THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

The black hole population in low-mass galaxies in large-scale cosmological simulations

Citation for published version:
Haidar, H, Habouzit, M, Volonteri, M, Mezcua, M, Greene, J, Neumayer, N, Angles-Alcazar, D, Martin-Navarro, I, Hoyer, N, Dubois, Y & Dave, R 2022, 'The black hole population in low-mass galaxies in large-scale cosmological simulations', *Monthly Notices of the Royal Astronomical Society* , vol. 514, no. 4, pp. 4912-4931. https://doi.org/10.1093/mnras/stac1659

Digital Object Identifier (DOI):
10.1093/mnras/stac1659

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Monthly Notices of the Royal Astronomical Society

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 27. Oct. 2022

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 16 June 2022 (MN LaTeX style file v2.2)

The black hole population in low-mass galaxies in large-scale cosmological simulations

Houda Haidar[1⋆], Mélanie Habouzit[2,3], Marta Volonteri[1], Mar Mezcua[4,5], Jenny Greene[6], Nadine Neumayer[3], Daniel Anglés-Alcázar[7,8], Ignacio Martin-Navarro[9,10], Nils Hoyer[3], Yohan Dubois[1], Romeel Davé[11]

[1] *Institut d'Astrophysique de Paris, IAP, Sorbonne Universités, CNRS, UMR 7095, 98 bis bd Arago, 75014 Paris, France*
[2] *Zentrum für Astronomie der Universität Heidelberg, ITA, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany*
[3] *Max-Planck-Institut für Astronomie, MPIA, Königstuhl 17, D-69117 Heidelberg, Germany*
[4] *Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Magrans, 08193 Barcelona, Spain*
[5] *Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capità, 08034 Barcelona, Spain*
[6] *Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ, 08544, USA*
[7] *Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA*
[8] *Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*
[9] *Instituto de Astrofísica de Canarias, Vía Láctea s/n, E-38205 La Laguna, Tenerife, Spain*
[10] *Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain*
[11] *Institute for Astronomy, Royal Observatory, University of Edinburgh, Edinburgh EH9 3HJ, UK*

arXiv:2201.09888v2 [astro-ph.GA] 14 Jun 2022

Accepted 2022 June 6. Received 2022 May 24; in original form 2022 January 24

ABSTRACT

Recent systematic searches for massive black holes (BHs) in local dwarf galaxies led to the discovery of a population of faint Active Galactic Nuclei (AGN). We investigate the agreement of the BH and AGN populations in the Illustris, TNG, Horizon-AGN, EAGLE, and SIMBA simulations with current observational constraints in low-mass galaxies. We find that some of these simulations produce BHs that are too massive, and that the BH occupation fraction at $z = 0$ is not inherited from the simulation seeding modeling. The ability of BHs and their host galaxies to power an AGN depends on BH and galaxy subgrid modeling. The fraction of AGN in low-mass galaxies is not used to calibrate the simulations, and thus can be used to differentiate galaxy formation models. AGN fractions at $z = 0$ span two orders of magnitude at fixed galaxy stellar mass in simulations, similarly to observational constraints, but uncertainties and degeneracies affect both observations and simulations. The agreement is difficult to interpret due to differences in the masses of simulated and observed BHs, BH occupation fraction affected by numerical choices, and an unknown fraction of obscured AGN. Our work advocates for more thorough comparisons with observations to improve the modeling of cosmological simulations, and our understanding of BH and galaxy physics in the low-mass regime. The mass of BHs, their ability to efficiently accrete gas, and the AGN fraction in low-mass galaxies have important implications for the build-up of the entire BH and galaxy populations with time.

Key words: black hole physics - galaxies: formation - galaxies: evolution - methods: numerical

1 INTRODUCTION

Over the last decade, compelling evidence has shown that massive black holes (BHs) are commonly found in the center of galaxies in the local Universe, and that their masses correlate with e.g. the stellar mass of their host galaxies or their bulges (e.g., Gültekin et al. 2009). Today, there is a systematic search for the presence of BHs in low-mass galaxies of $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$, with a focus on accreting BHs -commonly referred to as Active Galactic Nuclei (AGN)-, using multi-wavelength observations (Greene, Ho & Barth 2008; Reines & Volonteri 2015; Reines & Comastri 2016; Mezcua et al. 2016; Mezcua 2017; Chilingarian et al. 2018; Greene, Strader & Ho 2019; Reines et al. 2020). The least

