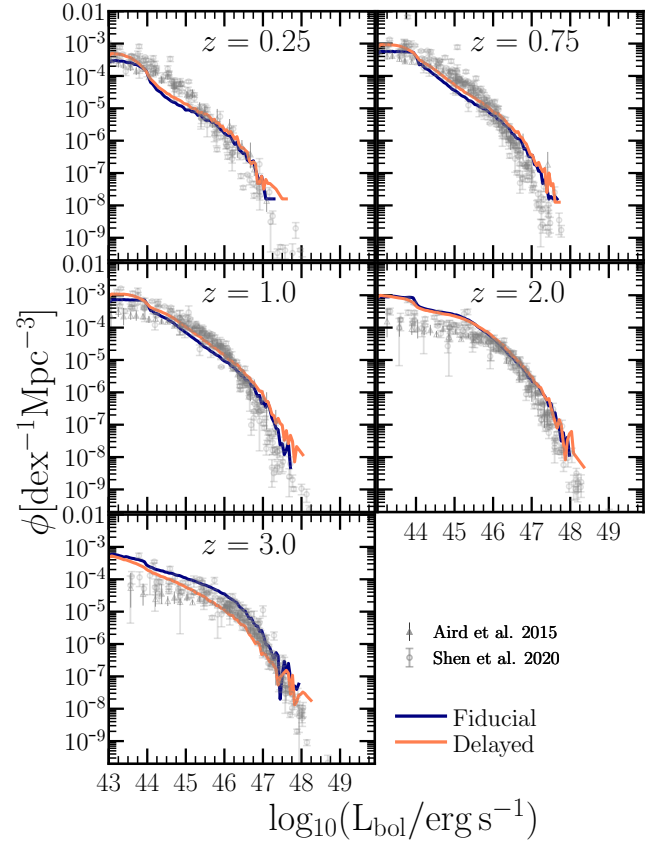


**Figure 3.** Values of  $t_{\text{delay}}$  as a function of redshift for different stellar mass bins. The black lines represent the median value while grey areas display the 32<sup>th</sup> – 68<sup>th</sup> percentile.

lar velocity of the halos in which these low-mass galaxies are hosted.

On top of the gas delay, in this model we relax the first phase of Eddington limited growth included in the *fiducial model*. Specifically, we assume that during that phase,  $f_{\text{Edd}}$  follows a log-normal distribution. Observational studies have suggested that at  $z < 3$ , MBH follow a log-normal distribution centered at 0.3 with a scatter of 0.25 dex (Kollmeier et al. 2006). However, when including these parameters in L-Galaxies, we did not find a good match with the observed luminosity functions at  $z > 2$ . The model under-predicted the number of objects with  $L_{\text{bol}} > 10^{46}$  erg/s. To avoid this, we decide to treat the median and scatter values as free parameters. We found that the model was still able to match the bright end of the  $z > 2$  luminosity functions when a median and scatter of 0.8 and 0.25 dex<sup>2</sup> were used. Finally, this procedure applies also in the presence of a MBHB: we allow a delay between the merger and the gas accretion and we assume that the accretion rate of the secondary MBH follow the same log-normal distribution as single nuclear MBHs.

To show that both *fiducial* and *delayed* models are able to reproduce the observed number density of active black holes, in Fig. 4 we present their predictions for the bolometric luminosity function at  $z < 3$ . Both models are in agreement with the current constraints provided by Aird et al. (2015) and Shen et al. (2020). At  $z > 2$ , the main differences between the two models are seen in the bright end of the luminosity function ( $L_{\text{bol}} > 10^{45}$  erg/s). Specifically, the *delayed model* displays a lower amplitude. This is caused by its smaller  $f_{\text{Edd}}$  values, which causes fainter objects at fix black hole mass and delays the assembly of MBHs with respect to an Eddington limited growth. On the other hand, at lower redshifts both models converge at  $L_{\text{bol}} < 10^{45}$  erg/s but they differ at lower luminosity, with the *delayed model* resulting in slightly larger values. This is caused by the lower accretion rates allowed in the *delayed model* which im-



**Figure 4.** Quasar bolometric luminosity functions ( $L_{\text{bol}}$ ) at  $z \approx 0.25, 0.75, 1.0, 2.0, 3.0$ . Luminosity functions are compared with the data of Shen et al. (2020) (circles) and Aird et al. (2015) (triangles). In all the plots the blue and orange lines correspond to the predictions for the *fiducial* and *delayed* model, respectively.

ply a slower consumption of the gas reservoir, thus resulting in an extended lifetime of the active phase.

#### 4 THE ENVIRONMENTS OF MASSIVE BLACK HOLE BINARIES

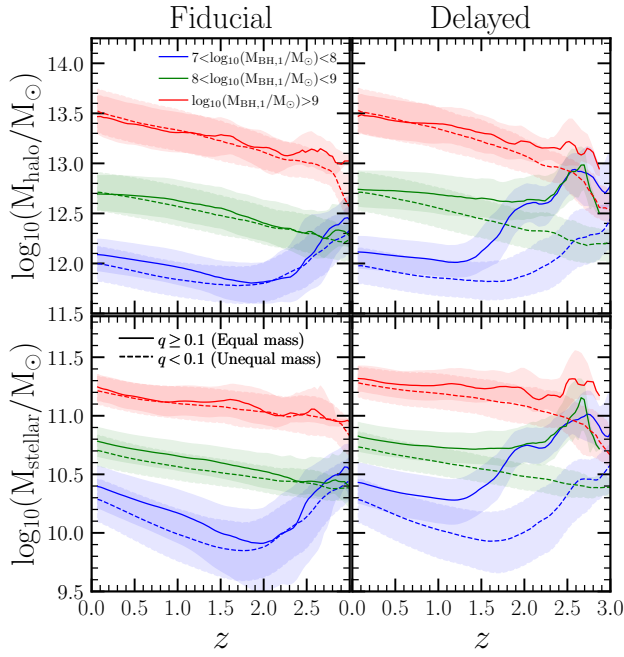
We now explore the properties of the hosts of MBHBs with  $M_{\text{BH},1} > 10^7 M_{\odot}$ . Specifically, we start by studying the stellar and halo content, we then move to the occupation fraction of MBHBs and finally we examine the galaxy morphology. All these properties are explored on based of the binary mass ratio,  $q = M_{\text{BH},2}/M_{\text{BH},1}$ . Here on, we denote as *unequal* binaries those having mass ratio  $q < 0.1$ , and *equal* mass when  $q \geq 0.1$ . We emphasize that, by taking into account Eq. 3 (see also Fig. 1), the typical separation of the binaries studied here is  $\lesssim 300$  pc. Indeed, we have checked that our studied MBHB population displays a peak close to 1 – 10 pc.

##### 4.1 Stellar and halo content of galaxies hosting massive black hole binaries

In Fig. 5 we present the halo and the stellar mass of the galaxies hosting MBHBs. As shown, regardless of  $q$  value and growth model, binaries with  $M_{\text{BH},1} > 10^9 M_{\odot}$  are placed in very massive galaxies and halos:  $10^{11} < M_{\text{stellar}} < 10^{11.5} M_{\odot}$  and

<sup>2</sup> In case the extracted  $f_{\text{Edd}}$  has a value larger than 1, we re-set it at  $f_{\text{Edd}} = 1$

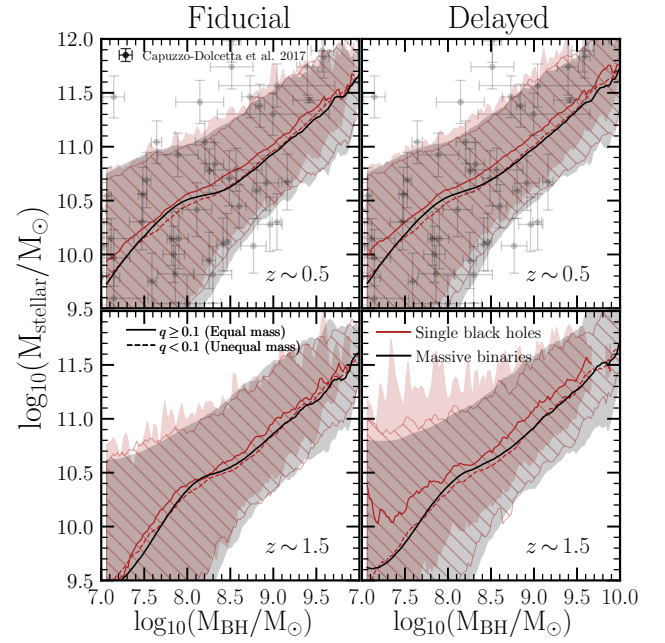




**Figure 5.** Median halo ( $M_{\text{halo}}$ ) and stellar ( $M_{\text{stellar}}$ ) mass of the MBHBs hosts. Each color represents a different mass bin:  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$  (blue),  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  (green) and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$  (red). In all the figures, solid and dotted lines correspond equal ( $q \geq 0.1$ ) and unequal ( $q < 0.1$ ) mass binaries, respectively. The left panel shows the results for the *fiducial model*, while the right panel does it for the *delayed model*.

$10^{13} < M_{\text{halo}} < 10^{13.5} M_{\odot}$ . In particular, the typical halo mass which hosts these MBHBs evolves with redshift. While at  $z > 1.5$  it has masses of  $\sim 10^{13} M_{\odot}$ , at  $z=0$  it increases by 0.5dex. On contrary, the stellar mass displays a milder evolution, reaching by  $z < 1.5$  a constant value of  $M_{\text{stellar}} \sim 10^{11.25} M_{\odot}$ . Such different behaviour is caused by the AGN feedback halting the stellar growth whereas leaving untouched the increase of the halo mass. For MBHBs with  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$ , the typical mass of their hosts increases monotonically towards low- $z$ , reaching by  $z \sim 0$   $M_{\text{halo}} \sim 10^{12.5} M_{\odot}$  and  $M_{\text{stellar}} \sim 10^{10.75} M_{\odot}$ . As happened for more massive systems, no strong differences are seen when the population is divided in  $q$  ratios. Finally, we note that binaries with  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  at  $z < 1$  are placed in Milky-Way type halos and galaxies:  $M_{\text{halo}} \sim 10^{12} M_{\odot}$  and  $M_{\text{stellar}} \sim 10^{10.5} M_{\odot}$ . However, the evolution of these quantities depends on the  $q$  parameter. In particular, unequal mass binaries are systematically placed in less massive halos and galaxies. Larger differences are seen in the stellar masses, where unequal mass binaries can be placed in systems up to 0.2dex less massive than their equal mass counterparts. Interestingly, this trend is more evident in the *delayed model*, where the differences reaches up to  $\sim 1$  dex at  $z > 1$ . Such large deviation seen in the *delayed model* is probably caused by the fact that MBHBs forming equal mass binaries are typically assembled earlier than the ones placed in unequal mass systems. Thus, the earlier formation time of the former population leaves an imprint on the host in which it resides.

