Astronomy & Astrophysics

Connecting low-redshift LISA massive black hole mergers to the nHz stochastic gravitational wave background

David Izquierdo-Villalba[1,2] , Alberto Sesana[1,2] , Monica Colpi[1,2], Daniele Spinoso[3], Matteo Bonetti[1,2], Silvia Bonoli[4,5], and Rosa Valiante[6,7]

[1] Dipartimento di Fisica "G. Occhialini", Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy
  e-mail: david.izquierdovillalba@unimib.it
[2] INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy
[3] Department of Astronomy, MongManWai Building, Tsinghua University, Beijing 100084, PR China
[4] Donostia International Physics Centre (DIPC), Paseo Manuel de Lardizabal 4, 20018 Donostia-San Sebastian, Spain
[5] IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain
[6] INAF/Osservatorio Astronomico di Roma, Via Frascati 33, 00040 Monte Porzio Catone, Italy
[7] INFN, Sezione Roma 1, Dipartimento di Fisica, "Sapienza" Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy

ABSTRACT

Pulsar Timing Array (PTA) experiments worldwide recently reported evidence of a nHz stochastic gravitational wave background (sGWB) compatible with the existence of slowly inspiralling massive black hole (MBH) binaries (MBHBs). The shape of the signal contains valuable information about the evolution of $z < 1$ MBHs above $10^8 \, M_\odot$, suggesting a faster dynamical evolution of MBHBs towards the gravitational-wave-driven inspiral or a larger MBH growth than usually assumed. In this work, we investigate if the nHz sGWB could also provide constraints on the population of merging lower-mass MBHBs ($<10^7 \, M_\odot$) detectable by LISA. To this end, we use the L-Galaxies semi-analytical model applied to the Millennium suite of simulations. We generate a population of MBHs compatible simultaneously with current electromagnetic and nHz sGWB constraints by including the possibility that, in favourable environments, MBHs can accrete gas beyond the Eddington limit. The predictions of this new model for the sGWB show that the global (integrated up to high-z) LISA detection rate is not significantly affected when compared to a fiducial model whose nHz sGWB signal is ∼2 times smaller. In both cases, the global rate yields ∼12 yr^{-1} and is dominated by systems of $10^{5-6} \, M_\odot$. The main differences are limited to low-z ($z < 3$), high-mass ($>10^6 \, M_\odot$) LISA MBHBs. The model compatible with the latest PTA results predicts up to ∼1.6 times more detections, with a rate of ∼1 yr^{-1}. We find that these LISA MBHB systems have 50% probability of shining with bolometric luminosities $>10^{43}$ erg s^{-1}. Hence, in case PTA results are confirmed and given the current MBH modelling, our findings suggest there will be higher chances to perform multimessenger studies with LISA MBHB than previously expected.

Key words. galaxies: dwarf – galaxies: evolution – galaxies: general – galaxies: interactions – quasars: general – quasars: supermassive black holes

1. Introduction

The large fraction of massive galaxies hosting massive black holes (MBHs) together with the hierarchical assembly of galaxies suggest that massive black hole binaries (MBHBs) are unavoidable products of galaxy evolution. The formation of these systems and their coalescence within the Hubble time is regulated by several processes (Begelman et al. 1980). Large-scale (∼kiloparsecs) dynamical friction, acting after galaxy mergers, causes the sinking of the MBHs towards the centre of the newly formed galaxy (Milosavljević & Merritt 2001; Yu 2002; Mayer et al. 2007; Callegari et al. 2009, 2011; Bortolas et al. 2022; Li et al. 2022; Amaro-Seoane et al. 2023). At smaller separations (<parsecs), the interaction with single stars and a putative circumbinary disc surrounding the two MBHs brings the two objects down to a distance where the emission of gravitational waves (GWs) lead to their final coalescence (Quinlan & Hernquist 1997; Sesana et al. 2006; Vasiliev et al. 2014; Sesana & Khan 2015; Escala et al. 2004, 2005; Dotti et al. 2007; Cuadra et al. 2009; Biava et al. 2019; Bonetti et al. 2020; Franchini et al. 2021, 2022).

Gravitational waves (GWs) emitted by MBHBs during the late inspiral, merger, and ringdown phase at the end of their lives are the main target of current and future experiments. On the one hand, the Laser Interferometer Space Antenna (LISA, Colpi et al. 2024), planned to be launched in 2035, will detect GWs at 0.1−100 mHz, probing MBHBs in the $10^4−10^7 \, M_\odot$ range. On the other hand, Pulsar Timing Array (PTA) experiments seek the nHz GWs produced by cosmologically nearby ($z < 1$) and slowly inspiraling more massive ($>10^8 \, M_\odot$) MBHBs. Currently, five main PTA experiments are taking data: the European Pulsar Timing Array (EPTA, Kramer & Champion 2013; Desvignes et al. 2016), the North American Nanohertz Observatory for Gravitational Waves (NANOGrav, McLaughlin 2013; Arzoumanian et al. 2015), Parkes Pulsar Timing Array (PPTA, Manchester et al. 2013; Reardon et al. 2016), the Indian Pulsar Timing Array (InPTA, Joshi et al. 2022) and Chinese Pulsar Timing Array (CPTA, Lee 2016) projects. The recent results reported by these collaborations provide evidence of a stochastic GW background (sGWB) with amplitude A that ranges between $[1.7−3.2] \times 10^{-15}$, at a reference frequency $f = 1 \, \text{yr}^{-1}$ (Agazie et al. 2023; EPTA Collaboration 2023; Reardon et al. 2023; Xu et al. 2023). Different theoretical studies have been carried out to interpret the nature of the sGWB. Despite the signal being compatible with a population of MBHBs (EPTA Collaboration & InPTA Collaboration 2024; Agazie et al. 2023),

modern, sophisticated hydrodynamical simulations and semi-analytical models, aiming at reproducing a wide array of cosmological observations, tend to produce smaller sGWBs, with typical amplitudes of $A \approx 1 \times 10^{-15}$ (see e.g. Sesana et al. 2009; Kelley et al. 2017a,b; Bonetti et al. 2018a; Izquierdo-Villalba et al. 2022; Curyło & Bulik 2024; Saeedzadeh et al. 2024; Li et al. 2024). Specifically, EPTA Collaboration & InPTA Collaboration (2024) showed that state-of-the-art semi-analytical models require important changes to reach the current sGWB reported by PTA collaborations. For instance, a fast dynamical evolution for MBHBs, a rapid and larger mass growth of MBHs, and a bigger normalization in the scaling relations are fundamental requirements in semi-analytical models to reconcile theoretical predictions and current observational constraints. However, these requisites are not easy to reach unless breaking current observational constraints of the MBH population, such as the black hole mass function or quasar luminosity function (see e.g. Izquierdo-Villalba et al. 2022; Sato-Polito et al. 2023).

Due to its high amplitude, the signal reported by the PTA collaborations is already providing valuable information about the possible formation and evolution of $>10^8 \, M_\odot$ MBHBs. In this paper, we investigate whether this signal has also implications for LISA. In particular, we study how the forecasts for LISA MBHBs with a potential observable electromagnetic counterpart are affected when a galaxy formation model is tuned to reproduce both the latest results on the nHz sGWB and the electromagnetic emission of MBHs. To this end, we make use of the L-Galaxies semi-analytical model which is a unique framework that includes at the same level of detail the physics involved in the assembly of galaxies, MBHs, and MBHBs (Henriques et al. 2015, 2020; Yates et al. 2021; Izquierdo-Villalba et al. 2020, 2022; Spinoso et al. 2023). Under the assumption that the whole PTA signal is coming from a low-z population of MBHBs, the merger trees of the large Millennium simulations (Springel 2005) are used to explore which conditions and model modifications are required in order to match the theoretical predictions about MBHs and MBHBs with current PTA and quasars/AGN electromagnetic constraints. The effect of the new modeling in the LISA MBHBs is explored by using the Millennium-II merger trees (Boylan-Kolchin et al. 2006) whose resolution is adequate to model smaller galaxies and MBHs down to $\approx 10^{3-4} \, M_\odot$.

The paper is organized as follows: in Sect. 2 we summarise the main physics included in L-Galaxies to trace the formation and evolution of galaxies, MBHs, and MBHBs. Furthermore, we show the difficulties of current galaxy formation models to reach the high level of nHz sGWB without mismatching the number density of active MBHs. In Sect. 3 we introduce a model able to generate an sGWB compatible with the latest PTA results, and to reduce the tension seen in the quasar bolometric luminosity function. In Sect. 4 we explore the effect of the nHz signal on the population of LISA MBHBs. In Sect. 5 we underline several caveats of the model. Finally, in Sect. 6 we summarise the main results of the paper. A Λ cold dark matter (ΛCDM) cosmology with parameters $\Omega_m = 0.315$, $\Omega_\Lambda = 0.685$, $\Omega_b = 0.045$, $\sigma_8 = 0.9$ and $h = H_0/100 = 67.3/100 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$ is adopted throughout the paper (Planck Collaboration XVI 2014).

2. L-Galaxies semi-analytical model

In this section, we present the semi-analytical model named L-Galaxies (SAM, Guo et al. 2011; Henriques et al. 2015, 2020; Yates et al. 2021). In brief, L-Galaxies is a code that tracks the cosmological assembly of galaxies through a set of analytical equations solved along the assembly history of dark matter halos, as given by their respective merger tree. On top of this, the latest modifications presented in Izquierdo-Villalba et al. (2020, 2022), and Spinoso et al. (2023) enable L-Galaxies to trace the formation and evolution of single and binary MBHs. We stress that all the physics described here, together with the default values of the free parameters, constitutes the SAM fiducial model tagged as Quiet model (see Table 1).

2.1. Dark matter

L-Galaxies is a flexible semi-analytical model run on top of different dark matter (DM) merger trees extracted from N-body DM-only simulations. In particular, its performance has been tested in the Millennium and TNG-DARK suit of simulations (see e.g. Henriques et al. 2015; Ayromlou et al. 2021). In this work, we use the merger trees extracted from the Millennium (MS, Springel 2005) and Millennium-II (MSII, Boylan-Kolchin et al. 2009) simulations whose minimum particle mass and large cosmological volumes allow to trace the assembly of halos in a broad mass range ($10^8 - 10^{14} \, M_\odot$). MS follows the cosmological evolution of 2160^3 DM particles of mass $8.6 \times 10^8 \, M_\odot \, h^{-1}$ inside a periodic box of $500 \, \mathrm{Mpc} \, h^{-1}$ on a side, from $z = 127$ to the present. MSII can be thought of as a high-resolution version of the MS, as it follows the same number of particles with a mass resolution 125 times higher ($6.885 \times 10^6 \, M_\odot \, h^{-1}$) in a box 125 times smaller ($100 \, \mathrm{Mpc} \, h^{-1}$). MS and MSII were stored at 63 and 68 epochs or snapshots, respectively. All the structures formed in these simulations were found by applying friend-of-fiend and SUBFIND algorithms and arranged with the L-HALOTREE code in the so-called merger trees (Springel et al. 2001). Given the coarse time resolution offered by the outputs of MS/MSII (snapshots are separated by $\approx 300 \, \mathrm{Myr}$), L-Galaxies performs an internal time interpolation of $5-50 \, \mathrm{Myr}$ (depending on redshift) to improve the tracing of the baryonic physics involved in galaxy evolution. Finally, both simulations, originally run with the WMAP1 and 2dFGRS concordance cosmology, were re-scaled with the procedure of Angulo & White (2010) to match the cosmological parameters provided by *Planck* first-year data (Planck Collaboration XVI 2014).

2.2. Gas and stars

L-Galaxies includes a sophisticated galaxy formation model, which is able to track the evolution of the gas and stellar components of structures forming along the DM merger trees. As soon as a DM halo collapses, the model assigns to it an amount of baryons consistent with the cosmological baryonic fraction (White & Frenk 1991). These baryons are initially distributed in a quasi-static hot gas atmosphere which is able to cool down at a rate that depends on the redshift and mass of the hosting DM halo (Guo et al. 2011). The gas that is cooled falls at the centre of the DM halo leading to the formation of a gas disc due to angular momentum conservation. Star formation events occur as soon as the gas disc exceeds a critical mass, giving rise to a stellar component distributed in a disc (Croton et al. 2006). As a consequence of star formation, massive and short-lived stars explode as supernovae injecting energy and metals into the cold gas disc, reheating it, and eventually pushing it beyond the virial radius of the DM halo (Guo et al. 2011). The cold gas component of massive galaxies is also regulated by the radio-mode feedback of the central MBH, which efficiently reduces or even suppresses

cooling flows (Croton 2006). Galaxies can also trigger star formation events through interaction with companion satellites. These events are divided into major and minor mergers. The first ones take place between galaxies with baryonic masses differing by less than a factor of 2 and the final result is the transformation of the remnant galaxy into a pure bulge. On the other hand, minor mergers occur during more extreme mass ratios and they are able to trigger the formation of bulges without destroying the stellar disc of the most massive galaxy. In addition to these two merger treatments, the model includes the prescription of smooth accretion presented by Izquierdo-Villalba et al. (2019) to deal with the physics of extreme-minor mergers. During these events it is expected that the stellar remnant of the satellite gets diluted inside the disc of the central galaxy before being able to reach the nucleus, thus, losing the possibility of making the bulge of the primary galaxy grow. Indeed, Izquierdo-Villalba et al. (2019) showed that the inclusion of these processes is important to recover the observed morphology of dwarf galaxies ($M_{\rm stellar} \leq 10^9 \, M_\odot$) in L-Galaxies when this one is run on top of the MSII simulation. Besides mergers, massive disc-dominated galaxies are allowed to develop galactic bulges through disc instability (Efstathiou et al. 1982). Finally, L-Galaxies models large-scale effects such as ram pressure stripping or galaxy tidal disruption (Henriques & Thomas 2010; Guo et al. 2011).

2.3. Massive black holes

In this section, we briefly describe the main physics included in L-Galaxies to deal with MBHs.