⋆ E-mail: houda.haidar@obspm.fr

© 0000 RAS

massive BHs observed are the BH found in RGG118 with a mass of $5 \times 10^4\,\mathrm{M}_\odot$ (Baldassare et al. 2015), and the BH located in NGC4395 (Filippenko & Sargent 1989) with a BH mass estimate in the range of $\sim 9100 - 4 \times 10^5\,\mathrm{M}_\odot$ (Woo et al. 2019). Several techniques are used to detect BHs and AGN in dwarf galaxies, and include optical spectroscopy, observations in X-ray, radio, and even infrared (IR) (e.g., Greene & Ho 2004; Greene, Ho & Barth 2008; Dong et al. 2012; Chilingarian et al. 2018; Liu et al. 2018; Mezcua et al. 2018; Mezcua, Suh & Civano 2019; Kaviraj, Martin & Silk 2019). Multi-wavelength observations are often used to confirm candidates, as for example optical tracers in young starburst regions in low-metallicity environment can resemble those of AGN. Winds from both supernovae (SN) and massive stars are also a source of contamination. Many AGN are detected based on their X-ray emission, but the presence of X-ray binaries (XRBs) can make the interpretation difficult for faint AGN (e.g., Gallo et al. 2008, 2010; Reines et al. 2011; Mezcua et al. 2016; Pardo et al. 2016; She, Ho & Feng 2017; Mezcua et al. 2018; Fornasini et al. 2018; Schirra et al. 2020). In the low-mass regime, the ubiquity of BHs is not unique to the center of galaxies. Reines et al. (2020) discovered that half a dozen of dwarf galaxies were hosting offset radio sources that could be AGN. Faint off-center AGN signatures[1] have been also recently detected in dwarf galaxies with MaNGA integral field unit (IFU) observations (Mezcua & Domínguez Sánchez 2020), confirming that the search for BHs should not be limited to central AGN.

The mechanisms of BH formation are still unknown, but the BH population in local dwarf galaxies provides us with indirect constraints (Volonteri, Lodato & Natarajan 2008a; van Wassenhove et al. 2010). One of the leading mechanisms forms BHs as remnants of the first generation of stars with initial masses of $M_{\rm BH} \sim 100\,\mathrm{M}_\odot$ ("PopIII remnant model" Madau & Rees 2001; Volonteri, Haardt & Madau 2003a; Johnson & Bromm 2007a,b). BHs could also form in slightly metal-enriched stellar compact clusters where runaway stellar collisions can trigger the formation of a very massive star that collapse onto a BH of $M_{\rm BH} \sim 10^2 - 10^4\,\mathrm{M}_\odot$ ("Compact stellar cluster model" Omukai, Schneider & Haiman 2008; Regan & Haehnelt 2009; Alister Seguel et al. 2020). In the "direct collapse" model (e.g., Begelman, Volonteri & Rees 2006a,b), a large amount of gas collapses isothermally without fragmenting and forms a supermassive star collapsing into a BH of $M_{\rm BH} \sim 10^4 - 10^6\,\mathrm{M}_\odot$. This model requires several key conditions such as the absence of molecular hydrogen coolant to avoid gas fragmentation. This can be achieved by the presence of a high Lyman-Werner radiation (11.2-13.6 eV photons) coming from nearby star-forming regions (Dijkstra et al. 2008; Visbal, Haiman & Bryan 2014; Agarwal et al. 2014; Habouzit et al. 2016; Chon, Hosokawa & Omukai 2021), but is not necessary in the case of dynamical heating in the host halo history (Wise et al. 2019).

The PopIII remnant and compact cluster models predict the formation of BHs in great abundance, inhabiting almost -if not all- galaxies in the high-redshift Universe (e.g., Greene 2012; Devecchi & Volonteri 2009). However, if BHs form through less common processes such as in the direct collapse model, only few galaxies could host BHs. While it is currently impossible to constrain the fraction of galaxies hosting a BH at redshifts close to those of BH formation, we can obtain indirect constraints in local low-mass galaxies using *archaeology*. These constraints include the mass of BHs in dwarf galaxies and the BH occupation fraction in these galaxies. Theoretically, the assembly and merger history of low-mass galaxies through cosmic time is predicted to be quieter than for more massive galaxies. First, the mass of BHs in dwarf galaxies could be close to their seeding mass. Second, dwarf galaxies with a possible low number of mergers in their history could provide us with constraints on the intrinsic occupation fraction of the high-redshift Universe.