To check if the MBHBs follow a different scaling relation than single MBHBs, in Fig. 6 we show at  $z \sim 1.5$  and  $z \sim 0.5$  the  $M_{\text{BH}} - M_{\text{stellar}}$  plane. For completeness we added the  $z=0$  data of Capuzzo-Dolcetta & Tosta e Melo (2017). The results show that

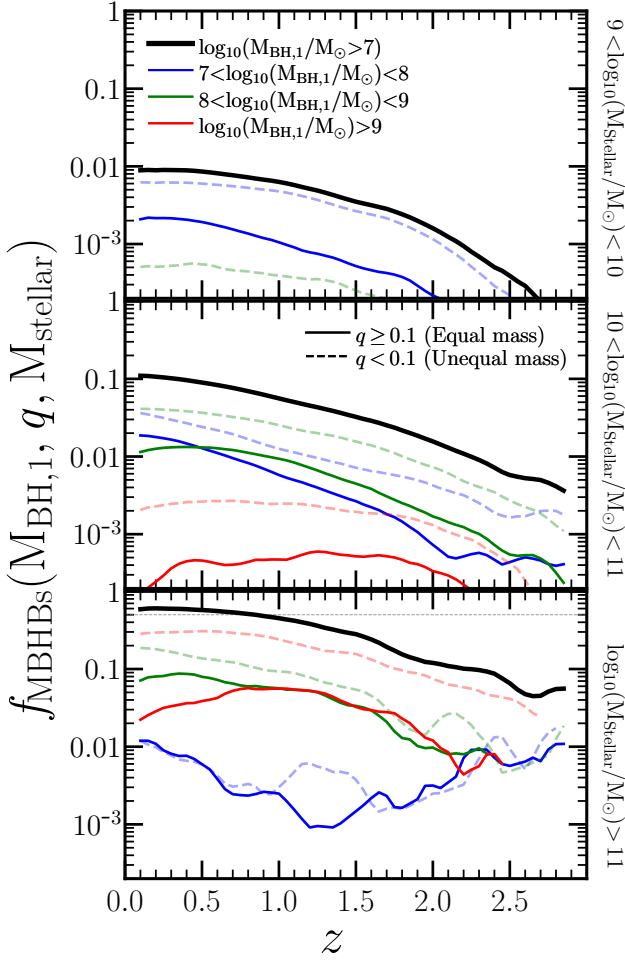


**Figure 6.** Median stellar mass ( $M_{\text{stellar}}$ ) versus the black hole mass ( $M_{\text{BH}}$ ) for single (black colors) and binary black holes (red colors) at  $z \sim 1.5$  and  $z \sim 0.5$ . For binaries, the value of  $M_{\text{BH}}$  corresponds to  $M_{\text{BH},1}$ . In all the panels, solid and dotted lines represent equal ( $q \geq 0.1$ ) and unequal ( $q < 0.1$ ) mass binaries, respectively. The black solid, red solid and red striped areas corresponds to the 10<sup>th</sup> – 90<sup>th</sup> percentile of single MBHBs, unequal and equal mass binaries, respectively. The left panel shows the results for the *fiducial model*, whereas the right panel does it for the *delayed model*. For reference, we have added the observational points of Capuzzo-Dolcetta & Tosta e Melo (2017) at  $z=0$ .

single and binary systems are placed in similar galaxies, regardless of redshift and growth model. When exploring the plane for unequal and equal mass systems, we can see that while the former does not show differences with respect to single MBHBs, the later is systematically placed in slightly more massive galaxies. This trend is particularly evident at  $z \sim 0.5$  in both models and at  $z \sim 1.5$  for the *delayed model*. Nevertheless, these differences are small and the median relation of equal mass systems lies inside the scatter of single MBHBs.

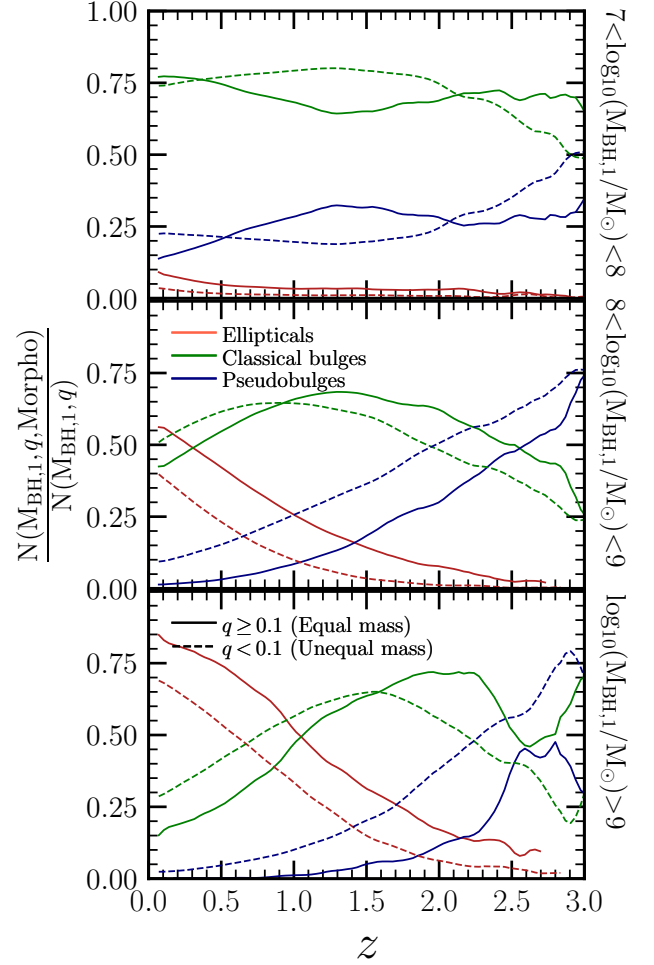
As discussed above, Fig. 5 and Fig. 6 give us useful information about the type of galaxies which host MBHBs, but it do not tell us how common such binaries are. For this, we explore the probability for a given galaxy to host a MBHB with  $M_{\text{BH},1} > 10^7 M_{\odot}$ , hereafter occupation fraction ( $f_{\text{MBHBs}}$ ), which we show in in Fig. 7. Here we define  $f_{\text{MBHBs}}$  as the ratio between galaxies hosting MBHBs and the total number of galaxies. For the sake of simplicity, we only show the results for the *fiducial model* given that we have checked that both growth models explored here display a similar behaviour. Looking at the upper panels of the figure, galaxies with  $10^9 < M_{\text{stellar}} < 10^{10} M_{\odot}$  rarely host a MBHBs with  $M_{\text{BH},1} > 10^7 M_{\odot}$ , regardless of redshift. Although the probability increases towards low redshifts, it never exceeds 1%. When dividing the population in different mass bins and  $q$  values, most of the binaries hosted in these galaxies are unequal mass ( $q < 0.1$ ) with masses  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$ . We highlight that this small occupation fraction is just a consequence of the adopted mass threshold ( $> 10^7 M_{\odot}$ ) and not because MBHBs are rarer in





**Figure 7.** Occupation fraction,  $f_{\text{MBHBs}}$ , as a function of redshift and stellar mass for the *fiducial model*. Black color displays the population of MBHBs with  $M_{\text{BH},1} > 10^7 M_\odot$  while colored ones refer to different mass bins:  $7 < \log_{10}(M_{\text{BH},1}/M_\odot) < 8$  (blue),  $8 < \log_{10}(M_{\text{BH},1}/M_\odot) < 9$  (green) and  $\log_{10}(M_{\text{BH},1}/M_\odot) > 9$  (red). In all the panels solid and dotted line corresponds to equal and unequal mass binaries, respectively.

these galaxies. Indeed, by reducing the mass of the primary black hole down to  $10^6 M_\odot$ , the occupation fraction increases  $\sim 1.5$  dex. For galaxies with  $10^{10} < M_{\text{stellar}} < 10^{11} M_\odot$  we see substantial changes, as the occupation fraction increases by a factor 10 in both the *fiducial* and the *delayed* models. For instance,  $f_{\text{MBHBs}}$  reaches up to 10% by  $z \sim 0$ , although it remains below 1% at  $z > 2$ . When dividing the population in mass and  $q$ , we can appreciate that the evolution of  $f_{\text{MBHBs}}$  is dominated by unequal mass binaries with  $10^7 < M_{\text{BH},1} < 10^9 M_\odot$ . Equal mass binaries are rarer, being present at most in  $\approx 2\%$  of the galaxies at  $z < 1$ . Finally, the lower panel of Fig. 7 shows  $f_{\text{MBHBs}}$  for the most massive galaxies in the lightcone, ( $M_{\text{stellar}} > 10^{11} M_\odot$ ). For these systems, by  $z < 1$  the model predicts a whopping  $f_{\text{MBHBs}} \sim 50\%$ . The typical binaries hosted in these galaxies have  $q < 0.1$  with  $M_{\text{BH},1} > 10^9 M_\odot$ . However, up to 10% of these galaxies can host equal mass MBHBs with primary mass of  $M_{\text{BH},1} > 10^8 M_\odot$ . These results highlight that, perhaps contrary to conventional wisdom, bound MBHBs are very common in massive galaxies at low redshift. In practice, every other massive galaxy at  $z < 1$  should host a binary with separation  $< 300$  pc, and one out of ten should host a relatively equal mass binary. Therefore, massive