2.3.1. Formation of massive black holes

L-Galaxies includes a refined physical model to follow the genesis of MBHs (Spinoso et al. 2023). Specifically, it tracks the spatial variations of metals and Lyman Werner radiation to account for the formation of massive seeds (10^3–$10^5 \, M_\odot$) via the direct collapse of pristine massive gas clouds and the collapse of dense, nuclear stellar clusters originated by early star-formation episodes. On the other hand, the formation of light seeds (10–$100 \, M_\odot$) after the explosion of the first generation of stars (also known as PopIII stars) is also accounted for by using a subgrid approach which takes as an input the results of GQD Press–Schechter based SAM (Valiante et al. 2021). Notice that this seeding model is only used in the MSII trees since their high halo resolution allows us to track self-consistently the genesis of the first MBHs. On the contrary, the mass resolution of MS merger trees hinders the possibility of using the seeding model presented in Spinoso et al. (2023). Therefore, when employing the MS simulation, we assume that all the newly resolved halos (which have mass of $\sim 10^{10} \, M_\odot$), regardless of redshift, host central MBHs with a fixed mass of $10^4 \, M_\odot$. We have checked that the specific value of the initial MBH mass in the MS merger trees has a marginal effect in the sGWB at nHz frequencies. The cause is the fact that PTA MBHBs are placed in massive galaxies that undergo an intense merger history, feeding MBHs with large amounts of gas that erase any memory of the initial seed mass in a few hundred Myr.

2.3.2. Growth of massive black holes in the Quiet model

As soon as the MBH forms, different processes can trigger its growth. In particular, in L-Galaxies the MBH growth is divided into three different channels: cold gas accretion, hot gas accretion, and mergers with other MBHs. Among these, the first

one is the main driver of black hole growth at any redshift and is triggered by galaxy mergers and disc instability (DI) events. On the one hand, after a galaxy merger, a fraction of cold gas is accreted by the nuclear black hole:

$$\Delta M_{\rm BH}^{\rm gas} = f_{\rm BH}^{\rm merger}(1 + z_{\rm merger})^\alpha \frac{m_{\rm R}}{1 + (V_{\rm BH}/V_{200})^2} \, M_{\rm gas}, \qquad (1)$$

where $m_{\rm R} \leq 1$ is the baryonic ratio of the two interacting galaxies, V_{200} the virial velocity of the host DM subhalo, $z_{\rm merger}$ the redshift of the galaxy merger, $M_{\rm gas}$ the cold gas mass of the galaxy. $f_{\rm BH}^{\rm merger}$, $V_{\rm BH}$ and α are three adjustable parameters set to $280 \, {\rm km \, s^{-1}}$, 0.02 and 5/2, respectively, in what we refer hereafter as Quiet model (see Table 1). While $f_{\rm BH}^{\rm merger}$ and $V_{\rm BH}$ characterize the efficiency of mergers in making gas lose angular momentum and flow towards the galactic nucleus, α takes into account the fact that at high-z galaxies are more compact (Mo et al. 1998; Shen et al. 2003; van der Wel et al. 2014; Lange et al. 2015) and thus, any high-z merger event should be more efficient in bringing gas onto the MBHs compared to low-z events (see Bonoli et al. 2009; Izquierdo-Villalba et al. 2020, 2022, for similar approaches).

On the other hand, during a disc instability, the black hole accretes an amount of cold gas proportional to the mass of stars that trigger the stellar disc instability (see Izquierdo-Villalba et al. 2020):

$$\Delta M_{\rm BH}^{\rm gas} = \max \left[f_{\rm BH}^{\rm DI}(1 + z_{\rm DI})^\alpha \frac{\Delta M_{\rm stars}^{\rm DI}}{1 + (V_{\rm BH}/V_{200})^2}, \, M_{\rm gas} \right], \qquad (2)$$

where $z_{\rm DI}$ is the redshift at which the disc instability takes place and $f_{\rm BH}^{\rm DI}$ is a free parameter that takes into account the gas accretion efficiency, set to 0.0015. $V_{\rm BH}$ and α have the same value as in the case of mergers. As described in Eq. (1), all the free parameters involved in Eq. (2) try to capture the efficiency of disc instabilities in feeding with cold gas the nuclear parts of the galaxy.

After a galaxy merger or a disc instability has occurred at a time t_0, the cold gas available for accretion ($\Delta M_{\rm BH}^{\rm gas}$) is assumed to settle in a reservoir around the black hole, $M_{\rm Res}$[1]. Instead of instantaneous gas consumption, the model considers that the gas reservoir is progressively consumed through an Eddington-limited growth phase, followed by a second phase of low accretion rates (Hopkins et al. 2005, 2006; Marulli et al. 2006; Bonoli et al. 2009). To characterize these two stages we introduce in L-Galaxies the parameter $f_{\rm Edd}$[2], defined as the

---

[1] Notice that $M_{\rm Res} = \Delta M_{\rm BH}^{\rm gas}$ is only satisfied if before the galaxy merger or disc instability the reservoir around the MBH was empty. On the contrary, $M_{\rm Res} = \Delta M_{\rm BH}^{\rm gas} + M_{\rm gas}^{\rm left-over}$ being $M_{\rm gas}^{\rm left-over}$ the leftover gas inside the reservoir, accumulated trough prior mergers or disc instabilities and not consumed by the MBH by the time at which the new merger or disc instability takes place.

[2] Given the value of $f_{\rm Edd}$, the mass of the MBH ($M_{\rm BH}$) and the radiative and accretion efficiency (η and ϵ, respectively) at a time t, the subsequent growth of a MBH (δt) is expressed as $M_{\rm BH}(t + \delta t) = M_{\rm BH}(t) \, e^{f_{\rm Edd} \frac{1-\eta(t)}{\epsilon(t)} \frac{\delta t}{t_{\rm Edd}}}$ where $t_{\rm Edd} = 0.45 \, {\rm Gyr}$. Notice that η accounts for the fraction of rest mass energy released by accretion (which depends on the MBH spin (a), see e.g. Fig. 5 of King et al. 2008), and $\epsilon \leq \eta$ accounts for the fact that not all of the available energy is necessarily radiated (which depends on the accretion disc geometry). Following Merloni & Heinz (2008), L-Galaxies assumes that at $f_{\rm Edd} > 0.03$ (thin disc regime) $\epsilon = \eta(a)$ whereas at $f_{\rm Edd} \leq 0.03$ (advected dominated accretion regime) $\epsilon = \eta(a) f_{\rm Edd}/0.03$ (see Izquierdo-Villalba et al. 2020).

Table 1. Set of values assigned to the free parameters in every model.

Model	$f_{\rm BH}^{\rm merger}$	$f_{\rm BH}^{\rm DI}$	A (at 1 yr^{-1})
Quiet	0.020	0.0015	$1.2_{0.8}^{1.5} \times 10^{-15}$
Boosted	0.045	0.0015	$2.0_{1.2}^{2.8} \times 10^{-15}$
Loud	0.042	0.0015	$1.8_{1.2}^{2.3} \times 10^{-15}$

Notes. The errors on the sGWB amplitude have been computed by dividing the `Millennium` box into sub-boxes of 100 Mpc side-length. Then, it was computed with all of them the 16th and 84th percentile of A at 1 yr^{-1}.

ratio between the bolometric ($L_{\rm bol}$) and the Eddington luminosity ($L_{\rm Edd}$):

$$f_{\rm Edd}(t) = \begin{cases} 1 & M_{\rm BH}(t) \le M_{\rm E} \\ \frac{1}{\left[1+((t-t_0)/t_Q)^{1/2}\right]^{2/\beta}} & M_{\rm BH}(t) > M_{\rm E}. \end{cases} \quad (3)$$

The duration of the Eddington limited phase is determined by the value of $M_{\rm E} = M_{\rm BH}(t_0) + \mathcal{F}_{\rm Edd}M_{\rm Res}(t_0)$, which is the mass reached by the MBH after consuming a fraction $\mathcal{F}_{\rm Edd}$ of its gas reservoir. Following Izquierdo-Villalba et al. (2020) the value of $\mathcal{F}_{\rm Edd}$ is set to 0.7. Note that if a galaxy undergoes a new merger or DI while the central MBH is still accreting mass from a previous event, the new cold gas driven around the MBH environment is added to the previous remnant $M_{\rm Res}$ and the growth re-starts under the new initial conditions. Finally, t_Q gives the time-scale at which $f_{\rm Edd}$ decreases and is defined as $t_Q = t_{\rm d}\, \xi^\beta/(\beta \ln 10)$, with $t_{\rm d} = 1.26 \times 10^8$ yr, $\beta = 0.4$ and $\xi = 0.3$. The choice of these values is based on Hopkins & Hernquist (2009) who showed that models of self-regulated MBH growth require $0.3 < \beta < 0.8$ and $0.2 < \xi < 0.4$. We highlight that any change in the values of β and ξ in the interval suggested by Hopkins & Hernquist (2009) has a small effect on our results since the bulk of the MBH growth happens during the Eddington-limited phase. Finally, during any of the events that make the MBH grow, L-Galaxies tracks the evolution of the black hole spin (a) in a self-consistent way. During gas accretion events, the model uses the approach presented in Dotti et al. (2013) and Sesana et al. (2014) which links the number of accretion events that spin-up or spin-down the MBH with the degree of coherent motion in the bulge. On the other hand, after an MBHB coalescence the final spin is determined by the expression of Barausse & Rezzolla (2009) where a distinction between wet and dry mergers is done to compute the alignment/anti-alignment between the two MBHs. For further details on the implementation of the spin model inside L-Galaxies, we refer the reader to Izquierdo-Villalba et al. (2020).

In Fig. 1 we compare the observational constraints of the quasar LFs (Hopkins et al. 2007; Aird et al. 2015; Shen et al. 2020) with the predictions of the Quiet model of L-Galaxies. As we can see, the model is in good agreement with the observations, being able to reproduce the redshift evolution of active MBHs from $z = 0.5$ up to $z = 3$.

2.4. Massive black hole binaries

The dynamical evolution of MBHB in L-Galaxies is divided into different stages (Begelman et al. 1980). The first one is called pairing phase and it starts after the merger of the two galaxies (see Sect. 2.2). During this process, the MBH

hosted by the less massive galaxy undergoes dynamical friction which causes its sinking toward the galactic centre of its new host. The process is modeled according to the standard Binney & Tremaine (1987) equation which depends on the mass and orbital circularity of the MBH, the initial position at which the satellite galaxy deposited the MBH after the merger ($\sim$kpc separation), and the velocity dispersion of the remnant galaxy. Notice that the model assumes that the nuclear MBH of the central galaxy is not displaced from the galactic centre after the galactic merger. This is a fair assumption since the dynamical friction time scale is dominated by the lighter MBH making irrelevant the displacement of the central and most massive MBH. Finally, we stress that the secondary MBH might be embedded inside nuclear stellar clusters (NSCs, not accounted in the model yet) leading to larger effective masses during the dynamical fiction than the ones assumed in the MBHB model of L-Galaxies. This is a caveat that will be addressed in future works by using a phenomenological model (see e.g. Polkas et al. 2023) or by including a self-consistent NSC modelling inside L-Galaxies (Hoyer et al., in prep.).

Once the dynamical friction phase ends, the satellite MBH reaches the galactic nucleus of the new galaxy and it binds with the central MBH ($\sim$pc separation) starting the so-called hardening phase and giving rise to a massive black hole binary. While the most massive MBH of the binary is flagged as primary, the lighter one is tagged as secondary. The initial eccentricity of the binary orbit is randomly drawn in the range [0, 0.99], while the initial semi-major axis is set to the scale at which the stellar content of the galaxy (distributed according to a Sérsic model) equals the mass of the secondary MBH. The eccentricity and separation of the MHBH are then evolved self consistently according to the environment in which the system is embedded. In case the gas reservoir around the binary ($M_{\rm Res}$) is larger than its total mass ($M_{\rm Bin}$) the system evolves by interacting with a circumbinary gaseous disc, following the prescription of Dotti et al. (2015). Otherwise, the system is driven by the interaction with stars according to the theory developed by Quinlan & Hernquist (1997), Sesana & Khan (2015). In both cases, GW emission takes over at smaller separations ($\lesssim$mpc), leading the binary to final coalescence. We refer to Izquierdo-Villalba et al. (2022) for a detailed description of the equations used to evolve the MBHB eccentricity and separation. In the case of repeated mergers, a third MBH can reach the nucleus of the remnant before the pre-existing binary completes its evolution. In this case, an MBH triplet forms and the outcome of the triple interaction is modeled according to the tabulated values of Bonetti et al. (2018a).