Constraining the galaxy BH occupation fraction in observations is challenging. The few existing constraints often rely on extrapolations from the AGN occupation fraction. BHs in dwarf galaxies have low masses, and thus are more likely to power faint AGN, which are difficult to identify because of the numerous sources of contamination described above. Current observational constraints indicate that more than 50% of galaxies with $M_\star \leqslant 10^{10}\,\mathrm{M}_\odot$ could host a BH (Miller et al. 2015; Trump et al. 2015; Nguyen et al. 2019), with large uncertainties and differences among these constraints. Assessing the galaxy BH occupation fraction theoretically is also difficult. This is because capturing the physics of BH formation and growth and galaxy evolution in a cosmological context is very demanding and requires high resolution simulations to capture low-mass galaxies and follow their evolution from the redshift of BH formation to the local Universe.

The BH occupation fraction for different BH formation mechanisms was first investigated with semi-analytical models (Volonteri, Haardt & Madau 2003b; Volonteri, Lodato & Natarajan 2008b). Later, Bellovary et al. (2011) used zoom-in cosmological simulations of low-mass halos to derive the number of low-mass galaxies that could host a BH in the local Universe. The occupation fraction was studied with large-scale cosmological simulations in Habouzit, Volonteri & Dubois (2017), from high-redshift down to $z = 2$. These simulations include a physical BH formation model mimicking the PopIII remnant and compact stellar cluster models. The AGN occupation fraction was investigated in the FABLE cosmological simulation (whose BH seeding model does not follow one of the theoretical models) down to $z = 0$ in Koudmani, Henden & Sijacki (2021). Recently, a model for the AGN occupation fraction was derived in Pacucci, Mezcua & Regan (2021), based on gas properties and angular momentum of galaxies and calibrated against the X-ray AGN occupation fractions.

Cosmological simulations are a great tool to study the evolution of massive BHs, as well as their interplay with their host galaxies. The rise of large-scale simulations such as Illustris (Genel et al. 2014; Vogelsberger et al. 2014; Sijacki et al. 2015; Nelson et al. 2015), TNG100, TNG300 (Pillepich et al. 2018b; Springel et al. 2018; Nelson et al. 2018; Naiman et al. 2018; Marinacci et al. 2018), Horizon-AGN (Dubois et al. 2014, 2016; Volonteri et al. 2016), EAGLE (Schaye et al. 2015; Crain et al. 2015; McAlpine et al. 2016) and SIMBA (Davé et al. 2019; Thomas et al. 2019) allows us to numerically bridge the gap between observational and theoretical studies. Such simulations have been

[1] Observational signatures of off-center AGN are in some cases also consistent with light echoes of past AGN activity (Mezcua & Sánchez 2020).

successful at producing galaxies in good agreement with the properties of observed galaxies. Nevertheless, discrepancies among the simulations exist and arise from the use of different galaxy and BH subgrid physics (Habouzit et al. 2021). The formation of BHs in these simulations does not follow the theoretical prescriptions of BH formation mechanisms. Nevertheless, as we will see in this paper, BH activity in low-mass galaxies can provide us with key hints on the interplay between BHs and their host galaxies in this regime. In particular, it can constrain the modeling of seeding, accretion, and feedback processes from SNe and AGN in simulations.

In this work, we analyze the presence of BHs and AGN, and their properties, in the dwarf galaxies of six state-of-the-art cosmological simulations. Throughout this paper, we refer to the low-mass regime as galaxies with $10^9 \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{10.5}$ and dwarf galaxies as $10^9 \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{9.5}$. We describe the simulations in Section 2. We compute the BH occupation fractions produced by the six simulations in Section 3. In Section 4, we derive the AGN occupation fractions of the simulations and assess their agreement with current observational constraints. In Section 5, we examine the properties of the AGN host galaxies. Discussion and conclusion points are given in Section 6 and 7.

2 METHODOLOGY: COSMOLOGICAL SIMULATIONS

Cosmological simulations follow the evolution of the dark matter and baryonic matter contents in an expanding space-time, and allow us to study the evolution of thousands of galaxies with a broad range of total stellar masses from $M_\star = 10^9$ to $10^{12}\,\mathrm{M}_\odot$. In this paper, we analyze the six Illustris, TNG100, TNG300 (larger volume and lower resolution with respect to TNG100), Horizon-AGN, EAGLE, and SIMBA simulations. They all have volumes of $\geqslant 100^3\,\mathrm{cMpc}^3$, dark matter mass resolutions of $\sim 5\times10^6 - 8\times10^7\,\mathrm{M}_\odot$, baryonic resolutions of $10^6 - 2\times10^7\,\mathrm{M}_\odot$, and spatial resolutions of $\sim 1-2$ ckpc. Physical processes taking place at small scales below the galactic scale are modeled as subgrid physics. This includes e.g., star formation, stellar and SN feedback, BH formation, evolution and feedback. Subgrid models vary from simulation to simulation (see Habouzit et al. 2021, for a summary of all these simulations). Detailed descriptions of the simulation's BH modelings can be found in Genel et al. (2014); Vogelsberger et al. (2014) for Illustris, Pillepich et al. (2018b); Weinberger et al. (2018) for TNG, Dubois et al. (2016); Volonteri et al. (2016) for Horizon-AGN, Schaye et al. (2015); Rosas-Guevara et al. (2016); McAlpine et al. (2018, 2017) for EAGLE, and Davé et al. (2019); Thomas et al. (2019, 2021) for SIMBA.