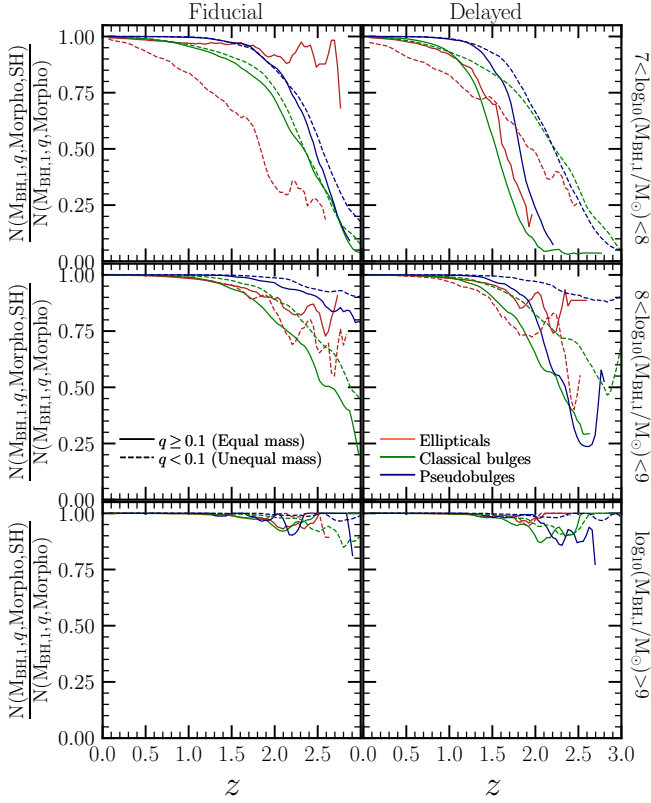


**Figure 8.** Redshift evolution of the morphology of galaxies hosting MBHBs. Red, green and blue color correspond to elliptical-, classical- and pseudo- bulge morphology. Upper, middle and lower panel represent the results for  $7 < \log_{10}(M_{\text{BH},1}/M_\odot) < 8$ ,  $8 < \log_{10}(M_{\text{BH},1}/M_\odot) < 9$  and  $\log_{10}(M_{\text{BH},1}/M_\odot) > 9$ , respectively. In all the panels solid and dotted lines correspond to equal and unequal mass binaries, respectively. We only show the results for the *fiducial model* given that the *delayed* one displays similar trends.

low redshift galaxies should be primary targets for searching MBHBs, although as we will discuss later on that only a small fraction of them is likely to display significant electromagnetic signatures.

## 4.2 The morphology of galaxies hosting massive black hole binaries

Once studied the mass of the MBHB hosts, we explore the redshift evolution of their morphology. As we did before, to avoid crowded plots we only present the results for the *fiducial model* since both growth models display similar trends. From hereafter, we define as elliptical (classical bulge) any galaxy with  $B/T > 0.7$  ( $0.01 < B/T < 0.7$ ), where  $B/T$  is the bulge-to-total stellar mass ratio. On the other hand, a galaxy is tagged as pseudobulge when  $2/3$  of the mass inside the bulge was assembled through disc instabilities



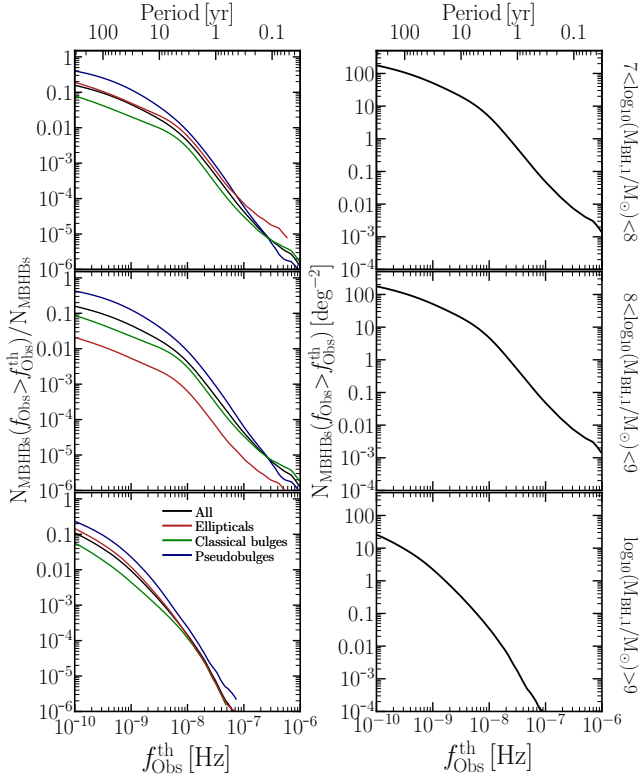
**Figure 9.** Redshift evolution of the fraction of MBHBs shrinking due to stellar hardening. Red, green and blue lines correspond to elliptical-, classical- and pseudo- bulge morphology. Upper, middle and lower panel represent the results for  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$ ,  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$ , respectively. The left panels display the results for the *fiducial* model while the right panels do it for the *delayed* model. In all the panels solid and dotted lines correspond to equal and unequal mass binaries, respectively.

(see Izquierdo-Villalba et al. 2019a, for more details).

In the upper panel of Fig. 8 we present the morphology of the galaxies hosting  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$ . Regardless of redshift and growth model, these binaries rarely inhabit elliptical structures, with less than 10% of the cases at any redshift. The large majority of these binaries (>70%) are hosted in spiral galaxies whose nuclear component is characterized by a classical bulge. The other 20% of the cases corresponds to spiral galaxies with a pseudobulge. Interestingly, this trend is maintained when the population is divided between equal and unequal mass binaries. MBHBs with  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  display similar trends, being spiral galaxies with a classical bulge the preferred hosts at any redshifts and  $q$  value. Nevertheless two differences can be appreciated: elliptical hosts become fairly common at  $z < 1$ , contributing to 25%-50% of the cases, and pseudobulges host >50% of binaries at  $z > 2$ . This is because in L-Galaxies disc instabilities are the main mechanism for bulge growth at  $z > 2$  (Izquierdo-Villalba et al. 2020). Therefore, disc instabilities taking place after a minor merger responsible of the MBHB formation can trigger the development of a pseudobulge, erasing the pre-existent classical bulge structure. A similar process can occur in elliptical galaxies thanks to the fast regeneration of stellar disc at high- $z$ . Finally, MBHBs with  $M_{\text{BH},1} > 10^9 M_{\odot}$  are

mainly placed in elliptical and classical bulge but with a different proportion depending on the redshift considered. While at  $z < 1$  the preferred morphology is the elliptical structure (>50% of the cases), at  $z > 1$  classical bulges take the main role.

As discussed in Section 2.2, binaries hosted in the galactic nucleus are able to reduce their separation through the interaction with a massive circumbinary gaseous discs (gas hardening) or by three-body interactions with single stars (stellar hardening). Given that the model links the formation of each bulge type with different formation scenarios (major/minor mergers or secular disc instabilities), it is interesting to study if the bulge morphology traces different environments around the MBHBs, and thus different processes in which binaries reduce their separation. To explore which shrinking mechanism is the predominant in each galaxy morphology, in Fig. 9 we computed the redshift evolution of the fraction of binaries shrinking through stellar hardening. Despite distinct assumptions on the timing and duration of gas accretion, no strong differences are seen between the *fiducial* and *delayed* model. However, the predictions depend on the mass of the primary black hole. Regardless of redshift and morphology,  $\sim 100\%$  of the binaries with  $M_{\text{BH},1} > 10^9 M_{\odot}$  reduce their separation through encounters with stars. This is the consequence of the fact that these MBHBs are hosted in the most massive galaxies of the simulation. Such systems are characterized by small gas fractions since they consumed most of the gas during the assembly of the stellar component (consistent with van Dokkum 2005; Lin et al. 2008, 2010) and further gas cooling events were efficiently suppressed by the activity of the central massive black hole. In this way, most of the mergers that lead to the formation of MBHBs with  $M_{\text{BH},1} > 10^9 M_{\odot}$  are gas poor, preventing the formation of massive circumbinary discs around the MBHB. This changes when we decrease  $M_{\text{BH},1}$ . For  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  we can see that at  $z > 2$  around 50–70% of the binaries hosted in elliptical galaxies and classical bulges (with slightly smaller values in the *delayed* model) reduce  $a_{\text{BH}}$  through stellar hardening. Interestingly, in pseudobulge structures this percentage increases to 90%. At  $z < 1$  the large majority of the population shrinks via stellar encounters, regardless of bulge morphology. The increasing weight of stellar hardening towards low redshifts is caused by two effects. The first one is that at  $z < 1$  the hosts of these MBHBs are massive enough to experience a similar gas reduction than the ones of  $M_{\text{BH},1} > 10^9 M_{\odot}$  MBHBs. The second effect is caused by the fact that circumbinary discs progressively decrease their mass (and eventually disappear) due to the binary mass accretion. These trends are confirmed for  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$ , where differences between high and low redshift are even larger. At  $z > 2$ , less than the half of the binaries is primarily driven by stellar hardening. The exact percentage varies depending on the galaxy morphology. For instance pseudobulge structures display values  $\sim 50\%$  whereas classical bulge have values  $\sim 20\%$ . When comparing *fiducial* and *delayed* model, we can see that the latter predicts smaller fractions at high- $z$ . This is primarily because in low mass galaxies the lagged gas consumption imposed by the *delayed* model makes more likely that a significant amount of cold gas is still in place when the MBHB forms. This gas forms a massive circumbinary disc that dominates the dynamical evolution of the binary over stellar scattering.