On top of the dynamical evolution of MBHBs, L-Galaxies allows MBHs in the pairing and hardening phase to increase their masses. Recent hydrodynamical simulations of merging galaxies with central MBHs have shown that the secondary galaxy suffers large perturbations during the pericenter passages around the central one (Callegari et al. 2009, 2011; Capelo et al. 2015; Gabor et al. 2016). Under these circumstances, the black hole of the secondary galaxy experiences accretion enhancements mainly correlated with the galaxy mass ratio. To include these findings, L-Galaxies assumes that right before the galaxy merger, the black hole of the secondary galaxy increases its gas reservoir according to Eq. (1). Once the satellite MBH is deposited in the new galaxy and starts its pairing phase, the gas reservoir is consumed according to the two-phase model described in Sect. 2.3.2. The accretion process onto the pairing MBH lasts until it consumes the total gas reservoir stored before the merger. On the other hand, gas accretion onto MBHBs has been extensively explored by different theoretical stud-
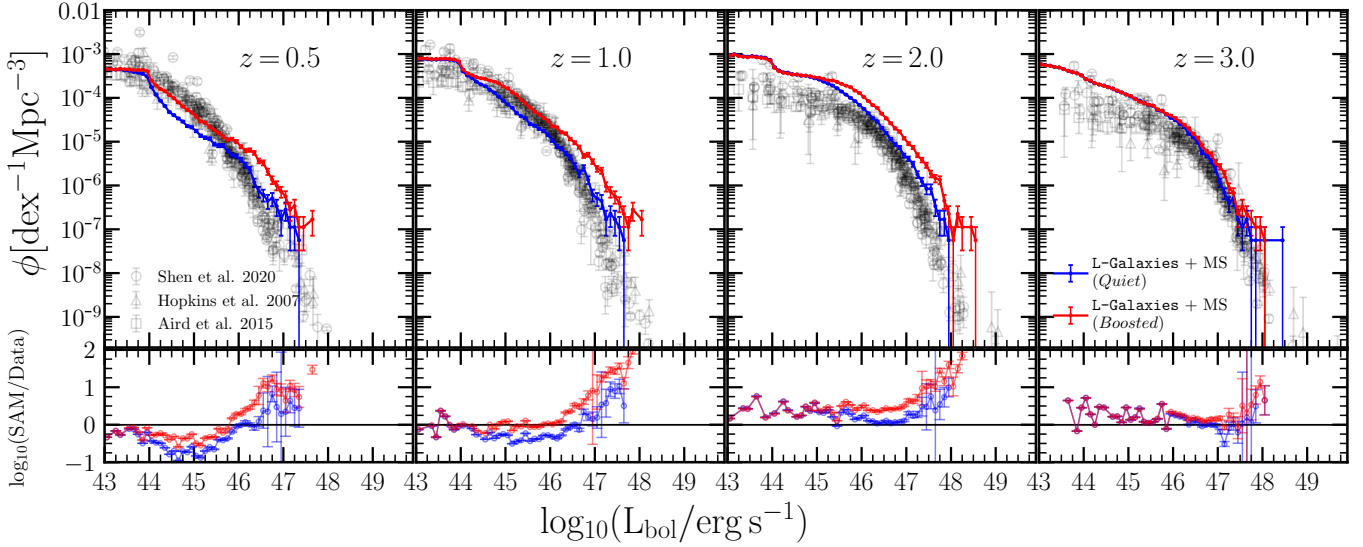
Fig. 1. Quasar luminosity functions at $z = 0.5, 1.0, 2.0, 3.0$ for the Quiet (blue) and Boosted model (red) run in the `Millennium` merger trees (i.e. `L-Galaxies` + MS). The error bars correspond to the Poissonian error. The results are compared with the observations of Hopkins et al. (2007) (triangles), Aird et al. (2015) (squares) and Shen et al. (2020) (circles). The lower panel represents the ratio between the model and the data points.

ies (D'Orazio et al. 2013; Farris et al. 2014; Moody et al. 2019; Muñoz et al. 2019; D'Orazio & Duffell 2021). The findings of these have shown that irrespective of the mass ratio of the binaries, the gas accretion onto the secondary MBH is sufficient to modify the final mass ratio of the binary, moving the initial values toward larger ones (see e.g. Farris et al. 2014; Duffell et al. 2020). Based on this picture, `L-Galaxies` assumes that an MBHB in the hardening phase featuring a gas reservoir progressively consumes it according to the results of Duffell et al. (2020). In brief, the accretion rate of a primary black hole ($\dot{M}_{\rm BH_1}$) is fully determined by the binary mass ratio (q) and the accretion rate of the secondary black hole ($\dot{M}_{\rm BH_2}$) by:

$$\dot{M}_{\rm BH_1} = \dot{M}_{\rm BH_2}(0.1 + 0.9q). \tag{4}$$

Except in the case of equal mass systems, the secondary MBHs are farther away from the binary centre of mass than primary ones. This causes them to be closer to the circumbinary disc edges and thus display high accretion rates. Based on this, `L-Galaxies` fix the accretion of the secondary black hole at the Eddington limit and determine the accretion onto the primary according to Eq. (4).

2.5. Model constraints from Pulsar Timing Arrays

Recent PTA results suggest the existence of an sGWB at nHz frequencies, compatible with a population of merging low-z MBHBs (Agazie et al. 2023; EPTA Collaboration 2023; Reardon et al. 2023; Xu et al. 2023)[3]. Despite the current significance is still below the canonical 5σ threshold, this signal provides galaxy formation models with a new condition to calibrate their underlying MBH growth physics, adding to the standard electromagnetic constraints coming from observations of galaxy properties and the quasar luminosity function.

The foundation to perform a comparison between recent PTA results and galaxy formation models resides in determining the comoving number density of MBHB mergers ($d^2n/dzd\mathcal{M}$)

per unit redshift, z, and rest-frame chirp mass, $\mathcal{M}$[4]. Following Sesana et al. (2008) and making the specific assumption that inspiralling MBHBs in the PTA band are in circular orbits evolving purely due to GW emission, the characteristic sGWB can be written as:

$$h_c^2(f) = \frac{4G^{5/3}f^{-4/3}}{3c^2\pi^{1/3}} \int\int dz d\mathcal{M} \frac{d^2n}{dzd\mathcal{M}} \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}}, \tag{5}$$

where f is the frequency of the GWs in the observer frame. This expression is often simplified as:

$$h_c(f) = A\left(\frac{f}{f_0}\right)^{-2/3}, \tag{6}$$

where A is the amplitude of the signal at the reference frequency f_0. Hereafter, we will set $f_0 = 1\,{\rm yr}^{-1}$ and refer A to that frequency. Under these assumptions, the Quiet model of `L-Galaxies` predicts an $d^2n/dzd\mathcal{M}$ which generates a sGWB of amplitude $A \sim 1.2 \times 10^{-15}$ (see Table 1), fully ruled by $z < 1$ MBHBs with $\mathcal{M} > 10^8\,M_\odot$ (see Izquierdo-Villalba et al. 2023). Despite agreeing with past theoretical studies (Wyithe & Loeb 2003; Sesana et al. 2008, 2009, 2016; Sesana 2013; Ravi et al. 2015; Bonetti et al. 2018b; Kelley et al. 2017a; Siwek et al. 2020) its value is lower than the one reported by PTA collaborations which span over the range $[1.7–3.2] \times 10^{-15}$ (EPTA Collaboration 2023; Agazie et al. 2023; Reardon et al. 2023; Xu et al. 2023). Interestingly, such tension is also present in other recent SAMs with MBHBs such as Li et al. (2024) or Curyło & Bulik (2024).

To reconcile theoretical predictions with the recent PTA results, EPTA Collaboration & InPTA Collaboration (2024) explored changes in the dynamical models of MBHBs included in `L-Galaxies` (e.g. faster dynamical friction phases and only stellar hardening) showing that these modifications are not enough to reduce the discrepancy. Despite this, in a recent paper, Barausse et al. (2023) showed that a Press–Schechter

---

[3] But see alternative origins of the sGWB from early cosmology (e.g. EPTA Collaboration & InPTA Collaboration 2024; Afzal et al. 2023 and references therein).

[4] The chirp mass of an MBHB system is defined as $\mathcal{M} = (M_{\rm BH_1}M_{\rm BH_2})^{3/5}(M_{\rm BH_1} + M_{\rm BH_2})^{-1/5}$, being $M_{\rm BH_1}$ and $M_{\rm BH_2}$ the mass of the primary and secondary MBH, respectively.

based SAM which includes a heavy MBH seeding scenario and assumes no delay between galaxy and MBH mergers would favor large sGWB, compatible with the latest PTA results. Another approach was reported in Izquierdo-Villalba et al. (2022) which proposed that MBHs should be more efficient in accreting cold gas after mergers and/or disc instabilities, thus becoming more massive when they enter the PTA frequency band. To increase the efficiency of MBH growth the authors explored in L-Galaxies a boosted model in which the two-phase growth model of Sect. 2.3.2 was untouched but the gas accretion efficiency, $f_{\rm BH}^{\rm merger}$ in Eq. (1), increases during mergers[5] (hereafter Boosted model, see Table 1). The results showed that the Boosted model was able to generate a stochastic GW background of amplitude $[1.9-2.6] \times 10^{-15}$, compatible with the latest PTA measurements. Despite this better agreement, the increase of the MBH growth was hindering the possibility of reproducing the number density of active MBHs. This is presented in Fig. 1 where we can see that the Quiet model is able to follow the observational trends presented in Hopkins et al. (2007), Aird et al. (2015), and Shen et al. (2020). However, the Boosted one displays a systematic overprediction (up to ~ 2 dex) in the number density of bright quasars ($L_{\rm bol} \gtrsim 10^{46}$ erg s^{-1}) at $z < 2$.

The results presented above suggest that increasing the mass of MBHs is a good avenue to reconcile theoretical models and PTA observations. However, the physical mechanism and the time scale by which the MBHs reach the necessary mass to generate a loud sGWB should be addressed carefully since the over-prediction of the quasar luminosity function seems to be a natural consequence. Taking this into account, in the next section we explore the effect of allowing super-Eddington accretion episodes to grow the population of single MBHs. Specifically, we investigate if these events enable the assembly of a population of MBHB compatible with the nHz sGWB and generate a population of active MBHs in agreement with the quasar luminosity functions. We stress that super-Eddington accretion is only enforced on single MBHs while the growth of MBHBs is modeled in the same way as described in Sect. 2.4.

3. A faster assembly for the MBH population

In this section, we explore the possibility of extending the Quiet model of L-Galaxies in such a way that galaxy mergers are more efficient in fuelling gas onto MBHs and some MBHs, under certain conditions, can undergo super-critical accretion events. While the former requirement is done in the same way as in the Boosted model (see Sect. 2.5), the latter is described in the next section. The interplay between these two processes is calibrated by running L-Galaxies on the Millennium merger trees. Finally, we stress again that super-Eddington accretion events will be allowed only to nuclear single MBHs, being the accretion onto MBHBs modelled in the same way as Sect. 2.4.

3.1. A toy model for super-Eddington growth

Super-Eddington accretion refers to growth episodes that proceed extremely rapidly, breaking the Eddington rate and increasing the MBH mass on time scales shorter than what is allowed by the standard Eddington-limited model (Abramowicz et al. 1988). Recent simulations have shown that not all the environments in which MBHs are embedded can trigger these extreme accretion events. Only dense and dusty gas environ-

ments around the MBHs replenished by large gas inflows after galactic mergers provide the ideal conditions to trigger super-Eddington growth (see e.g. Inayoshi et al. 2016; Takeo et al. 2018; Regan et al. 2019; Toyouchi et al. 2021; Sassano et al. 2023; Massonneau et al. 2023). To account for these requirements in the L-Galaxies SAM, we assume that a super-Eddington phase is active only if: (i) at the moment of the merger/disc instability (t_0) the gas reservoir around the single MBH exceeds by a factor of $\mathcal{R}^{\rm th}$ the MBH mass:

$$\frac{M_{\rm Res}(t_0)}{M_{\rm BH}(t_0)} > \mathcal{R}^{\rm th}, \tag{7}$$

and (ii) the rate at which the gas is infalling towards the galactic centre, $M_{\rm infolw}$, overcomes a certain threshold given by

$$M_{\rm inflow} = \frac{\Delta M_{\rm BH}^{\rm gas}}{t_{\rm dyn}} > M_{\rm inflow}^{\rm th}, \tag{8}$$

where $\Delta M_{\rm BH}^{\rm gas}$ corresponds to the cold gas that fuels the new accretion event onto the nuclear single MBH and it is determined by Eqs. (1) and (2). We set the dynamical time of the gas to inflow towards the galactic nucleus, $t_{\rm dyn}$, as $V_{\rm disc}/R_{\rm gas}^{\rm sl}$ being $V_{\rm disc}$ the maximum circular velocity of the cold gas (assuming an exponential disc profile) and $R_{\rm gas}^{\rm sl}$ the cold gas scale length radius (see Guo et al. 2011, for the detailed model of galactic sizes included in L-Galaxies).

In case both conditions are satisfied, the MBH does not follow the Eddington-limit growth of Eq. (3) but undergoes a super-critical accretion event whose lightcurve is characterized by the following $f_{\rm Edd}$:

$$f_{\rm Edd}(t) = \begin{cases} B(a)\left[\frac{0.985}{\dot{M}_{\rm Edd}/\dot{M}+C(a)} \\ \quad + \frac{0.015}{\dot{M}_{\rm Edd}/\dot{M}+D(a)}\right] & M_{\rm BH}(t) \leq M_{\rm SE} \\ \frac{1}{\left[1+((t-t_0)/t_Q)^{1/2}\right]^{2/\beta}} & M_{\rm BH}(t) > M_{\rm SE}. \end{cases} \tag{9}$$

This two-phase-lightcurve tries to mimic the fact that even in an environment favourable for super-Eddington accretion, the AGN feedback resulting from the gas accretion makes the MBH reach a self-regulated phase within a few Myr[6] (see e.g. Massonneau et al. 2023). The parameter $M_{\rm SE}$ is the maximum mass reached by the MBH during the super-critical accretion and it is defined as $M_{\rm SE} = M_{\rm BH}(t_0) + \mathcal{F}_{\rm SE} M_{\rm Res}(t_0)$, being $M_{\rm BH}(t_0)$ and $M_{\rm Res}(t_0)$ the mass of the MBH and the reservoir at the moment of the (major/minor) merger and/or disc instability (occurring at t_0). $\mathcal{F}_{\rm SE}$ determines the fraction of the gas reservoir consumed by the MBH throughout the super-Eddington phase, before the large energy released during the accretion swaps the gas material around the MBH and hinders a subsequent Eddington limit phase (Lupi et al. 2016; Regan et al. 2019). For simplicity, and to reduce the large number of parameters we fix $\mathcal{F}_{\rm SE} = 0.1$. The functions $B(a)$, $C(a)$ and $D(a)$ are taken from Madau et al. (2014) and they scale with the spin of the MBH (a) as $B(a) = (0.9663 - 0.9292a)^{-0.5639}$, $C(a) = (4.627 - 4.445a)^{-0.5524}$ and $D(a) = (827.3 - 718.1a)^{-0.7060}$. Notice that we do not make any assumption about the spin value of the MBH since it is

---

[5] The conclusions presented here apply also when increasing the gas accretion during disc instabilities.