**Figure 10.** **Left panel:** For the *fiducial model*, the fraction of MBHBs with observed frequency,  $f_{\text{obs}}$ , larger than a threshold  $f_{\text{obs}}^{\text{th}}$ . The red, green and blue colors correspond to galaxies with a elliptical-, classical- and pseudo- bulge morphology. **Right panel:** For the *fiducial model*, the number of binaries per square degree (deg $^2$ ) emitting GW at a frequency larger than  $f_{\text{obs}}^{\text{th}}$ . Upper, middle and lower panel represent the results for  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$ ,  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$ , respectively.

### 4.3 The gravitational wave signature of massive black hole binaries

A large fraction of the MBHBs studies here are GW sources at nHz frequencies, the regimen probed by PTAs (Sesana et al. 2008). Given that the common red signal recently detected by these experiments might indeed be the stochastic GW background generated by inspiraling supermassive black hole binaries, it is interesting to study what is the frequency distribution of our MBHBs and which fraction of them will be accessible to PTAs. For that, in the left panels of Fig. 10 we present for the *fiducial model*<sup>3</sup> the cumulative fraction of MBHBs emitting GWs above a given observed frequency,  $f_{\text{obs}}^{\text{th}}$ . The specific value of  $f_{\text{obs}}$  for each MBHBs has been computed as:

$$f_{\text{obs}} = \frac{n}{2\pi(1+z)} \left( (1+q) \frac{M_{\text{BH},1}}{a_{\text{BH}}^3} \right)^{1/2} \quad (13)$$

where  $z$  is the redshift of the binary and  $n$  the harmonic number. Despite our model computes self-consistently the evolution of  $e_{\text{BH}}$ , here, for simplicity, we assume that all the MBHBs are in circular orbits ( $e_{\text{BH}}=0$ ) and, thus radiating only at the  $n=2$  harmonic. As

<sup>3</sup> The *delayed model* displays the same trends and only minor differences are seen.

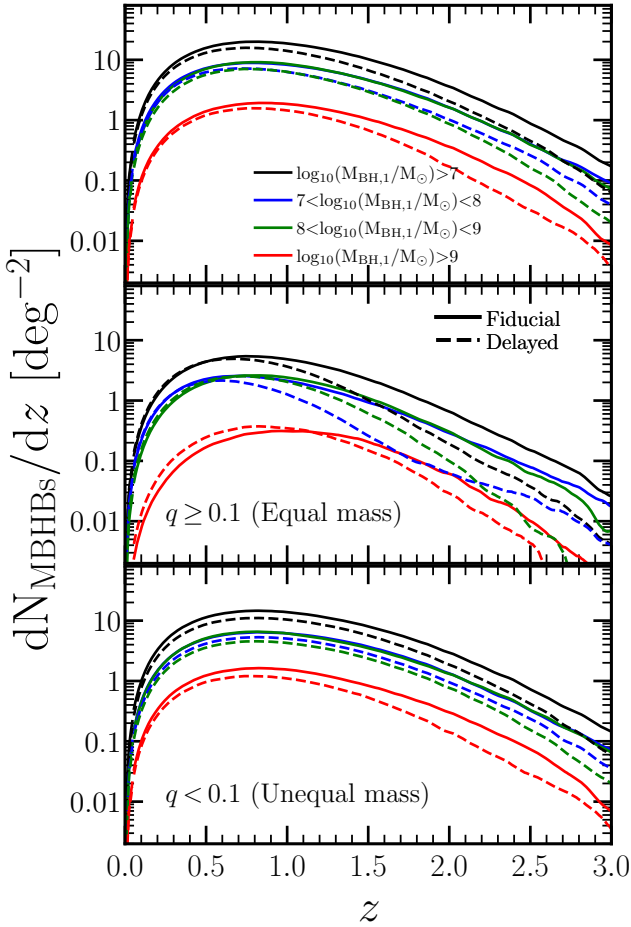
shown, less than 10% of the MBHBs are emitting at frequencies  $f_{\text{obs}}^{\text{th}} > 10^{10}$  Hz, regardless of  $M_{\text{BH},1}$ . This percentage drops down to 1% at  $f_{\text{obs}}^{\text{th}} > 10^9$  Hz. Interestingly, the specific shape of the distribution depends on the galaxy morphology. About 10% of the MBHB hosted in pseudobulge structures emit GW at frequencies  $> 10^9$  Hz, but this percentage drops up to 2dex in the  $10^{-9} < f_{\text{obs}}^{\text{th}} < 10^{-8}$  Hz range. For elliptical and classical bulge structures, the curve displays a smoother behaviour with a distribution that only drops 0.5dex in the same frequency interval. The right panels of Fig. 10 show the number of MBHBs per square degree emitting GWs above a given frequency thresholds. For  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  systems, the model predicts up to 100 objects per deg $^2$  with frequencies  $> 10^{-9}$  Hz. Similar values are found for  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$ . However, binaries with  $M_{\text{BH},1} > 10^9 M_{\odot}$  are much rarer, with only  $\sim 1$  objects per deg $^2$  at  $> 10^{-9}$  Hz. We stress that an in-depth study on the relation between GW and electromagnetic signatures of the PTA sources is deferred in a companion paper in preparation.

## 5 MASSIVE BLACK HOLE BINARIES IN THE SKY: REDSHIFT DISTRIBUTION AND ELECTROMAGNETIC SIGNATURES

In this section, we explore redshift evolution of the number of massive black hole binaries and the fraction of those showing an electromagnetic signature. Besides, we study the occurrence of an active phases in MBHBs and examine if these active systems can be distinguished from quasars/AGNs triggered by single MBHBs.

### 5.1 The redshift distribution of massive black hole binaries

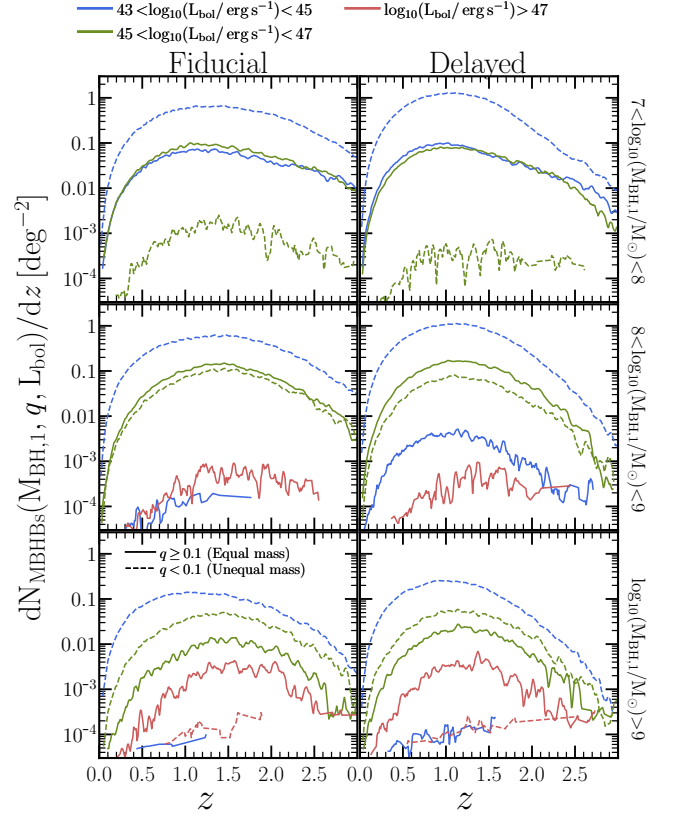
From an observational point of view, one of the first questions that arises when studying MBHBs is at which redshift the probability of finding these systems is the largest. To shed light on this question, in Fig. 11 we present the predicted number of dynamically bound MBHBs per deg $^2$  at different redshifts,  $N_{\text{MBHBs}}$ . Independently of the growth model, L-Galaxies shows an increasing trend until  $z \sim 0.5$ , where up to  $\sim 20$  objects can be found in 1 deg $^2$ . After  $z \sim 0.5$ , the expected number drops very quickly, vanishing at  $z \sim 0$ . When dividing the population in different mass bins, we can see that binaries with  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  and  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  display very similar  $N_{\text{MBHBs}}$  values. However, systems with  $M_{\text{BH},1} > 10^9 M_{\odot}$  are rarer and the largest probability of finding them is at  $z \sim 0.7$  with  $\sim 1$  object per deg $^2$ . In the same plot we have divided the population between equal ( $q \geq 0.1$ ) and unequal mass ( $q < 0.1$ ) binaries. Regardless of mass and model, unequal mass binaries dominate the population, being typically 2–3 times more common. Interestingly, while the predictions for unequal mass systems are not strongly affected by the growth model, in the case of equal mass binaries we find some differences. Specifically, at high redshift ( $z > 1$ ) they are less frequent in the *delayed model*, with the largest differences affecting the  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  and  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  populations. For instance, the *delayed model* predicts 5–7 times less equal mass binaries at  $z > 1$ . Such effect is just a natural consequence of assuming a log-normal distribution in  $f_{\text{Edd}}$  (see Section 3). In the *delayed model* the population of massive black holes are less efficient in assembly their masses, and thus the formation of massive equal mass binaries at high- $z$  is more difficult.



**Figure 11.** Number of massive black hole binaries per square degree ( $\text{deg}^2$ ) for the *fiducial* and *delayed* model (solid and dashed lines, respectively). Black color displays the population of MBHBs with  $M_{\text{BH},1} > 10^7 M_\odot$  while colored ones refer to different mass bins:  $7 < \log_{10}(M_{\text{BH},1}/M_\odot) < 8$  (blue),  $8 < \log_{10}(M_{\text{BH},1}/M_\odot) < 9$  (green) and  $\log_{10}(M_{\text{BH},1}/M_\odot) > 9$  (red). The two lower panels represent the same but dividing the population between equal and unequal mass binaries.

## 5.2 The electromagnetic signatures of massive black hole binaries

Even though the previous section provides useful information about the redshift distribution of MBHBs, it is fundamental to explore how it changes when we only consider systems displaying electromagnetic signatures. In this section we refer to  $L_{\text{bol}}$  as the total luminosity of the system computed as the sum of the primary and secondary black hole luminosity. This choice is motivated by the fact that the projected angular separation of our MBHBs ( $< 0.01 \text{ arcsec}$ ) hampers the detection of dual AGNs by current electromagnetic observatories. In Fig. 12 we present the number of active binaries per square degree radiating at different bolometric luminosity. Regardless of the mass and  $q$  parameter the *fiducial* model shows a maximum at  $z \sim 1.5$ , coinciding with the peak of the galaxy merger rate (see Figure 2 in Izquierdo-Villalba et al. 2019a). By exploring the model predictions at different mass bins, we can see that most of the active population with  $10^7 < M_{\text{BH},1} < 10^8 M_\odot$  displays  $10^{43} < L_{\text{bol}} < 10^{45} \text{ erg/s}$  with  $\sim 0.5$  objects per  $\text{deg}^2$  at

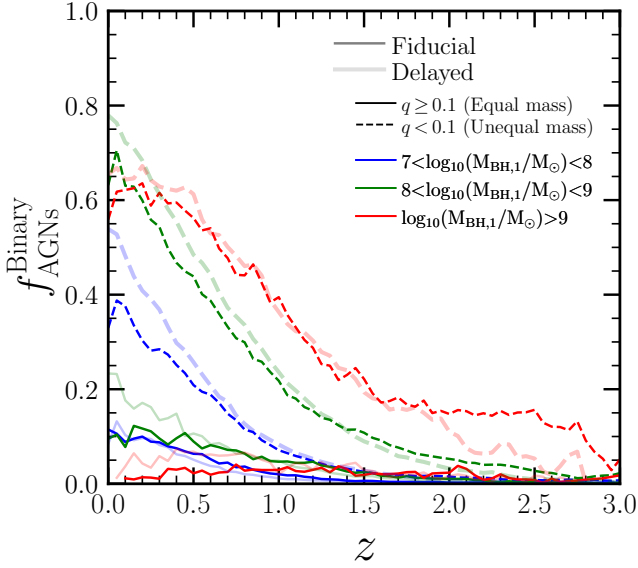


**Figure 12.** Redshift distribution of the number of binaries per square degree radiating at different bolometric luminosity bins. Upper right, lower left and lower right panels represents the results for  $7 < \log_{10}(M_{\text{BH},1}/M_\odot) < 8$ ,  $8 < \log_{10}(M_{\text{BH},1}/M_\odot) < 9$  and  $\log_{10}(M_{\text{BH},1}/M_\odot) > 9$  binaries, respectively. In all panels, solid and dotted lines correspond to equal and unequal mass binaries, respectively. Left and right figures show the results for the *fiducial* and *delayed* model, respectively.

$z \sim 1.5$ . This number decreases quickly towards low- $z$ , reaching by  $z \sim 0.1$  less than  $10^{-3}$  accreting binary per  $\text{deg}^2$ . Interestingly, the MBHBs triggering this faint population are mostly unequal mass binaries. The largest contribution for equal mass systems is found at  $10^{45} < L_{\text{bol}} < 10^{47} \text{ erg/s}$  but it is a much rarer population with  $\lesssim 0.08$  systems per  $\text{deg}^2$  at  $z \sim 1.5$ . Similar trends are found in the  $10^8 < M_{\text{BH},1} < 10^9 M_\odot$  population where most active binaries correspond to unequal mass systems radiating at  $L_{\text{bol}} < 10^{45} \text{ erg/s}$ . Brighter objects are less common (by  $\sim 1$  dex) with the large majority being powered by equal mass systems. Finally, active binaries with  $M_{\text{BH},1} > 10^9 M_\odot$  are  $\sim 1$  dex less frequent than those in the other mass bins. Despite this, the model shows that such kind of binaries are able to power very bright quasars ( $L_{\text{bol}} > 10^{47} \text{ erg/s}$ ) at  $z \sim 2$  but their occurrence drops rapidly after  $z \sim 1$ .

The predictions for the *delayed* model are presented in the right panels of Fig. 12. As shown, the peak of the MBHBs activity does not coincide with the one of the galaxy merger rate. The former happens  $\sim 2 \text{ Gyr}$  after the latter. Besides this, the shape of the distribution has strong differences with respect to the one shown for the *fiducial* model. In particular, the activity of MBHBs is heavily suppressed at high redshifts and enhanced at lower ones. By exploring the predictions as a function of mass, the behavior for  $10^7 < M_{\text{BH},1} < 10^8 M_\odot$  is similar to the one shown by the *fiducial*





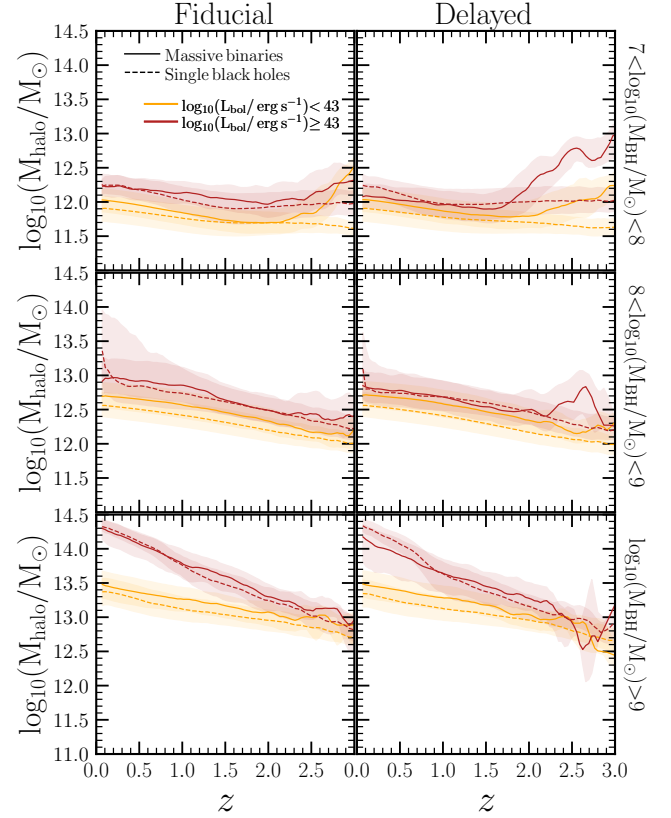
**Figure 13.** Redshift evolution of the probability that a given observed quasar/AGN is powered by a binary system ( $f_{\text{AGNs}}^{\text{Binary}}$ ). Each curve represent  $f_{\text{AGNs}}^{\text{Binary}}$  for different black hole populations:  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$  (blue),  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  (green) and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$  (red). In all panels, solid and dotted lines correspond to equal and unequal mass binaries, respectively. Solid and transparent lines represent the results for the *fiducial* and *delayed* model, respectively.

one: faint and unequal binaries are the main active population. The main difference is in the  $10^{43} < L_{\text{bol}} < 10^{45}$  erg/s population at  $z < 1$  which is up to 5 times more abundant in the *delayed* model. A similar enhancement of sources at low- $z$  is seen for  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  with an important increase of equal mass binaries with  $10^{45} < L_{\text{bol}} < 10^{47}$  erg/s. Finally, we can see that equal mass systems with  $M_{\text{BH},1} > 10^9 M_{\odot}$  take a larger relevance at low- $z$  than in the *fiducial* model. This is particularly evident at  $z \sim 1$  for  $10^{45} < L_{\text{bol}} < 10^{47}$  erg/s where the *delayed* model finds 0.02 equal mass binaries per  $\text{deg}^2$ , while the *fiducial* one predicts few  $10^{-3}$ .

The small projected angular separation of our MBHBs challenges their detection as dual active galactic nuclei. Consequently, without studies addressing periodic lightcurves or Doppler-shifting of AGN broad emission lines, it is not easy differentiate between AGNs powered by single or binary massive black holes. Taking into account this limitation, we explore the probability that a given observed quasar/AGN is powered by a binary system. To this end, we define the quantity  $f_{\text{AGNs}}^{\text{Binary}}$  as:

$$f_{\text{AGNs}}^{\text{Binary}} = \frac{N_{\text{MBHBs}}(L_{\text{bol}} > L^{\text{th}}, M_{\text{BH},1})}{N_{\text{AGNs}}(L_{\text{bol}} > L^{\text{th}}, M_{\text{BH},1})} \quad (14)$$