[6] We have checked that in the model the typical time spent by the MBH in the super-critical accretion is $\sim 70-100$ Myr. Notice that the time resolution of the SAM is ~ 20 Myr. These values align with the results of Pezzulli et al. (2016) and Lupi et al. (2024) which pointed out that super-Eddington accretion events can be sustained over time scales of a few tens of Myr.

computed self-consistently in `L-Galaxies` after any MBHB merger and gas accretion episode (see Izquierdo-Villalba et al. 2020, for further details)[7]. Finally, $\dot{M}$ and $\dot{M}_{Edd}$ are the accretion rate and Eddington accretion rate onto the MBH, respectively. To determine $\dot{M}$ we extract a random number between $[0-10^5]$ distributed according to $\dot{M}^{-1}$. This choice is motivated by recent theoretical works that show a power-law decline in the distribution of simulated populations of accreting MBHs for values above the Eddington-limit (see e.g. Fanidakis et al. 2012; Griffin et al. 2019 or Shirakata et al. 2019)[8].

3.2. Setting up the super-Eddington conditions: The Loud model

The large number of possible galaxy merger histories provided by `L-Galaxies` and the `Millennium` merger trees enable us to explore which are the most common gas environments around MBHs and the typical amount of gas flowing toward the galactic centre after any secular or merger process. To illustrate the interplay between these quantities, Fig. 2 shows the plane M_{inflow} versus $M_{Res}(t_0)/M_{BH}(t_0)$ for the three different events able to trigger the MBH growth: major merger, minor mergers and disc instabilities. These quantities have been computed under the assumption of a more efficient fuelling of gas onto MBHs (see the free parameters used in the third row of Table 1). As shown, at $z > 3$ galaxy interactions can trigger large gas inflows ($>10\,M_{\odot}\,\mathrm{yr}^{-1}$) that bury MBHs in large gas reservoirs of up to $>10^3$ times more massive than the MBH itself. This trend is the result of the fact that interacting high-z galaxies are compact, gas-rich, and host MBHs whose mass is several orders of magnitude smaller than the cold gas component. On the other hand, disc instabilities occurring at high-z do not trigger the large gas inflows shown in mergers, with rates that are typically 2 orders of magnitude smaller ($0.1\,M_{\odot}\,\mathrm{yr}^{-1}$). This points out that in the high-z universe, galactic mergers between gas-rich systems are the unique systems that can sustain significant gas inflows to trigger potential super-critical accretion events onto MBHs. These results align with the findings of other works such as Pezzulli et al. (2016, 2017) and Trinca et al. (2022) who showed, by using a semi-analytical model, that small MBH seeds at high-z can increase several orders of magnitude their masses trough super-critical accretion after gas-rich galactic mergers. Furthermore, similar conclusions have been drawn from the hydrodynamical simulations of Lupi et al. (2024), which pointed out that gas-rich environments at high-z redshift can support long-lasting (tens of Myr) super-Eddington accretion phases, speeding up the growth of intermediate-mass MBHs. Figure 2 also shows that at intermediate redshifts, $1 < z < 3$, the picture is similar to the one seen at higher-z. However, there is an important decrease of events with large inflows and a rise of cases where MBHs are surrounded by small gas reservoirs. This is the consequence of the fact that a large fraction of the galaxy population has already transformed its gas component into stars; as a consequence, mergers between gas-rich systems are less common than at higher redshifts. Specifically, in the case of major mergers, the peak of the M_{inflow} distribution is displaced down to $\sim 0.1\,M_{\odot}\,\mathrm{yr}^{-1}$, about 1.5 orders of magnitude smaller than the higher-z case.

---

[7] Taken into account Madau et al. (2014), the radiative efficiency of the MBH during the super-Eddington accretion is given by $\epsilon = \frac{\dot{m}}{16}A(a)\left(\frac{0.985}{\dot{m}+B(a)} + \frac{0.015}{\dot{m}+C(a)}\right)$, being $\dot{m} = \dot{M}_{Edd}/\dot{M}$.

[8] We have also checked that the results presented here do not change significantly when the accretion rate is extracted according to a distribution following $\dot{M}^{-2}$.
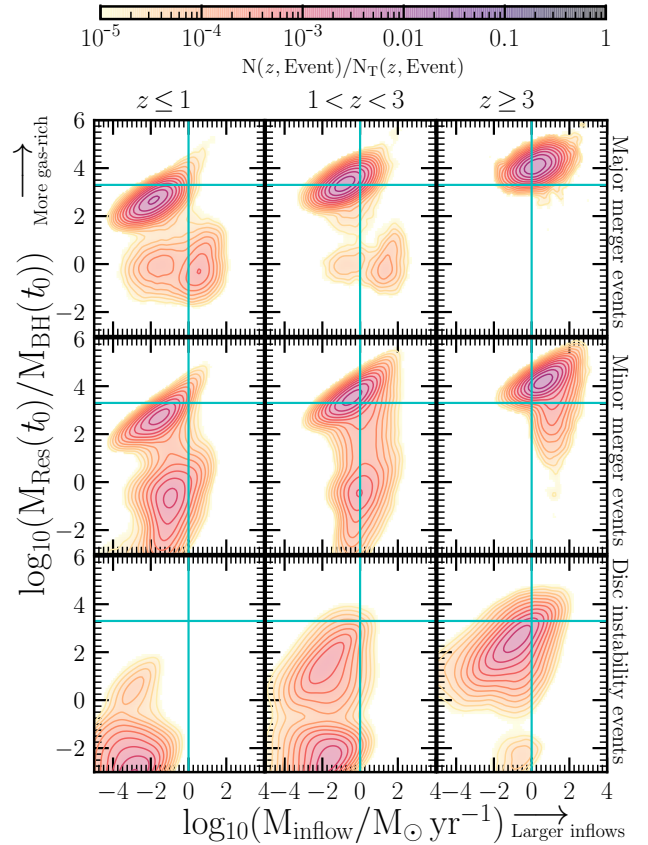
Fig. 2. $M_{inflow}-M_{Res}(t_0)/M_{BH}(t_0)$ plane in three different redshifts bins according to `L-Galaxies` run on the `Millennium` merger trees. While the y-axis depicts how gas-rich is the environment around the MBH, the x-axis illustrates how powerful are the gas inflows towards the MBH. The blue horizontal and vertical lines correspond to the chosen $\mathcal{R}^{th}$, and M_{inflow}^{th}. Each row corresponds to each of the events that can trigger cold gas inflows toward the MBH: major mergers (top), minor mergers (middle) and disc instabilities (bottom). The plot has been computed by using the free parameters of the Boosted model reported in Table 1. To guide the reader, major (minor) mergers correspond to galaxy interactions which involve galaxies with baryonic masses differing by less (more) than a factor of 2.

Finally, mergers and disc instabilities occurring at $z < 1$ are very inefficient in sustaining large inflows towards the galactic nucleus, with rates that rarely exceed $0.01\,M_{\odot}\,\mathrm{yr}^{-1}$. Such low inflows imply that MBHs are systematically embedded in small gas reservoirs whose masses can be two orders of magnitude smaller than the central MBH.
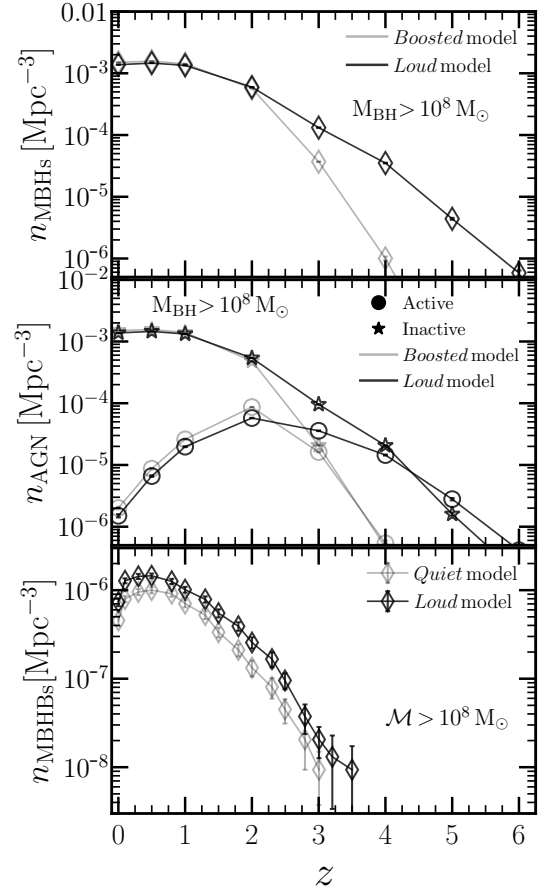
Taking into account the trends presented above and how the LFs evolve with different thresholds in M_{inflow}^{th} and $\mathcal{R}^{th}$ (presented in Appendix A for the sake of brevity), we have chosen $M_{inflow} = 10\,M_{\odot}\,\mathrm{yr}^{-1}$ and $\mathcal{R}^{th} = 2 \times 10^3$ as the best set of parameters to reproduce the evolution of the quasar population. The resulting model is tagged as Loud (see Table 1) and its LFs are presented in Fig. 3. As shown, the match between predictions and observations improves with the new model. Specifically, the Loud model matches the $z > 4$ LFs and predicts at $z \leq 2$ up to a factor of 2 less objects for any bin of luminosity larger than $10^{46}\,\mathrm{erg\,s^{-1}}$, reducing the tension seen in the Boosted case. These improvements are the result of the faster MBH population assembly which consumes most of its gas at high-z and evolves quiescently in the low-z Universe (see Appendix A). This effect can be seen in the upper and middle panel of Fig. 4
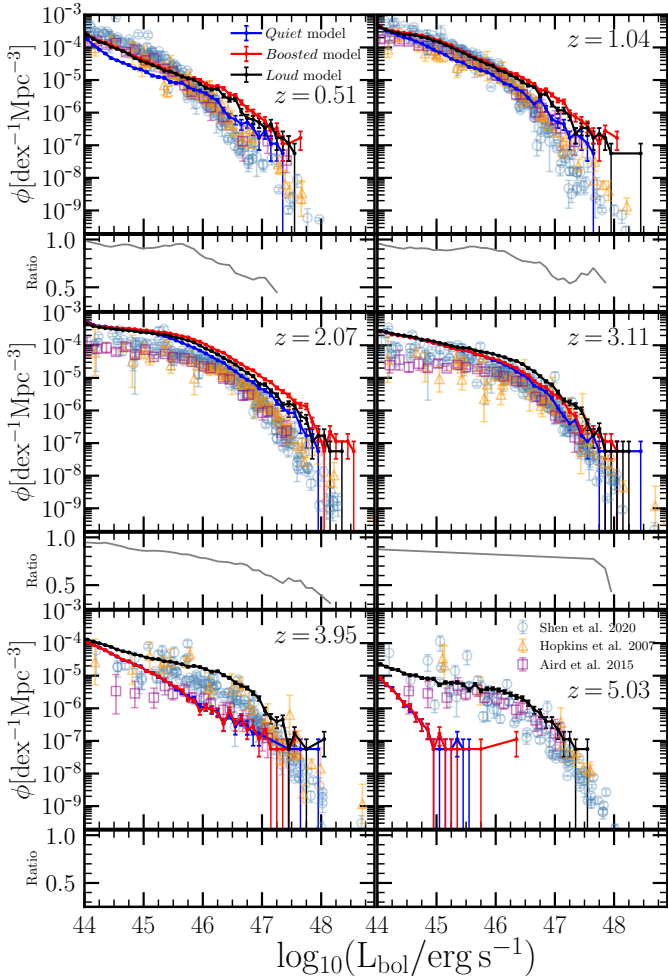
Fig. 3. Evolution of the quasar luminosity function produced by the Loud model (black line), Boosted model (red line) and Quiet model (blue line) when run on the `Millennium` merger trees. The error bars correspond to the Poissionian error. The results are compared with the observations of Hopkins et al. (2007) (orange triangles), Aird et al. (2015) (purple squares) and Shen et al. (2020) (blue circles). The panels below each luminosity function represent the ratio of the luminosity functions predicted by Loud and Boosted model. Notice that the two lower ones do not display any line because their ratio goes beyond the limit 1.

Fig. 4. Number density of active and inactive MBHs/MBHBs. Upper panel: number density of MBHs with mass $>10^8 \, M_\odot$ as a function of redshift, z, for Loud (black) and Boosted model (grey). Middle panel: number density of active ($f_{\rm Edd} > 0.01$) and inactive ($f_{\rm Edd} < 0.01$) MBHs with mass $>10^8 \, M_\odot$ as a function of redshift, z, for Loud (black) and Boosted model (grey). Lower panel: number density of MBHBs with chirp mass $\mathcal{M} > 10^8 \, M_\odot$ as a function of redshift, z, for Loud (black) model. For reference, it has been shown the results for the Quiet model (grey). In all panels, the error bars correspond to the Poisson error. We remind the reader that the comoving volume provided by the `Millennium` simulation corresponds to $\sim 3.6 \times 10^8 \, \rm Mpc^3$.

which presents the number density of MBHs with mass $>10^8 \, M_\odot$ in the Loud and Boosted model (see the global assembly of the MBH population of the Loud model in Appendix B). As shown, the population of $>10^8 \, M_\odot$ is in place much earlier in the former model than in the latter, with number densities that can be up to $\sim$1 dex larger. Regarding the AGN activity of such population, we can see that they are mainly active ($f_{\rm Edd} > 0.01$) at $z > 3$ but they become rapidly inactive ($f_{\rm Edd} < 0.01$) at $z < 2$.

The better agreement between the Loud model and the electromagnetic constraints does not imply a small value of the sGWB amplitude which instead has a value $A = 1.8 \times 10^{-15}$, consistent with the 90% credible interval reported by all the PTA collaborations. The reason why the sGWB is higher in the Loud model with respect to the Quiet one can be seen in the lower panel of Fig. 4, which shows the number density of MBHBs with $\mathcal{M} > 10^8 \, M_\odot$, i.e. the systems which contribute the most to the nHz sGWB signal (see Sesana et al. 2008; Sesana 2013; Izquierdo-Villalba et al. 2022). As we can see, these MBHBs

are in place much earlier in the Loud model than in the Quiet one with number densities up to 2 times larger at any redshift (especially at $z < 1$). Finally, the implications that our new modelling has about the population of super-Eddington sources can be found in Appendix C.

4. Constraining low-redshift LISA MBHB mergers from the nHz sGWB

In this section, we study the implications of the Loud model, featuring a sGWB of amplitude 1.8×10^{-15} at $f = 1 \, \rm yr^{-1}$, on the expected merger rate and electromagnetic counterpart detection of LISA MBHBs. Since we are interested in a population of MBHs up to 4 orders of magnitude smaller than the PTA one, instead of using the `Millennium` merger trees we will make use of the `Millennium-II` ones, which offer the possibility of resolving halos down to $\sim 10^8 \, M_\odot$ (see Sect. 2.1). On top of this, the high-resolution offered by `Millennium-II` enables us to seed MBHs in newly formed galaxies according to the multi-flavour seeding model of Spinoso et al. (2023).

4.1. The detection rate of LISA MBHBs

The redshift versus total rest-frame MBHB mass plane of merging binaries in the Loud model is presented in Fig. 5. As shown, at $M_{Bin} < 10^6 \, M_\odot$ the Loud model displays many mergers occurring at $z > 3$ but the large majority of the events lie at $1 < z < 3$ (see similar results from hydrodynamical simulations in Salcido et al. 2016). For more massive systems, most mergers occur at $z < 1$. To explore how the PTA signal affects the LISA predictions, in Fig. 6 we present the $M_{Bin}-z$ plane where each pixel represents the ratio between the number of MBH mergers predicted by the Loud and Quiet model. While the former displays an sGWB of $A = 1.8 \times 10^{-15}$, compatible with PTA results, the latter produces a signal with $A = 1.2 \times 10^{-15}$ and is in line with a large number of past theoretical works (see e.g. Jaffe & Backer 2003; Sesana et al. 2008; Sesana 2013; Roebber et al. 2016; Kelley et al. 2017a; Barausse et al. 2020). As shown, the Loud model produces $2-5$ times more mergers of MBHBs with total masses $>10^6 \, M_\odot$ than the Quiet one. These differences decrease for $10^5 < M_{Bin} < 10^6 \, M_\odot$ where the ratio varies between $0.9-1.5$. For the lighter systems ($<10^4 \, M_\odot$) the Loud model displays a deficit with respect to the Quiet one. Specifically, the number of MBHB mergers at these masses can be down to 0.5 less frequent at $z < 4$. The decrease of light MBHB mergers in favour of more massive ones is the result of an earlier and faster assembly of the MBH population in the Loud model than in the Quiet one (see Fig. 4). The different MBHB populations generated by the two models point towards different merger rates. These are summarised in Table 2. As shown, the Loud and Quiet model displays a similar global rate of 12.7 yr^{-1}. However, as expected from Fig. 6 the differences are more evident when dividing the population by masses. Specifically, the Loud model predicts merger rates than are ~ 1.2 smaller than the Quiet case when MBHBs of total mass $10^{3-5} \, M_\odot$ are considered. However, it boosts by a factor of 1.5 the coalescences involving systems with total mass $>10^5 \, M_\odot$ ($0.8-4.3$ yr^{-1}, depending on the mass).

Having determined the MBHB merger rate, it is important to calculate how many events are detectable by LISA. To this end, we compute the signal-to-noise ratio (S/N) of all the simulated binaries, defined as:

$$S/N = \left(\int_{f_0}^{f_f} \left[\frac{h_c(f'_{obs})}{h_n(f'_{obs})} \right]^2 \frac{df'_{obs}}{f'_{obs}} \right)^{1/2}, \quad (10)$$

where h_n corresponds to the characteristic strain noise parameterized as in Babak et al. (2021), while h_c represents the characteristic strain amplitude of the source defined as $h_c(f) = 4f^2|\tilde{h}(f)|^2$ being $\tilde{h}(f)$ the Fourier transform of the strain signal, here computed according to the phenomenological frequency-domain gravitational waveform model PhenomC (Santamaría et al. 2010). f_0 represents the starting frequency of inspiralling binaries, set for simplicity to the low-frequency cutoff limit of the LISA sensitivity curve (10^{-4} Hz). Instead, f_f is the maximum frequency of the signal set to $0.15 \, c^3/G(1+z) \, M_{Bin}$[9]. The resolution frequency bin, df, used to integrate Eq. (10) is set to $df = 1/T_{obs}$ where $T_{obs} = 4$ yr corresponds to the length of LISA observations. Finally, the effective spin of the MBHBs is taken from the predictions of L-Galaxies and the eccentricity is set to 0, for simplicity.

---

[9] This specific choice about the initial and final frequencies is made since all events that are going to merge in the LISA band will generally span the whole frequency range from the cut-off up to the merger frequency.
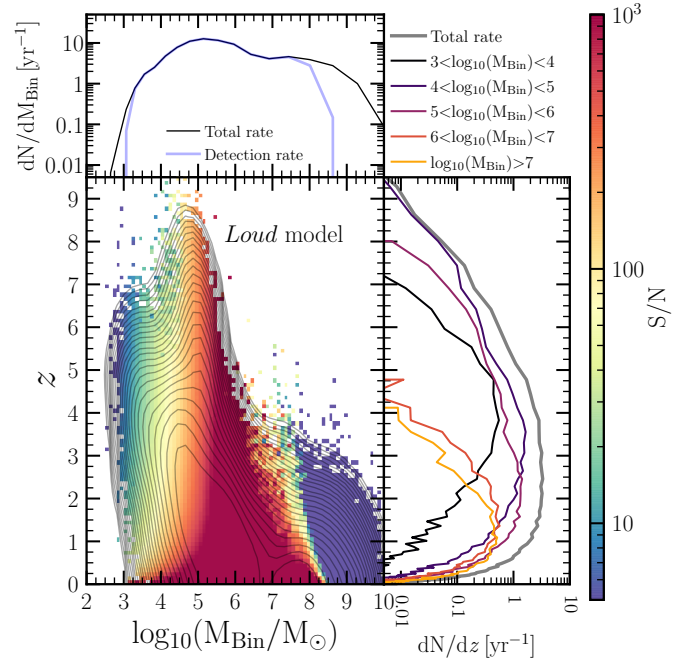
Fig. 5. Distribution of MBHB mergers predicted by the model Loud in the plane redshift (z) versus total rest-frame MBHB mass (M_{Bin}). Each pixel is encoded by the median signal-to-noise ratio (S/N). The contours represent the number of objects in the $z-M_{Bin}$ plane. The small plots around the $z-M_{Bin}$ plane correspond to the redshift (right panel) and source-frame mass (top panel) distribution of the differential number of MBHBs. In the right panel, the total distribution has been divided into 5 different mass bins.
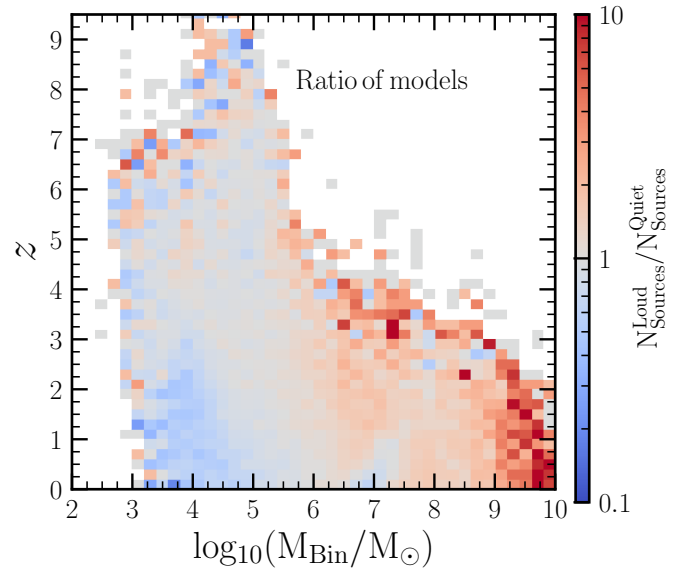
Fig. 6. Distribution of the $z-M_{Bin}$ when each pixel is encoded by the ratio of the number of MBHB mergers in the Loud and Quiet model.

The S/N of the MBHBs generated with the Loud model are depicted in Fig. 5. Each bin of the $M_{Bin}-z$ plane encodes the median S/N value of the MBHBs falling within it. As shown in Fig. 5, MBHBs of $>10^8 \, M_\odot$ have S/Ns that rarely surpass a value of 10. Conversely, MBHB mergers with masses $10^5 < M_{Bin} < 10^7 \, M_\odot$ display large S/Ns, with values that can span between $100 < S/N < 10^4$. For lighter systems ($<10^4 \, M_\odot$) the S/N decreases and the values are systematically <100, being the

smallest for mergers of MBHBs with total mass $<5 \times 10^3 \, M_\odot$ at $z > 2$. Taking into account the S/N distribution of the MBHBs, Table 2 presents the LISA detection rate assuming a minimum signal-to-noise ratio of 10. As we can see, the Loud model predicts a detection rate of 12.24 MBHBs per year, being the systems with $10^4 < M_{BHBs} < 10^5 \, M_\odot$ and $10^5 < M_{Bin} < 10^6 \, M_\odot$ the ones with largest rates (5.10 yr^{-1} and 4.37 yr^{-1}, respectively). As expected, MBHBs with masses $>10^7 \, M_\odot$ are the most affected by the LISA sensitivity curve. In this mass range, LIA can detect only half of the total events (0.4 yr^{-1}). As discussed in Fig. 6, the main difference between the Quiet and Loud models resides in the fact that the latter predicts larger merger events only for MBHBs of $>10^5 \, M_\odot$. Since LISA is not very sensible to MBHBs of mass $>10^7 \, M_\odot$ it implies that the overall detected rate predicted by Quiet and Loud model does not differ much: 12.25 yr^{-1} versus 12.24 yr^{-1}.

The results presented in this section highlight how the recent PTA signal is not significantly informative about the expected global merger rate which will be inferred from LISA observations (see opposite conclusions in Steinle et al. 2023). However, under the assumption that an MBHB population is entirely responsible for the signal observed by PTAs, our detailed galaxy formation model suggests that the loud PTA signal would favour a more numerous population of MBHs and MBHBs of $>10^8 \, M_\odot$ at low-z that usually expected (see Figs. 4 and B.1) which increases the possibilities (1.5 times larger than expected) of LISA to detect such kind of events during its lifetime. However, the physical processes involved in the efficient growth of large MBHs do not act in the same way for the smallest MBHs ($<10^5 \, M_\odot$) which in turn are the preferred targets of LISA. These objects are not able to increase their mass as fast as the most massive MBHs because the small galaxies where they are hosted are not capable of sustaining large and continuous gas inflows towards their centres after galaxy interactions. As a result, the population of low-mass MBHs (and consequently low-mass MBHBs, $<10^5 \, M_\odot$) is only mildly affected by the modifications of MBH growth presented in this work. Thus, small changes are seen in the global population of LISA MBHB mergers (dominated by low-mass MBHs) with respect to the Quiet model.

4.2. Electromagnetic emission of low-z LISA MBHBs

Our results suggest that the expected merger rate detected by LISA is not significantly affected by the specific level of the nHz sGWB, as shown in Table 2. Despite this, Fig. 6 shows that a model in agreement with current constraints on the nHz sGWB predicts that the LISA detection rate of MBHBs above $10^5 \, M_\odot$ at $z < 3$ could be larger compared to a model with a smaller amplitude of the sGWB. The detection and merger rates at $z < 3$ are depicted in the middle panel of Table 2. As shown, for MBHBs of $M_{Bin} > 10^5 \, M_\odot$ Loud model displays up to a factor [1.3−1.5] larger rates than the Quiet one. Since these LISA MBHBs are the ones with the brightest electromagnetic emission, this implies that during the lifetime of the LISA mission, there will be higher chances to detect the electromagnetic emission of MBHBs than previously estimated. To explore the electromagnetic detectability of LISA systems in the Loud model, Fig. 7 presents the cumulative distribution function of the bolometric (and hard X-rays) luminosity of detectable MBHBs with $M_{Bin} > 10^6 \, M_\odot$. The figure shows that these systems have a 50% (20%) probability of shining at $L_{bol} > 10^{43}$ erg s^{-1} ($L_{bol} > 10^{44}$ erg s^{-1}). In X-rays (2−10 keV, see Merloni et al. 2004 for

Table 2. Merger rate predicted by the model (middle column) and detected by LISA ($S/N > 10$, right column) at a different total mass of the binary (M_{Bin}).