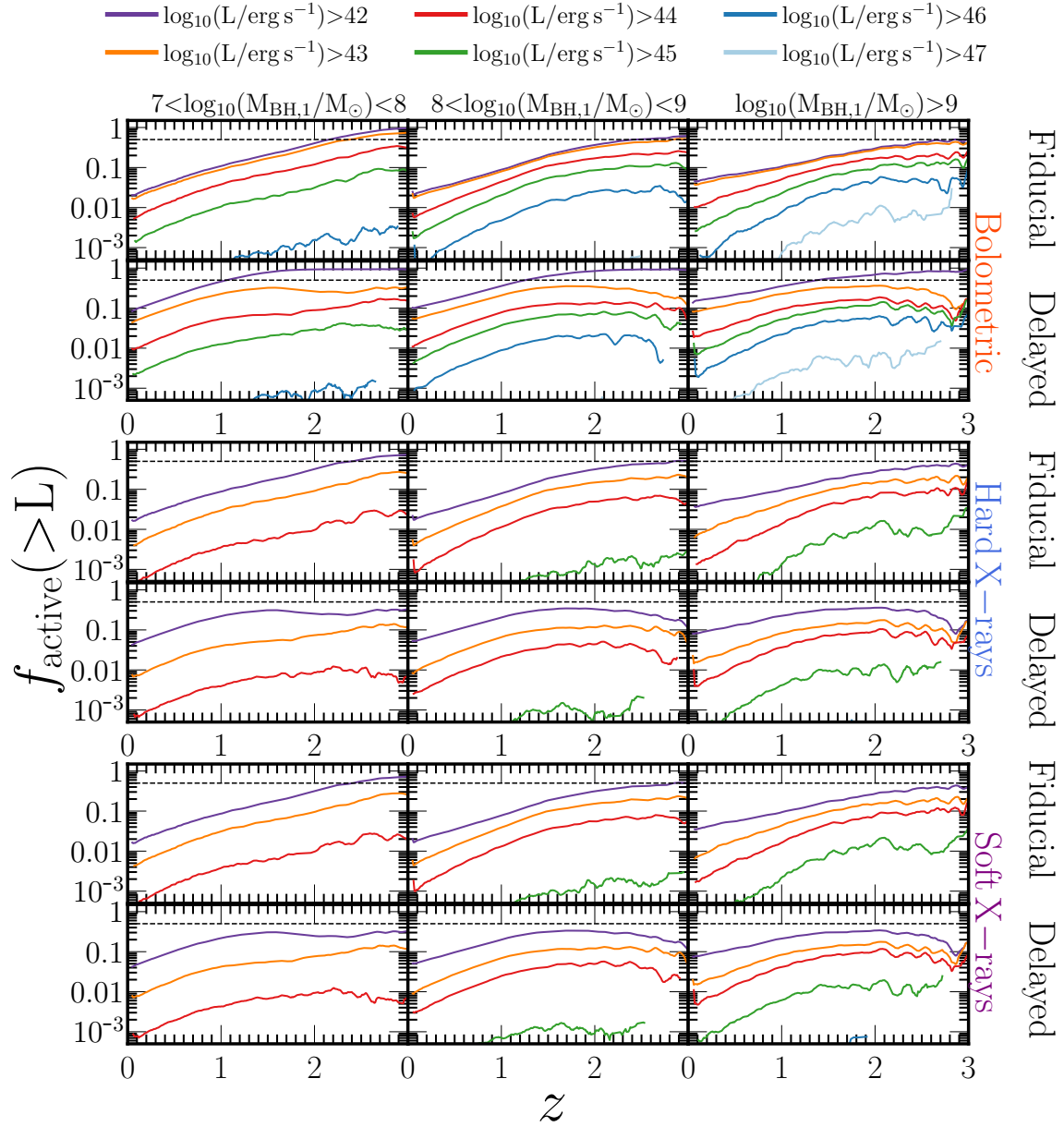
where  $N_{\text{MBHBs}}(L_{\text{bol}} > L^{\text{th}}, M_{\text{BH},1})$  is the total number of active MBHBs with primary MBH mass  $M_{\text{BH},1}$ , and  $N_{\text{AGNs}}(L_{\text{bol}} > L^{\text{th}}, M_{\text{BH},1})$  is the total number of accreting MBHs (binaries and single systems) with mass  $M_{\text{BH},1}$ . Here we assume  $L^{\text{th}} = 10^{43}$  erg/s as the threshold to define an active system but we have checked that similar trends are found when we increase  $L^{\text{th}}$  up to  $10^{45}$  erg/s. The results presented in Fig. 13 show that there are only mild differences between *fiducial* and *delayed* models. Regardless of mass, equal mass systems contribute  $< 20\%$  to the



**Figure 14.** Median halo mass ( $M_{\text{halo}}$ ) of active ( $L_{\text{bol}} > 10^{43}$  erg/s, red) and inactive ( $L_{\text{bol}} < 10^{43}$  erg/s, orange) MBHBs (solid lines) and MBHs (dotted lines). Top, middle and bottom panels represent the results for  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$ ,  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$ , respectively. The left panels show the results for the *fiducial* model whereas the right ones are for the *delayed* model. In all the panels the shaded areas correspond to the 32<sup>th</sup> – 68<sup>th</sup> percentile of the distributions.

total population of active MBHBs. Conversely, unequal mass binaries have a much larger importance with a contribution that strongly correlates with  $M_{\text{BH},1}$ . For instance, MBHBs with  $q < 0.1$  and  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  can represent at  $z=0$  ( $z \sim 0.5$ ) up to 40% (20%) of active MBHBs within their mass range. Larger values are found for binaries with  $M_{\text{BH},1} > 10^9 M_{\odot}$  whose contribution grows up to 60% as redshift goes to zero. Interestingly, this large importance of unequal mass binaries quickly drops towards high redshift. At  $z > 1.5$  the value of  $f_{\text{AGNs}}^{\text{Binary}}$  decreases down to 20% – 5%, regardless of the mass.

To explore indirect ways to disentangle between quasars/AGNs powered by MBHBs and single MBHs, in Fig. 14 we present the evolution of the hosting halo mass for active ( $L_{\text{bol}} > 10^{43}$  erg/s) and inactive ( $L_{\text{bol}} < 10^{43}$  erg/s) MBHBs and MBHs. Despite the halo mass can be a difficult parameter to estimate from observations, it clearly correlates with the environment, i.e number of galaxies around our target (Vale & Ostriker 2004; Shankar et al. 2006; Conroy et al. 2006). Therefore, the halo mass will give us potential information about the typical overdensities in which these systems are hosted. As shown, regardless of mass and model, the active population of both MBHBs and single MBHs is systematically hosted in more massive halos than the inactive one. Nevertheless, by com-



**Figure 15.** Redshift evolution of the active fraction of MBHBs,  $f_{\text{active}}$ , at three different masses of the primary MBH,  $M_{\text{BH},1}$ :  $7 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 8$  (left panels),  $8 < \log_{10}(M_{\text{BH},1}/M_{\odot}) < 9$  (middle panels) and  $\log_{10}(M_{\text{BH},1}/M_{\odot}) > 9$  (right panels). Horizontal dotted line highlights the value of  $f_{\text{active}} = 0.5$ . Different colors display the predictions when different thresholds in luminosity are assumed. The predictions of  $f_{\text{active}}$  have been explored for bolometric luminosity (top panels), hard X-rays (middle panels) and soft X-rays (lower panels).

paring single and binary systems we do not see strong differences. Therefore, the results suggest that the environment of quasars/AGNs will not be a good tracer for pointing out the presence of an active MBHBs with  $M_{\text{BH},1} > 10^7 M_{\odot}$ . We have checked that the same trends are kept when we divide the population of active MBHBs between equal and unequal mass.

### 5.3 The occurrence of active massive black hole binaries

In this section we study how frequent is the active phase of MBHBs. In Fig. 15 we present for different  $M_{\text{BH},1}$  the fraction of active

MBHBs,  $f_{\text{active}}$ , defined as

$$f_{\text{active}} = \frac{N_{\text{MBHBs}}(M_{\text{BH},1}, > L_{\text{bol}})}{N_{\text{MBHBs}}(M_{\text{BH},1})}, \quad (15)$$

where  $N_{\text{MBHBs}}(M_{\text{BH},1}, > L_{\text{bol}})$  is the number of MBHBs with  $M_{\text{BH},1}$  radiating at luminosity larger than  $L_{\text{bol}}$ . On the other hand,  $N_{\text{MBHBs}}(M_{\text{BH},1})$  is the total number of MBHBs with primary MBH mass  $M_{\text{BH},1}$ . As shown, at  $z > 2$  the *fiducial model* predicts that a large fraction of the MBHB population with  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  displays some level of gas accretion. For instance, the values of  $f_{\text{active}}$  can be  $\gtrsim 0.5$  when a threshold of  $L_{\text{bol}} = 10^{43} \text{ erg/s}$  is assumed. However, in the same redshift range,  $f_{\text{active}}$  can drop down to 0.1 (0.01) when the more strict limiting luminosity  $L_{\text{bol}} > 10^{45} \text{ erg/s}$

( $L_{\text{bol}} > 10^{46}$  erg/s) is imposed. At lower redshifts, the value of  $f_{\text{active}}$  decreases very quickly at any luminosity threshold, reaching  $f_{\text{active}} \leq 0.1$  by  $z < 0.5$ . This fast drop is caused by the combination of two different processes. The first one is the formation of new MBHBs embedded in gas poor environments unable to trigger AGN activity. The second is the fast decrease of gas reservoir around MBHBs caused by the gas consumption (see Section 3). This latter effect is particularly evident when we compare the predictions of  $f_{\text{active}}$  between the *fiducial* and *delayed model*. As shown, the latter one displays values of  $f_{\text{active}}$  systematically larger regardless of the bolometric luminosity and redshift. When considering  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  and  $M_{\text{BH},1} > 10^9 M_{\odot}$ , we find similar trends to those seen in lighter MBHBs: at any luminosity there is a decreasing trend of  $f_{\text{active}}$  towards low- $z$  being earlier and more pronounced in the *fiducial model*. Interestingly, in both growth models the most luminous MBHBs ( $L_{\text{bol}} > 10^{45}$  erg/s) powered by  $M_{\text{BH},1} > 10^8 M_{\odot}$  at  $z < 1$  contribute less than 10% of the total MBHBs population within their mass range. All these results highlight the challenges of future multi-messenger studies as a consequence of the elusive nature of the most massive MBHBs.

Given current and future facilities exploring the X-ray sky with large sensitivity (see, e.g. eROSITA, Merloni et al. 2012, Athena, Nandra et al. 2013 or Lynx The Lynx Team 2018), in Fig. 15 we also show the model predictions when using X-rays luminosity as threshold. Specifically, to transform bolometric to soft (0.5–2 keV) and hard (2–10 keV) X-ray luminosity we make use of the corrections derived in Marconi et al. (2004):

$$\log_{10}(L_{\text{Hx}}/L_{\text{bol}}) = -1.54 - 0.24\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3 \quad (16)$$

$$\log_{10}(L_{\text{Sx}}/L_{\text{bol}}) = -1.64 - 0.22\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3 \quad (17)$$

where  $\mathcal{L} = \log_{10}(L_{\text{bol}}/L_{\odot}) - 12$ , and  $L_{\text{Hx}}$  ( $L_{\text{Sx}}$ ) is the hard (soft) X-ray luminosity. As we can see, soft and hard X-rays predictions display very similar trends. As for the bolometric luminosity, the predictions differ between the *fiducial model* and *delayed model*, especially at low- $z$ . Despite that, both growth models predict that  $f_{\text{active}}$  rarely overpass values of 0.5, regardless of redshift and X-ray luminosity.

## 6 SUMMARY AND CONCLUSIONS

In this paper we studied the hosts and electromagnetic signatures of parsec-scale MBHBs with primary black hole mass  $M_{\text{BH},1} > 10^7 M_{\odot}$ . To this purpose, we made use of the L-Galaxies semi-analytical model applied on the Millennium merger trees. Specifically, we used the L-Galaxies version presented in Izquierdo-Villalba et al. (2019b) and Izquierdo-Villalba et al. (2022) where a lightcone construction and several prescriptions for tracing the formation and dynamical evolution of MBHBs had been included. Since the evolution of the MBHs and MBHBs depends on the specific prescriptions implemented for describing the mass growth, we explored two different models: a *fiducial model* where MBHs and MBHBs exhaust their gas reservoir in a *fast* and efficient way, and a *delayed model* allowing for longer periods of gas accretion. These two different models allowed us to explore how the uncertainties in the gas consumption rate affect our predictions. The main results of this work can be summarized as follows:

- Number of MBHBs versus redshift: The population of parsec-scale MBHBs is especially common at  $0.5 < z < 1$ , ( $\sim 20$  objects per deg<sup>2</sup>) with unequal mass binaries ( $q \leq 0.1$ ) a factor 2–3 more

abundant than equal-mass systems. In the *delayed model*, as a consequence of the retarded mass growth, the number of systems with  $q \geq 0.1$  drops by one order of magnitude by  $z \sim 3$ .

- Masses of the host galaxies: Binaries with  $M_{\text{BH},1} > 10^8 M_{\odot}$  are located in very massive systems, with halo and stellar masses  $\sim 10^{12.5} M_{\odot}$  and  $\sim 10^{11} M_{\odot}$ , respectively. Conversely, binaries with  $10^7 < M_{\text{BH},1} < 10^8 M_{\odot}$  are placed galaxies with halos and stellar masses  $\sim 10^{12} M_{\odot}$  and  $\sim 10^{10} M_{\odot}$ , respectively. On top of this, single MBHs and MBHBs populate the same plane of the  $M_{\text{BH}} - M_{\text{stellar}}$  scaling relation, ruling out that binaries are outliers of the scaling relations.

- Morphology of the host: Spiral galaxies with a classical bulge component are the preferred hosts of MBHBs with  $M_{\text{BH},1} > 10^7 M_{\odot}$ . However, the model predicts that ellipticals become important in the local Universe, especially for systems with  $M_{\text{BH},1} > 10^9$  where they account for the  $\sim 75\%$  of the hosts. Interestingly, *fiducial* and *delayed model* display a compatible behaviour. Thus, the morphology of the galaxy where a MBHB is placed is just driven by the galaxy merger history, regardless of the growth path followed by the binary system.

- Occupation fraction: Less than 1% of galaxies with  $10^9 < M_{\text{stellar}} < 10^{10} M_{\odot}$  host a MBHBs with  $M_{\text{BH},1} > 10^7 M_{\odot}$ . On contrary, up to 10% of galaxies with  $10^{10} < M_{\text{stellar}} < 10^{11} M_{\odot}$  can host a MBHBs, with unequal mass binaries of  $10^8 < M_{\text{BH},1} < 10^9 M_{\odot}$  the preferred systems. For the most massive galaxies in the model ( $M_{\text{stellar}} > 10^{11} M_{\odot}$ ), the MBHB occupation fraction can reach up to 50%, suggesting that they are the perfect targets for finding low- $z$  parsec-scale binaries. In particular, MBHBs hosted by these massive galaxies are unequal systems with  $M_{\text{BH},1} > 10^9 M_{\odot}$ , but a still significant number of galaxies ( $\sim 10\%$ ) host equal mass binaries with  $M_{\text{BH},1} > 10^8 M_{\odot}$ . We checked that these results are independent of the growth model explored.

- Hardening mechanisms: Regardless of the galaxy morphology, the main shrinking process of binaries with  $M_{\text{BH},1} > 10^9 M_{\odot}$  is through encounters with single stars. For lighter systems, the process displays a redshift evolution. While  $z < 2$  the vast majority of the systems reduce their separation through stellar hardening, at  $z > 2$  the situation inverts and more than 50% experience a gas driven hardening. Interestingly, this percentage is larger in classical bulges than in elliptical and pseudobulge structures.

- MBHBs in the PTA (nHz) band: Looking at the distribution of binary periods, ranging between 100 and 1 years, and their associated GW emission frequency, we found that less than 1% of MBHBs hosted in elliptical galaxies and classical bulges have observed frequencies accessible by PTA experiments ( $10^{-9} < f_{\text{Obs}} < 10^{-7}$  Hz). Moreover, this fraction is systematically smaller in elliptical structures than in classical bulges. Our model indicates that independently of the growth model used,  $\sim 30$  binary systems per deg<sup>2</sup> should be emitting gravitational waves at the PTA frequencies.

- MBHBs as active AGNs: In the *fiducial model* the maximum of the MBHB AGN activity is reached at  $z \sim 1.5$ . In the *delayed model*, important accretion events onto MBHBs are highly suppressed at high redshifts causing a delay of  $\sim 2$  Gyr in the peak of the activity. In both models, active MBHBs are mainly unequal-mass systems

emitting primarily at sub-Eddington level with luminosity spanning between  $10^{43} < L_{\text{bol}} < 10^{45}$  erg/s. At  $z \sim 1$ , we find up to 0.2 – 0.5 (1 – 2) of these binaries per deg<sup>2</sup> in the *fiducial (delayed) model*. For equal-mass systems, the largest contribution is at  $L_{\text{bol}} > 10^{45}$  erg/s but less than 0.06 – 0.08 objects per deg<sup>2</sup> are predicted.

- Fraction of active MBHBs: The fraction of active binaries strongly depends on the luminosity threshold. In general, at  $z > 2$  a large fraction of MBHBs (>40%) should be in an active phase. On the other hand, at  $z < 1$  and taking a fiducial threshold of  $10^{43}$  erg/s, less than 10% of the systems can be considered active in any band, including hard and soft X-rays. When comparing the halo mass of the galaxies hosting active MBHBs and single MBHs we find that both samples display similar values, highlighting that the environments of AGNs are not good tracers to reveal the presence of an active MBHBs.

Given all the results summarized above we can conclude that parsec-scale massive black hole binaries with primary mass  $> 10^7 M_{\odot}$  are a common population at  $z < 1$ , being hosted in an important fraction of the most massive galaxies in the local universe. However, most of these MBHBs are expected to be inactive or very dim, challenging future multi-messenger astronomy studies. Future observations with JWST will be crucial to shed light on that faint AGN population, guiding our understanding of active MBHBs (Seth et al. 2021).

## ACKNOWLEDGEMENTS

The authors warmly thank Silvia Bonoli for the useful discussions and comments. D.I.V. and A.S. acknowledge financial support provided under the European Union’s H2020 ERC Consolidator Grant “Binary Massive Black Hole Astrophysics” (B Massive, Grant Agreement: 818691). D.I.V. acknowledges also financial support from INFN H45J18000450006. M.C. acknowledges funding from MIUR under the Grant No. PRIN 2017-MB8AEZ.

## DATA AVAILABILITY

The simulated data underlying this article will be shared on reasonable request to the corresponding author. This work used the 2015 public version of the Munich model of galaxy formation and evolution: L-Galaxies. The source code and a full description of the model are available at <http://galformod.mpa-garching.mpg.de/public/LGalaxies/>.

## REFERENCES

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, *MNRAS*, **451**, 1892  
 Amaro-Seoane P., et al., 2017, arXiv e-prints, p. [arXiv:1702.00786](https://arxiv.org/abs/1702.00786)  
 Angulo R. E., White S. D. M., 2010, *MNRAS*, **405**, 143  
 Arzoumanian Z., et al., 2015, *ApJ*, **810**, 150  
 Arzoumanian Z., et al., 2020, *ApJ*, **905**, L34  
 Bailes M., et al., 2016, in MeerKAT Science: On the Pathway to the SKA. p. 11 ([arXiv:1803.07424](https://arxiv.org/abs/1803.07424))  
 Barausse E., Rezzolla L., 2009, *ApJ*, **704**, L40  
 Begelman M. C., Blandford R. D., Rees M. J., 1980, *Nature*, **287**, 307  
 Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition  
 Bogdanović T., Eracleous M., Sigurdsson S., 2009, *ApJ*, **697**, 288