	All redshifts	
	Merger rate [yr^{-1}]	LISA detection rate [yr^{-1}]
M_{BHBT} [$M_\odot$]	Loud/Quiet	Loud/Quiet
No mass cut	12.74/12.54	12.28/12.23
$10^3 < M_{Bin} \leq 10^4 \, M_\odot$	1.40/1.56	1.34/1.50
$10^4 < M_{Bin} \leq 10^5 \, M_\odot$	5.12/5.48	5.12/5.48
$10^5 < M_{Bin} \leq 10^6 \, M_\odot$	4.38/4.17	4.38/4.17
$10^6 < M_{Bin} \leq 10^7 \, M_\odot$	1.00/0.76	1.00/0.76
$M_{Bin} > 10^7 \, M_\odot$	0.81/0.54	0.44/0.31
	Low-z universe ($z < 3$)	
	Merger rate [yr^{-1}]	LISA detection rate [yr^{-1}]
M_{BHBT} [$M_\odot$]	Loud/Quiet	Loud/Quiet
No mass cut	6.55/6.45	6.19/6.22
$10^3 < M_{Bin} \leq 10^4 \, M_\odot$	0.27/0.43	0.27/0.43
$10^4 < M_{Bin} \leq 10^5 \, M_\odot$	1.97/2.32	1.97/2.32
$10^5 < M_{Bin} \leq 10^6 \, M_\odot$	2.56/2.44	2.56/2.44
$10^6 < M_{Bin} \leq 10^7 \, M_\odot$	0.93/0.72	0.93/0.72
$M_{Bin} > 10^7 \, M_\odot$	0.77/0.53	0.42/0.31
	Low-z universe ($z < 3$) and $L_{bol} > 10^{43}$ erg s^{-1}	
	Merger rate [yr^{-1}]	LISA detection rate [yr^{-1}]
M_{BHBT} [$M_\odot$]	Loud/Quiet	Loud/Quiet
No mass cut	2.55/2.19	2.41/2.13
$10^3 < M_{Bin} \leq 10^4 \, M_\odot$	–/–	–/–
$10^4 < M_{Bin} \leq 10^5 \, M_\odot$	–/–	–/–
$10^5 < M_{Bin} \leq 10^6 \, M_\odot$	1.52/1.48	1.52/1.48
$10^6 < M_{Bin} \leq 10^7 \, M_\odot$	0.67/0.53	0.67/0.52
$M_{Bin} > 10^7 \, M_\odot$	0.35/0.20	0.22/0.13

Notes. Whereas the upper part of the table depicts all the results without any cut in redshift, the middle and lower parts show the results at $z < 3$. Furthermore, the lower part of the panel displays an extra cut in luminosity ($L_{bol} > 10^{43}$ erg s^{-1}) referring in this way to the low-z active merging MBHBs.

the bolometric correction)[10] we see similar trends, with 50% of the MBHB with $M_{Bin} > 10^5 \, M_\odot$ shining at an X-ray luminosity $>10^{42}$ erg s^{-1}. These prospects imply that future X-ray observatories such as *Athena* (Nandra et al. 2013) or optical surveys such as the *Vera C. Rubin* Observatory (Ivezić et al. 2019) which feature low limiting fluxes will be able to detect the electromagnetic emission coming from several LISA MBHBs.

Having seen that around 50% of the MBHBs detected by LISA will have an observable electromagnetic counterpart, it is interesting to determine the detection rate of these multimessenger MBHBs. To this end, the lower part of Table 2 summarises for the Loud model the detection rate of MBHBs at $z < 3$ featuring a bolometric luminosity $>10^{43}$ erg s^{-1}. The results show that LISA will be able to detect that type of system at a rate of 2.41 yr^{-1} (against the 2.13 yr^{-1} displayed by the Quiet model). When the population is divided into bins of mass, we can see that MBHBs with $10^5 - 10^6 \, M_\odot$ have a detection rate of 1.5 yr^{-1} whereas MBHBs of $10^6 - 10^7 \, M_\odot$ have a rate up a factor 2 smaller. For the case of MBHBs with $M_{bin} > 10^7 \, M_\odot$, the Loud model shows that the detection rate can be up to 0.22 yr^{-1} (a factor 2 larger than in the Quiet model).

5. Caveats

In this section, we discuss several caveats to take into consideration when interpreting the results presented in this work.

---

[10] According to Merloni et al. (2004) the bolometric correction to determine the hard X-ray luminosity of an AGN is given by: $\log_{10}(L_{2-10 \, keV}/L_{bol}) = -1.69 - 0.257\mathcal{L} - 0.0078\mathcal{L}^2 + 0.0018\mathcal{L}^3$, where L_{bol} is the binary bolometric luminosity and $\mathcal{L} = \log_{10}(L_{bol}/L_\odot) - 12$.
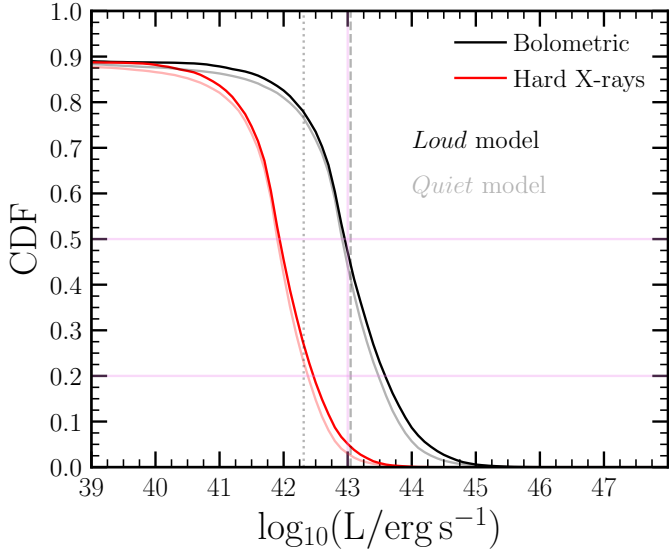
Fig. 7. Cumulative distribution function (CDF) of bolometric (black) and hard X-ray (2–10 KeV, red) luminosity of LISA detectable ($S/N >$ 10) MBHBs with masses $M_{Bin} > 10^5\,M_\odot$. Darker lines correspond to the Loud model whereas the light curves depict the predictions of the Quiet one. Horizontal lines highlight the CFR values of 0.5 and 0.2. The vertical pink line highlights the luminosity value of $10^{43}\,\mathrm{erg\,s^{-1}}$, while black dashed and dotted lines represent the minimum luminosity that a source must have respectively at $z = 1$ and $z = 0.5$ to be detected by *Athena* X-ray observatory assuming a flux limit in the hard X-ray of $2 \times 10^{-15}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ (see Lops et al. 2023).

5.1. Multiple avenues to enlarge the nano-Hz stochastic GWB

In the recent work, EPTA Collaboration & InPTA Collaboration (2024) explored whether state-of-the-art galaxy and MBH formation and evolution models could reproduce the nHz signal reported by the EPTA collaboration. Under the assumption that the whole signal is generated by an MBHB population, the authors showed that under standard assumptions on the MBH and MBHB evolution, SAM models generally predict a signal approximately a factor of two smaller than what is detected. To overcome this limitation, it was shown that a faster dynamical evolution of MBHBs after a galaxy merger and/or a rapid and larger growth of MBHs should be required. In this work, we have explored the latter, by studying the possibility that a quick assembly of the MBH population could be caused by a larger efficiency of galactic mergers in bringing gas towards the galactic nucleus, triggering super-Eddington accretion events onto single MBHs. However, this approach is not the unique avenue to reconcile a high sGWB amplitude with galaxy formation models. For instance, the dynamical models for MBHBs, generally relying on oversimplified assumptions, could be revisited. The changes in the dynamics of MBHBs would imply variations in the population of MBHBs that could enlarge the nHz sGWB without imposing any change in the population of single MBHs. Therefore, different approaches used to reach large nHz sGWB would imply different consequences for low-z LISA MBHBs than those found in this work. In a future paper, we plan to revisit the dynamical and growth model of MBHBs, exploring which of the requirements involved in these processes increase the nHz signal.

Finally, the results presented in this work do not follow the same trends as the ones reported in Barausse et al. (2023).

By using a Press–Schechter based SAM calibrated against the PTA results, the authors showed that LISA forecasts are strongly affected by the underlying PTA signal. Specifically, Barausse et al. (2023) reported that no time delays between galactic and MBH mergers, higher accretion rates onto MBHs, and heavy MBH seeding scenarios would favour large nHz sGWBs with LISA detection rates varying between $3\,\mathrm{yr^{-1}}$ up to $9600\,\mathrm{yr^{-1}}$. In the case of L-Galaxies, a model without delays causes a decrease in the PTA signal as the result of the smaller chirp masses of MBHBs at the time of merger ($h_c \propto \mathcal{M}^{5/3}$, see Fig. 5 of Izquierdo-Villalba et al. 2022): the time spent by large MBHBs in the dynamical friction phase enlarges the mass of the satellite and central MBH, and the gas accretion during the hardening phase tends to make more equal mass systems, increasing in this way the chirp mass of the MBHB at merging time. The discrepancies seen between this work and Barausse et al. (2023) point out that further investigations are required to shed light on how PTA detections would impact the forecast about LISA MBHBs. Specifically, an analysis of why two SAMs that follow very accurately galaxy, MBH and MBHB evolution provide so different results would help in sharpening our knowledge about what is the main physics shaping the population of MBHBs.

5.2. LISA merger rate: Effects of the underlying assumptions

In this work, we have given predictions about the expected MBHB mergers that LISA will detect. However, several assumptions included in L-Galaxies could affect these results. One of them concerns the MBH seeding model. Whereas assumptions in the seed mass do not have any relevant impact in PTA studies, the range of masses and redshifts covered by LISA makes the results sensible to the specific physics assumed about the MBH formation. To have an idea about how the seeding model of L-Galaxies affects the LISA merger rates, we have explored a pessimistic scenario in which the formation of MBHs in L-Galaxies halos is suppressed[11]. In this pessimistic scenario, the LISA merger rate can drop a factor of 3 ($\sim 4\,\mathrm{yr^{-1}}$) with respect to the values presented in this work.

Another caveat to take into account concerns the growth of small MBHs. Recent works pointed out that SN feedback can influence the environments around small MBHs and hinder their growth (see e.g. Habouzit et al. 2017). This scenario, not taken into account by L-Galaxies, could generate a population of lighter MBHs compared to our fiducial runs. Consequently, the dynamical friction phase could be longer and hinder the formation of MBHBs. To take into account this, we have explored a toy model in which the growth of MBHs in L-Galaxies is suppressed as long as the escape velocity of its hosting galaxy is smaller than $270\,\mathrm{km\,s^{-1}}$ (see similar approach presented in Barausse et al. 2020). The results showed that the merger rates dropped down to $9\,\mathrm{yr^{-1}}$ and the peak of the distribution was delayed a few Gyr. Further analysis on the suppression of MBH growth in dwarf galaxies will be explored in Spinoso et al. (in prep.).

Finally, the predicted LISA detection rate could be also affected by the halo mass resolution provided by the merger trees of the Millennium-II simulation. To have a tentative

---

[11] We follow the work of Spinoso et al. (2023) in which they dumped the MBH formation in Millennium-II halos according to a probability, $\mathcal{P} = \mathcal{G}(M_{vir}/10^8\,M_\odot)$. Here we have explored $\mathcal{G} = 0.25$ which implies that a newly resolved halo of $M_{vir} = 10^8\,M_\odot$ ($10^9\,M_\odot$) has 75% (7.5%) less probability of forming a seed with respect to the Quiet/Loud models presented here.

estimation of this, we have used `L-Galaxies` on top of the merger trees of the `TNG50-Dark` and `TNG50-2-Dark` simulations (Nelson et al. 2015)[12]. `TNG50-Dark-2` can be approximated to `Millennium-II` given their similar particle mass resolutions. On the other hand, `TNG50-Dark` can be interpreted as the high-resolution run of `Millennium-II` given its 10 times better mass resolution. To set an upper limit of MBHB mergers lost just by resolution we have compared the MBHB merger rates predicted by `L-Galaxies` on top of `TNG50-Dark` and `TNG50-2-Dark` by assuming that all the halos are seeded with an MBH of $100\,M_\odot$ (maximal MBH occupation) and no delays in the formation of MBHB after the galaxy merger (efficient sinking). Such comparison showed that the main differences were concentrated at $z > 1$ where the MBHB merger rate was a factor 1.1 larger in the `TNG50-Dark` than in `TNG50-2-Dark` (i.e. `Millennium-II` like run). Considering this resolution test, we can conclude that the LISA merger events discussed in this study remain almost unchanged in case the mass resolution of the underlying merger trees is raised by one order of magnitude.

6. Conclusions

In this paper, we explored if the constraints on the nHz sGWB provided by the latest PTA measurements can give valuable information about the population of low-mass MBHBs ($<10^7\,M_\odot$) that will be detectable by the LISA space-based mission. To this end, we made use of the `L-Galaxies` semi-analytical model which runs on top of the `Millennium` suite of simulations and includes detailed physical models to trace galaxy, MBH, and MBHB formation and evolution.

The starting point consisted in creating a population of MBHBs, which produces a nHz sGWB amplitude compatible with the latest PTA results ($A = 1.7-3.2 \times 10^{-15}$). To this end, we followed EPTA Collaboration & InPTA Collaboration (2024) which pointed out that the recent sGWB would imply a faster and larger mass growth of MBHs than usually assumed. However, Izquierdo-Villalba et al. (2023) showed that raising the growth efficiency of MBHs to match a louder sGWB tends to over-predict key electromagnetic constraints such as the quasar bolometric luminosity function or the local MBH mass function. To avoid this shortcoming and taking as reference the fiducial MBH growth model of `L-Galaxies`, we constructed a new framework in which the increase of the galaxy merger efficiency in fuelling gas onto MBHs triggers super-Eddington accretion events. This new model, calibrated by making use of the large volume provided by the `Millennium` simulation, allowed us to create a population of MBHs which generates a sGWB amplitude of 1.8×10^{-15} and reduces the tension with the electromagnetic constraints on the quasar luminosity function.

With this new model, denoted as Loud, we explored the predictions for the LISA MBHB merger rate and compared them against our former fiducial model (Quiet case), which produces a smaller sGWB amplitude. To reach the range of MBHB masses targeted by LISA, we applied our models on top of the `Millennium-II` simulation (i.e. the high-resolution version of `Millennium`) whose merger trees offer the possibility of tracing the cosmological assembly of galaxies and MBHs placed in halos of $[10^7-10^{14}]\,M_\odot$. The main results can be summarised as follows:

- The overall LISA detection rate of MBHBs is not significantly affected by the underlying PTA signal. The Loud model predicts a LISA MBHB detection rate of $12.3\,\mathrm{yr}^{-1}$ whereas the Quiet model forecasts $12.2\,\mathrm{yr}^{-1}$. Therefore, under the assumption of a faster MBH assembly, our results suggest that LISA rates cannot be constrained by using the latest nHz sGWB results.