Bogdanovic T., Miller M. C., Blecha L., 2021, arXiv e-prints, p. [arXiv:2109.03262](https://arxiv.org/abs/2109.03262)  
 Bonetti M., Haardt F., Sesana A., Barausse E., 2018, *MNRAS*, **477**, 3910  
 Bonetti M., Sesana A., Haardt F., Barausse E., Colpi M., 2019, *MNRAS*, **486**, 4044  
 Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, *MNRAS*, **396**, 423  
 Boylan-Kolchin M., Ma C.-P., Quataert E., 2006, *MNRAS*, **369**, 1081  
 Capuzzo-Dolcetta R., Tosta e Melo I., 2017, *Monthly Notices of the Royal Astronomical Society*, **472**, 4013  
 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  
 Charisi M., Taylor S. R., Runnoe J., Bogdanovic T., Trump J. R., 2022, *MNRAS*, **510**, 5929  
 Chen S., et al., 2021, *Monthly Notices of the Royal Astronomical Society*, **508**, 4970  
 Colpi M., 2014, *Space Sci. Rev.*, **183**, 189  
 Colpi M., Mayer L., Governato F., 1999, *ApJ*, **525**, 720  
 Comerford J. M., Gerke B. F., Stern D., Cooper M. C., Weiner B. J., Newman J. A., Madsen K., Barrows R. S., 2012, *ApJ*, **753**, 42  
 Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, **647**, 201  
 Cuadra J., Armitage P. J., Alexander R. D., Begelman M. C., 2009, *MNRAS*, **393**, 1423  
 De Rosa A., et al., 2019, *New Astron. Rev.*, **86**, 101525  
 Desvignes G., et al., 2016, *MNRAS*, **458**, 3341  
 Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, **433**, 604  
 Dotti M., Colpi M., Haardt F., Mayer L., 2007, *MNRAS*, **379**, 956  
 Dotti M., Merloni A., Montuori C., 2015, *MNRAS*, **448**, 3603  
 Dressler A., Richstone D. O., 1988, *ApJ*, **324**, 701  
 Duffell P. C., D’Orazio D., Derdzinski A., Haiman Z., MacFadyen A., Rosen A. L., Zrake J., 2020, *ApJ*, **901**, 25  
 Eracleous M., Boroson T. A., Halpern J. P., Liu J., 2012, *ApJS*, **201**, 23  
 Escala A., Larson R. B., Coppi P. S., Mardones D., 2004, *ApJ*, **607**, 765  
 Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, *ApJ*, **630**, 152  
 Franchini A., Sesana A., Dotti M., 2021, *MNRAS*, **507**, 1458  
 Gadotti D. A., 2009, *MNRAS*, **393**, 1531  
 Genzel R., Townes C. H., 1987, *ARA&A*, **25**, 377  
 Genzel R., Hollenbach D., Townes C. H., 1994, *Reports on Progress in Physics*, **57**, 417  
 Goldstein J. M., Sesana A., Holgado A. M., Veitch J., 2019, *MNRAS*, **485**, 248  
 Goncharov B., et al., 2021, *ApJ*, **917**, L19  
 Graham M. J., et al., 2015, *MNRAS*, **453**, 1562  
 Guo Q., et al., 2011, *MNRAS*, **413**, 101  
 Haehnelt M. G., Rees M. J., 1993, *Monthly Notices of the Royal Astronomical Society*, **263**, 168  
 Häring N., Rix H.-W., 2004, *ApJ*, **604**, L89  
 Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, **451**, 2663  
 Hernquist L., 1989, *Nature*, **340**, 687  
 Hobbs G., et al., 2010, *Classical and Quantum Gravity*, **27**, 084013  
 Hopkins P. F., Quataert E., 2010, *MNRAS*, **407**, 1529  
 Hopkins P. F., Narayan R., Hernquist L., 2006, *ApJ*, **643**, 641  
 Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, *ApJ*, **691**, 1168  
 Hu W.-R., Wu Y.-L., 2017, *National Science Review*, **4**, 685  
 Ivezić Ž., et al., 2019, *ApJ*, **873**, 111  
 Izquierdo-Villalba D., Bonoli S., Spinoso D., Rosas-Guevara Y., Henriques B. M. B., Hernández-Monteagudo C., 2019a, *MNRAS*, **488**, 609  
 Izquierdo-Villalba D., et al., 2019b, *A&A*, **631**, A82  
 Izquierdo-Villalba D., Bonoli S., Dotti M., Sesana A., Rosas-Guevara Y., Spinoso D., 2020, *MNRAS*, **495**, 4681  
 Izquierdo-Villalba D., Sesana A., Bonoli S., Colpi M., 2022, *MNRAS*, **509**, 3488  
 Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, *MNRAS*, **307**, 529  
 Kelley L. Z., Blecha L., Hernquist L., 2017, *MNRAS*, **464**, 3131



- Kelley L. Z., Blecha L., Hernquist L., Sesana A., Taylor S. R., 2018, *MNRAS*, **477**, 964
- Kelley L. Z., Haiman Z., Sesana A., Hernquist L., 2019, *MNRAS*, **485**, 1579
- Kollmeier J. A., et al., 2006, *ApJ*, **648**, 128
- Kormendy J., 1988, *ApJ*, **325**, 128
- Kormendy J., Ho L. C., 2013, *ARA&A*, **51**, 511
- Kormendy J., Richstone D., 1992, *ApJ*, **393**, 559
- Koss M., Mushotzky R., Treister E., Veilleux S., Vasudevan R., Tripp M., 2012, *ApJ*, **746**, L22
- Liao W.-T., et al., 2021, *MNRAS*, **500**, 4025
- Lin L., et al., 2008, *ApJ*, **681**, 232
- Lin L., et al., 2010, *ApJ*, **718**, 1158
- Liu T., et al., 2016, *ApJ*, **833**, 6
- Liu T., et al., 2019, *ApJ*, **884**, 36
- Luò J., et al., 2016, *Classical and Quantum Gravity*, **33**, 035010
- Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, *MNRAS*, **351**, 169
- Márquez I., Lima Neto G. B., Capelato H., Durret F., Gerbal D., 2000, *A&A*, **353**, 873
- Martínez Palafox E., Valenzuela O., Colín P., Gottlöber S., 2014, arXiv e-prints, p. [arXiv:1410.3563](https://arxiv.org/abs/1410.3563)
- Marulli F., Bonoli S., Branchini E., Moscardini L., Springel V., 2008, *MNRAS*, **385**, 1846
- Merloni A., et al., 2012, arXiv e-prints, p. [arXiv:1209.3114](https://arxiv.org/abs/1209.3114)
- Mo H., van den Bosch F. C., White S., 2010, *Galaxy Formation and Evolution*
- Montuori C., Dotti M., Colpi M., Decarli R., Haardt F., 2011, *MNRAS*, **412**, 26
- Naab T., Burkert A., 2001, in Knapen J. H., Beckman J. E., Shlosman I., Mahoney T. J., eds, *Astronomical Society of the Pacific Conference Series* Vol. 249, *The Central Kiloparsec of Starbursts and AGN: The La Palma Connection*. p. 735 ([arXiv:astro-ph/0110374](https://arxiv.org/abs/astro-ph/0110374))
- Nandra K., et al., 2013, arXiv e-prints, p. [arXiv:1306.2307](https://arxiv.org/abs/1306.2307)
- O'Dowd M., Urry C. M., Scarpa R., 2002, *ApJ*, **580**, 96
- Perera B. B. P., et al., 2019, *MNRAS*, **490**, 4666
- Peters P. C., Mathews J., 1963, *Phys. Rev.*, **131**, 435
- Peterson B. M., et al., 2004, *ApJ*, **613**, 682
- Planck Collaboration et al., 2014, *A&A*, **571**, A16
- Quinlan G. D., Hernquist L., 1997, *New Astron.*, **2**, 533
- Reardon D. J., et al., 2016, *MNRAS*, **455**, 1751
- Rodriguez C., Taylor G. B., Zavala R. T., Peck A. B., Pollack L. K., Romani R. W., 2006, *ApJ*, **646**, 49
- Roedig C., Dotti M., Sesana A., Cuadra J., Colpi M., 2011, *MNRAS*, **415**, 3033
- Rosado P. A., Sesana A., 2014, *MNRAS*, **439**, 3986
- Rosado P. A., Sesana A., Gair J., 2015, *MNRAS*, **451**, 2417
- Ruan W.-H., Guo Z.-K., Cai R.-G., Zhang Y.-Z., 2018, arXiv e-prints, p. [arXiv:1807.09495](https://arxiv.org/abs/1807.09495)
- Savorgnan G. A. D., Graham A. W., Marconi A. r., Sani E., 2016, *ApJ*, **817**, 21
- Sersic J. L., 1968, *Atlas de Galaxias Australes*
- Sesana A., Khan F. M., 2015, *MNRAS*, **454**, L66
- Sesana A., Haardt F., Madau P., Volonteri M., 2004, *ApJ*, **611**, 623
- Sesana A., Haardt F., Madau P., 2006, *ApJ*, **651**, 392
- Sesana A., Vecchio A., Colacino C. N., 2008, *MNRAS*, **390**, 192
- Sesana A., Vecchio A., Volonteri M., 2009, *MNRAS*, **394**, 2255
- Sesana A., Roedig C., Reynolds M. T., Dotti M., 2012, *MNRAS*, **420**, 860
- Sesana A., Barausse E., Dotti M., Rossi E. M., 2014, *ApJ*, **794**, 104
- Seth A. C., et al., 2021, *Revealing Low Luminosity Active Galactic Nuclei (ReveaLLAGN)*, JWST Proposal. Cycle 1, ID. #2016
- Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, *ApJ*, **643**, 14
- Shen Y., Liu X., Loeb A., Tremaine S., 2013, *ApJ*, **775**, 49
- Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, *MNRAS*, **495**, 3252
- Spinoso D., Bonoli S., Valiante R., Schneider R., Izquierdo-Villalba D., 2022, arXiv e-prints, p. [arXiv:2203.13846](https://arxiv.org/abs/2203.13846)
- Springel V., 2005, *MNRAS*, **364**, 1105
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, **328**, 726
- Susobhanan A., et al., 2021, *Publ. Astron. Soc. Australia*, **38**, e017
- Tanaka T., Haiman Z., 2009, *ApJ*, **696**, 1798
- Tanaka T., Menou K., Haiman Z., 2012, *MNRAS*, **420**, 705
- Terzić B., Graham A. W., 2005, *MNRAS*, **362**, 197
- The Lynx Team 2018, arXiv e-prints, p. [arXiv:1809.09642](https://arxiv.org/abs/1809.09642)
- Tsalmantza P., Decarli R., Dotti M., Hogg D. W., 2011, *ApJ*, **738**, 20
- Vale A., Ostriker J. P., 2004, *MNRAS*, **353**, 189
- Valtonen M. J., et al., 2008, *Nature*, **452**, 851
- Vasiliev E., Antonini F., Merritt D., 2014, *ApJ*, **785**, 163
- Vaughan S., Uttley P., Markowitz A. G., Huppenkothen D., Middleton M. J., Alston W. N., Scargle J. D., Farr W. M., 2016, *MNRAS*, **461**, 3145
- Vestergaard M., Peterson B. M., 2006, *ApJ*, **641**, 689
- Volonteri M., Haardt F., Madau P., 2003, *ApJ*, **582**, 559
- Volonteri M., et al., 2020, *MNRAS*, **498**, 2219
- White S. D. M., Frenk C. S., 1991, *ApJ*, **379**, 52
- White S. D. M., Rees M. J., 1978, *MNRAS*, **183**, 341
- Witt C. A., Charisi M., Taylor S. R., Burke-Spolaor S., 2021, arXiv e-prints, p. [arXiv:2110.07465](https://arxiv.org/abs/2110.07465)
- van Dokkum P. G., 2005, *AJ*, **130**, 2647

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.