- The underlying PTA signal causes some differences in the mass distribution of the detected LISA MBHBs. In the Loud model the number of $10^{3-5}\,M_\odot$ detectable MBHBs decreases by a factor of 1.2 with respect to what is predicted by the Quiet model. Conversely, the number of coalescence of $10^{5-7}\,M_\odot$ MBHBs is boosted by a factor of 1.5.

- The increase of detectable merging MBHBs with masses $>10^5\,M_\odot$ found in the Loud model implies better prospects for multimessenger astronomy. Specifically, the model predicts that MBHBs of $10^{5-7}\,M_\odot$ potentially detected by LISA ($S/N > 10$) have 50% (10%) probability of displaying an electromagnetic emission with $L_{\mathrm{bol}} > 10^{43}\,\mathrm{erg\,s}^{-1}$ ($L_{\mathrm{bol}} > 10^{44}\,\mathrm{erg\,s}^{-1}$). Furthermore, the LISA detection rate of such type of systems at $z < 3$ is expected to be $2.4\,\mathrm{yr}^{-1}$.

The results listed above point out that, under the assumption of a faster MBH assembly, the PTA signal cannot constrain the expected LISA merger rate. Indeed, we have shown that a fast MBH growth can only be attained in already-massive systems, where gas is efficiently fuelled onto $>10^6\,M_\odot$ MBHs and prompted to lead super-Eddington accretion episodes. On the contrary, MBHs with mass $<10^6\,M_\odot$ (i.e. the main contributors to LISA events) are not able to increase their mass as fast as the larger MBHs because the small galaxies where they reside are not capable of sustaining massive and continuous gas inflows towards their centres after galaxy interactions. As a result, the population of low-mass MBHs and MBHBs is just mildly affected by an efficient growth model with episodic super-Eddington accretion events. Consequently, the latter model only leads to small changes in the global LISA merger rate, despite producing a 1.5 louder sGWB at nHz frequencies with respect to our fiducial model. Regardless of these small differences, our results show that an efficient mass-growth model induces an increased number of merging systems with $>10^6\,M_\odot$ that can be effectively detected by LISA. Interestingly, MBHs of these masses are also the systems which are more prone to exhibit detectable EM counterparts easily accessible with current and future astronomical facilities. Therefore, our work suggest that, if the astrophysical nature of the PTA signal is confirmed, the possibility of performing multi-messenger analysis with MBHBs could be larger than currently envisioned.

Finally, we stress that in this work we have included a fast assembly of the MBHs to reach the recent sGWB level reported by the PTA collaborations. Nonetheless, different dynamical models for MBHBs could result in a similar enhancement of the nHz signal without invoking any change to the whole MBH population. In an upcoming paper, we will explore this possibility by investigating the conditions leading to a rise in the nHz signal as a consequence of a modified description of the MBHB dynamics.

---

[12] Both of them are run with the same conditions and box size (51.7 Mpc) but their mass resolution differs by one order of magnitude: $m_{\mathrm{p}} = 5.4 \times 10^5\,M_\odot$ for `TNG50-Dark` and $m_{\mathrm{p}} = 4.3 \times 10^6\,M_\odot$ for `TNG50-2-Dark`.

References

Abramowicz, M. A., Czerny, B., Lasota, J. P., & Szuszkiewicz, E. 1988, ApJ, 332, 646

Afzal, A., Agazie, G., Anumarlapudi, A., et al. 2023, ApJ, 951, L11

Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023, ApJ, 951, L8

Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892

Amaro-Seoane, P., Andrews, J., Arca Sedda, M., et al. 2023, Liv. Rev. Relat., 26, 2

Angulo, R. E., & White, S. D. M. 2010, MNRAS, 405, 143

Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2015, ApJ, 810, 150

Ayromlou, M., Nelson, D., Yates, R. M., et al. 2021, MNRAS, 502, 1051

Babak, S., Hewitson, M., & Petiteau, A. 2021, ArXiv e-prints [arXiv:2108.01167]

Barausse, E., & Rezzolla, L. 2009, ApJ, 704, L40

Barausse, E., Dvorkin, I., Tremmel, M., Volonteri, M., & Bonetti, M. 2020, ApJ, 904, 16

Barausse, E., Dey, K., Crisostomi, M., et al. 2023, Phys. Rev. D, 108, 103034

Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307

Biava, S., Colpi, M., Capelo, P. R., et al. 2019, MNRAS, 487, 4985

Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton University Press)

Bonetti, M., Haardt, F., Sesana, A., & Barausse, E. 2018a, MNRAS, 477, 3910

Bonetti, M., Sesana, A., Barausse, E., & Haardt, F. 2018b, MNRAS, 477, 2599

Bonetti, M., Rasskazov, A., Sesana, A., et al. 2020, MNRAS, 493, L114

Bonoli, S., Marulli, F., Springel, V., et al. 2009, MNRAS, 396, 423

Bortolas, E., Bonetti, M., Dotti, M., et al. 2022, MNRAS, 512, 3365

Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2006, MNRAS, 369, 1081

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Callegari, S., Mayer, L., Kazantzidis, S., et al. 2009, ApJ, 696, L89

Callegari, S., Kazantzidis, S., Mayer, L., et al. 2011, ApJ, 729, 85

Capelo, P. R., Volonteri, M., Dotti, M., et al. 2015, MNRAS, 447, 2123

Colpi, M., Danzmann, K., Hewitson, M., et al. 2024, ArXiv e-prints [arXiv:2402.07571]

Croton, D. J. 2006, MNRAS, 369, 1808

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Cuadra, J., Armitage, P. J., Alexander, R. D., & Begelman, M. C. 2009, MNRAS, 393, 1423

Curyło, M., & Bulik, T. 2024, MNRAS, 528, 1053

Desvignes, G., Caballero, R. N., Lentati, L., et al. 2016, MNRAS, 458, 3341

D'Orazio, D. J., & Duffell, P. C. 2021, ApJ, 914, L21

D'Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, MNRAS, 436, 2997

Dotti, M., Colpi, M., Haardt, F., & Mayer, L. 2007, MNRAS, 379, 956

Dotti, M., Colpi, M., Pallini, S., Perego, A., & Volonteri, M. 2013, ApJ, 762, 68

Dotti, M., Merloni, A., & Montuori, C. 2015, MNRAS, 448, 3603

Duffell, P. C., D'Orazio, D., Derdzinski, A., et al. 2020, ApJ, 901, 25

Efstathiou, G., Lake, G., & Negroponte, J. 1982, MNRAS, 199, 1069

EPTA Collaboration (Antoniadis, J., et al.) 2023, A&A, 678, A50

EPTA Collaboration & InPTA Collaboration 2024, A&A, 685, A94

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

Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, MNRAS, 419, 2797

Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, ApJ, 783, 134

Franchini, A., Sesana, A., & Dotti, M. 2021, MNRAS, 507, 1458

Franchini, A., Lupi, A., & Sesana, A. 2022, ApJ, 929, L13

Gabor, J. M., Capelo, P. R., Volonteri, M., et al. 2016, A&A, 592, A62

Griffin, A. J., Lacey, C. G., Gonzalez-Perez, V., et al. 2019, MNRAS, 487, 198

Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101

Habouzit, M., Volonteri, M., & Dubois, Y. 2017, MNRAS, 468, 3935

Henriques, B. M. B., & Thomas, P. A. 2010, MNRAS, 403, 768

Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663

Henriques, B. M. B., Yates, R. M., Fu, J., et al. 2020, MNRAS, 491, 5795

Hopkins, P. F., & Hernquist, L. 2009, ApJ, 698, 1550

Hopkins, P. F., Hernquist, L., Martini, P., et al. 2005, ApJ, 625, L71

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJ, 639, 700

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731

Inayoshi, K., Haiman, Z., & Ostriker, J. P. 2016, MNRAS, 459, 3738

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111

Izquierdo-Villalba, D., Bonoli, S., Spinoso, D., et al. 2019, MNRAS, 488, 609

Izquierdo-Villalba, D., Bonoli, S., Dotti, M., et al. 2020, MNRAS, 495, 4681

Izquierdo-Villalba, D., Sesana, A., Bonoli, S., & Colpi, M. 2022, MNRAS, 509, 3488

Izquierdo-Villalba, D., Sesana, A., & Colpi, M. 2023, MNRAS, 519, 2083

Jaffe, A. H., & Backer, D. C. 2003, ApJ, 583, 616

Joshi, B. C., Gopakumar, A., Pandian, A., et al. 2022, JApA, 43, 98

Kelley, L. Z., Blecha, L., & Hernquist, L. 2017a, MNRAS, 464, 3131

Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2017b, MNRAS, 471, 4508

King, A. R., Pringle, J. E., & Hofmann, J. A. 2008, MNRAS, 385, 1621

Kramer, M., & Champion, D. J. 2013, Class. Quant. Grav., 30, 224009

Lange, R., Driver, S. P., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2603

Lee, K. J. 2016, in Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015, eds. L. Qain, & D. Li, ASP Conf. Ser., 502, 19

Li, K., Bogdanović, T., Ballantyne, D. R., & Bonetti, M. 2022, ApJ, 933, 104

Li, Z., Jiang, Z., Fan, X.-L., et al. 2024, MNRAS, 527, 5616

Lops, G., Izquierdo-Villalba, D., Colpi, M., et al. 2023, MNRAS, 519, 5962

Lupi, A., Haardt, F., Dotti, M., et al. 2016, MNRAS, 456, 2993

Lupi, A., Quadri, G., Volonteri, M., Colpi, M., & Regan, J. A. 2024, A&A, in press, https://doi.org/10.1051/0004-6361/202348788

Madau, P., Haardt, F., & Dotti, M. 2014, ApJ, 784, L38

Manchester, R. N., Hobbs, G., Bailes, M., et al. 2013, PASA, 30, e017

Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169

Marulli, F., Crociani, D., Volonteri, M., Branchini, E., & Moscardini, L. 2006, MNRAS, 368, 1269

Massonneau, W., Volonteri, M., Dubois, Y., & Beckmann, R. S. 2023, A&A, 670, A180

Mayer, L., Kazantzidis, S., Madau, P., et al. 2007, Science, 316, 1874

McLaughlin, M. A. 2013, Class. Quant. Grav., 30, 224008

Merloni, A., & Heinz, S. 2008, MNRAS, 388, 1011

Merloni, A., Rudnick, G., & Di Matteo, T. 2004, MNRAS, 354, L37

Milosavljević, M., & Merritt, D. 2001, ApJ, 563, 34

Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319

Moody, M. S. L., Shi, J.-M., & Stone, J. M. 2019, ApJ, 875, 66

Muñoz, D. J., Miranda, R., & Lai, D. 2019, ApJ, 871, 84

Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints [arXiv:1306.2307]

Nelson, D., Pillepich, A., Genel, S., et al. 2015, Astron. Comput., 13, 12

Pezzulli, E., Valiante, R., & Schneider, R. 2016, MNRAS, 458, 3047

Pezzulli, E., Valiante, R., Orofino, M. C., et al. 2017, MNRAS, 466, 2131

Planck Collaboration XVI. 2014, A&A, 571, A16

Polkas, M., Bonoli, S., Bortolas, E., et al. 2023, A&A, submitted [arXiv:2312.13242]

Quinlan, G. D., & Hernquist, L. 1997, New Astron., 2, 533

Ravi, V., Wyithe, J. S. B., Shannon, R. M., & Hobbs, G. 2015, MNRAS, 447, 2772

Reardon, D. J., Hobbs, G., Coles, W., et al. 2016, MNRAS, 455, 1751

Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, ApJ, 951, L6

Regan, J. A., Downes, T. P., Volonteri, M., et al. 2019, MNRAS, 486, 3892

Roebber, E., Holder, G., Holz, D. E., & Warren, M. 2016, ApJ, 819, 163

Saeedzadeh, V., Mukherjee, S., Babul, A., Tremmel, M., & Quinn, T. R. 2024, MNRAS, 529, 4295

Salcido, J., Bower, R. G., Theuns, T., et al. 2016, MNRAS, 463, 870

Santamaría, L., Ohme, F., Ajith, P., et al. 2010, Phys. Rev. D, 82, 064016

Sassano, F., Capelo, P. R., Mayer, L., Schneider, R., & Valiante, R. 2023, MNRAS, 519, 1837

Sato-Polito, G., Zaldarriaga, M., & Quataert, E. 2023, ArXiv e-prints [arXiv:2312.06756]

Sesana, A. 2013, MNRAS, 433, L1

Sesana, A., & Khan, F. M. 2015, MNRAS, 454, L66

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., Barausse, E., Dotti, M., & Rossi, E. M. 2014, ApJ, 794, 104

Sesana, A., Shankar, F., Bernardi, M., & Sheth, R. K. 2016, MNRAS, 463, L6

Shankar, F., Salucci, P., Granato, G. L., De Zotti, G., & Danese, L. 2004, MNRAS, 354, 1020

Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, ApJ, 690, 20

Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2013, MNRAS, 428, 421

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, MNRAS, 343, 978

Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, MNRAS, 495, 3252

Shirakata, H., Kawaguchi, T., Oogi, T., Okamoto, T., & Nagashima, M. 2019, MNRAS, 487, 409

Siwek, M. S., Kelley, L. Z., & Hernquist, L. 2020, MNRAS, 498, 537

Spinoso, D., Bonoli, S., Valiante, R., Schneider, R., & Izquierdo-Villalba, D. 2023, MNRAS, 518, 4672
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726
Steinle, N., Middleton, H., Moore, C. J., et al. 2023, MNRAS, 525, 2851
Takeo, E., Inayoshi, K., Ohsuga, K., Takahashi, H. R., & Mineshige, S. 2018, MNRAS, 476, 673
Toyouchi, D., Inayoshi, K., Hosokawa, T., & Kuiper, R. 2021, ApJ, 907, 74
Trinca, A., Schneider, R., Valiante, R., et al. 2022, MNRAS, 511, 616
Valiante, R., Colpi, M., Schneider, R., et al. 2021, MNRAS, 500, 4095
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Vasiliev, E., Antonini, F., & Merritt, D. 2014, ApJ, 785, 163
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Wyithe, J. S. B., & Loeb, A. 2003, ApJ, 590, 691
Xu, H., Chen, S., Guo, Y., et al. 2023, Res. Astron. Astrophys., 23, 075024
Yates, R. M., Henriques, B. M. B., Fu, J., et al. 2021, MNRAS, 503, 4474
Yu, Q. 2002, MNRAS, 331, 935

Appendix A: Effect of the thresholds

In Fig. A.1 we explore how different values of $\mathcal{R}^{th}$ and M_{inflow}^{th} affect the number of super-critical accretion events and, thus, the evolution of the luminosity functions. Notice that when varying a threshold we do not impose any limit for the other. In this way, it is possible to marginalize the impact that each parameter. Regarding the effect of $\mathcal{R}^{th}$, we can see that the smaller the value, the larger the number density of $z > 3$ AGNs. This is because a small value of $\mathcal{R}^{th}$ permits a large fraction of MBHs that are embedded in relatively gas-poor environments to trigger super-critical accretion. As a result, the population of $>10^7 M_\odot$ MBHs ($L_{bol} > 10^{45}$ erg/s) is in place earlier in the Universe, having the possibility of triggering more and brighter AGNs than in

a case of a model run with large $\mathcal{R}^{th}$. Interestingly, small values of $\mathcal{R}^{th}$ (e.g., 1) have an opposite effect at low-z, i.e. the normalization of the LFs is smaller than for the cases of large $\mathcal{R}^{th}$ thresholds (e.g., $>10^2$). This smaller number density is the consequence of the faster assembly of MBHs which consumed most of their gas reservoirs at high-z and thus became inactive (or quiescent) at lower redshifts. Regarding the effect of M_{inflow}^{th}, we can see similar trends to the ones shown for $\mathcal{R}^{th}$. Allowing that small inflows fuel super-critical accretion causes a faster assembly of a large fraction of the MBH population, and a rise, at $z > 3$, in the normalization of the LF at any bolometric luminosity. The drawback of the fast MBH assembly with small M_{inflow}^{th} is that at low-z ($z < 2$) the number of AGNs is diminished, a result of the fact that MBHs consumed more of their reservoirs at high-z.
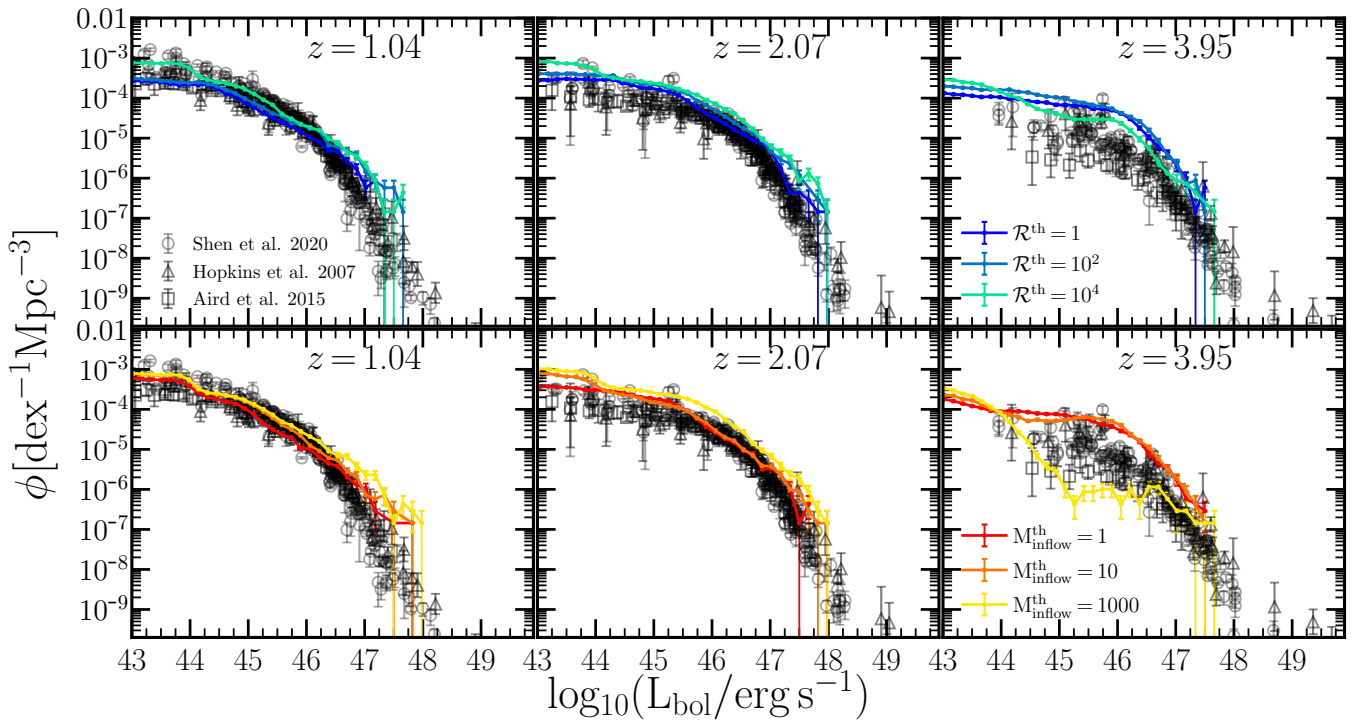
Fig. A.1. AGN luminosity functions at $z = 1.0, 2.0, 4.0$ for the model which includes super-Eddington events. The merger trees used correspond to the ones of `Millennium`. The error bars display the Poissionian error. The results are compared with the observations of Hopkins et al. (2007) (triangles) Aird et al. (2015) (squares) and Shen et al. (2020) (circles). The upper panels correspond to the predictions when varying $\mathcal{R}^{th}$ and not imposing any M_{inflow}^{th} limit. The lower ones represent the LFs of the model when changing M_{inflow}^{th} without any $\mathcal{R}^{th}$ threshold.

Appendix B: The population of massive black holes
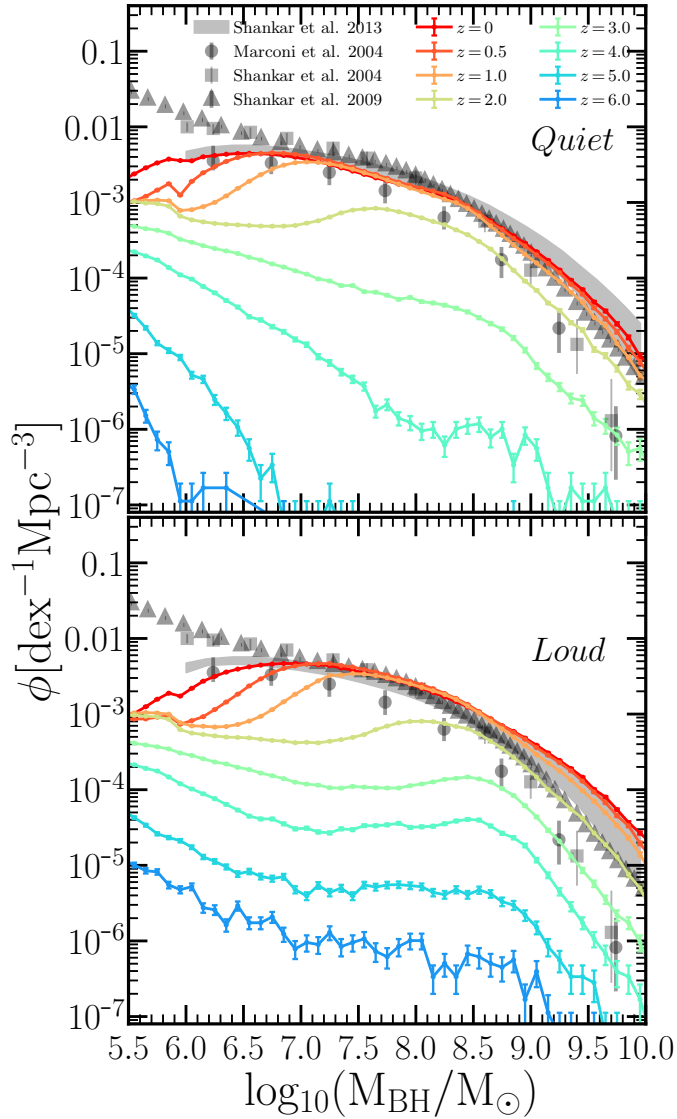
Fig. B.1. Redshift evolution of the black hole mass function compared to the observational results of Marconi et al. (2004), Shankar et al. (2004, 2009) and Shankar et al. (2013) for the Loud model.

In this appendix, we present the population of MBHs generated by `L-Galaxies` and `Millennium` merger trees by making use of the Quiet and Loud model. Fig. B.1 depicts the evolution of the black hole mass function. As we can see, regardless of the model, the predictions at $z = 0$ are consistent with the observational constraints provided by Marconi et al. (2004), Shankar et al. (2004) and Shankar et al. (2013). As described in Izquierdo-Villalba et al. (2020), the mass function displays a fast growth until $z \lesssim 1$ moment at which it slows down and a very small evolution is seen in the massive end ($M_{BH} > 10^7 M_{\odot}$) and the main evolutionary role is taken by the small MBH population ($10^{5.5} < M_{BH} < 10^7 M_{\odot}$). As expected, the biggest difference between Quiet and Loud is that the latter displays a faster assembly of the massive end of the black hole mass function. Specifically, at $z \sim 5-6$ the Loud displays a population of MBHs ($10^{7-8} M_{\odot}$) that is absent in the Quiet case.

Appendix C: Implications for the MBHs: Rare events in particular hosts
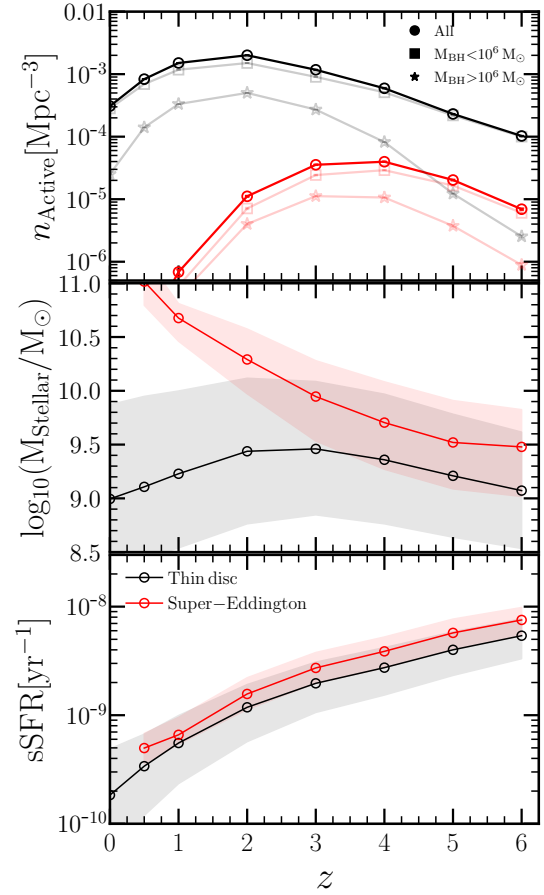
Fig. C.1. Comparison between the population of MBHs accreting during the thin disc mode ($0.03 < f_{Edd} \leq 1.0$) and super-Eddington ($f_{Edd} > 1$). The results correspond to the ones of `L-Galaxies` applied on the `Millennium` simulation. Upper panel: Number density of AGNs accreting at super-Eddington (red) and at thin disc (black) regimen. While circles represent the whole population of AGNs, squares, and stars correspond to AGNs triggered by MBHs with $M_{BH} > 10^6 M_{\odot}$ and $M_{BH} < 10^6 M_{\odot}$, respectively. Middle and bottom panels: Median stellar mass (specific star formation rate, sSFR) of the galaxies hosting AGNs in the super-Eddington (red) and thin disc regimen (black). The shaded area corresponds to the percentile 16th–84th.

Since Loud model has been calibrated using electromagnetic and GW constraints we can make predictions about the expected comoving number density of MBHs undergoing a Super-Eddington phase at different cosmological times. Fig. C.1 shows the results. As we can see, the vast majority of active MBHs are in a thin disc regime ($0.03 < f_{Edd} < 1$), with a peak occurring at $1 < z < 2$ coinciding with the peak of star formation and galaxy mergers. The number density of MBHs undergoing a super-Eddington phase is up to one order of magnitude smaller and its shape displays a peak at $3 < z < 4$. This maximum is followed by a sharp decrease with a number densities below 10^{-5} Mpc^{-3} at $z \sim 2$. In the upper panel of Fig. C.1 the population has been divided between massive ($>10^6 M_{\odot}$) and light ($<10^6 M_{\odot}$) MBHs. Interestingly, only a small fraction of light MBHs undergo a super-critical accretion. This is presumably caused by the fact that the conditions required to sustain super-Eddington episodes are not easily reachable by the small

galaxies where these light MBHs reside. On the contrary, for MBHs of masses $>10^6\,M_\odot$ the requirements for super-Eddington episodes are easier to full-field provoking that the relative difference between AGNs triggered by thin discs and super-critical accretion episodes to be smaller than in the case of light MBHs.

Regarding the hosts of the MBHs undergoing super-Eddington accretion, Fig. C.1 shows the median stellar mass as a function of redshift. As shown, at high-z ($z > 3$) the population harboring super-Eddington MBHs does not display differences with respect to the one in a thin disc regime: the hosts display a stellar mass of $10^{8.5-9.5}\,M_\odot$. For lower redshifts, the galaxies where super-Eddington MBHs are placed are up to 1 dex more massive than the ones of normal AGNs ($10^9\,M_\odot$ versus $10^{10.5}\,M_\odot$). Besides stellar mass, Fig. C.1 shows the specific star formation rate (sSFR) of galaxies hosting super-Eddington and thin disc accretion. Interestingly, the former displays systematically larger values. This is the result of the fact that large galaxy inflows able to trigger super-critical accretion onto MBHs are principally related to gas-rich major mergers which, in turn, are linked with intense bursts of star formation.