2.1 Modeling of BH physics

The modeling of BH seeding, accretion and feedback are defined differently across simulations. Once the BH seed is planted, it can start growing through feeding on its surrounding gas and BH mergers. In cosmological simulations, BH accretion is often modeled with the Bondi model or a variation of it. For example, the accretion rate onto TNG BHs includes a magnetic field component as the TNG simulations are magneto-hydrodynamical simulations (Pillepich et al. 2018b). Moreover, EAGLE includes a viscous disk component (Rosas-Guevara et al. 2015), while SIMBA incorporates two modes of accretion: the Bondi model for the hot gas component ($T > 10^5$ K) and a gravitational torque limited model for the cold gas ($T < 10^5$ K, Hopkins & Quataert 2011; Anglés-Alcázar et al. 2015, 2017a). All simulations rely on AGN feedback to produce a population of quenched massive galaxies, but each simulation has its own implementation and strength of the feedback. Simulations either use a single-mode feedback as EAGLE (Schaye et al. 2015), or a two-mode feedback for the others (Sijacki et al. 2015; Weinberger et al. 2017; Dubois et al. 2016; Davé et al. 2019). The released energy can be e.g., thermal and isotropic, and/or kinetic with collimated jets, or non-collimated outflows.

In this work, we are particularly interested in BH seed modeling. BHs are seeded in massive halos of $7\times10^{10}\,\mathrm{M}_\odot$ in Illustris, TNG100, and TNG300, and in galaxies of $M_\star \geqslant 10^{9.5}\,\mathrm{M}_\odot$ in SIMBA. Notwithstanding this condition, SIMBA shows a small fraction of galaxies with $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$ harboring a BH. These can be seeded galaxies that lost mass due to interactions and stripping processes. These galaxies can also be the result of approximate mass measurement from the on-the-fly friends-of-friends algorithm used to identify galaxies for the seeding; the mass of these galaxies would be larger than the post-processing masses. We do not consider these galaxies in our analysis. In Horizon-AGN, the location of BH formation is based on the properties of the local gas cells instead of the halo or galaxy mass. Initial BH masses are fixed and comprised in the range $M_{\mathrm{BH}} = 10^4 - 10^6\,\mathrm{M}_\odot$, with $\sim 10^4\,\mathrm{M}_\odot$ for SIMBA, $\sim 10^5\,\mathrm{M}_\odot$ for Illustris, Horizon-AGN, EAGLE, and $\sim 10^6\,\mathrm{M}_\odot$ for TNG100, and TNG300. These simulations do not employ one of the theoretical BH formation models, such as the PopIII remnant or direct collapse models. The initial mass of BHs and the prescriptions for seeding are instead often chosen to reproduce one of the BH to stellar mass empirical scaling relations observed in the local Universe.

For what follows, we only consider galaxies with stellar mass $M_\star \geqslant 10^9\,\mathrm{M}_\odot$. It is also important to mention that we do not distinguish between central and satellite galaxies. Among all the simulations, Horizon-AGN is the only one that forms BHs based on the local properties of the gas. There is an exclusion radius of 50 ckpc preventing the formation of a BH in a galaxy that already hosts one. In Horizon-AGN, galaxies can have multiple BHs as a result of galaxy mergers. When this occurs, we only consider the most massive BH within the galaxies. Two BHs merge if their separation is smaller than the spatial resolution of the simulation (Dubois et al. 2014). In most simulations, BHs are re-positioned to the galaxy potential well at every time step, which means that these simulations can not be used to assess the number of wandering BHs in galaxies.

2.2 AGN luminosity

We follow the model of Churazov et al. (2005) to compute the AGN luminosity. As such, we explicitly differentiate between radiatively efficient and inefficient AGN. For radiatively efficient AGN with Eddington ratios of $f_{\mathrm{Edd}} = \dot{M}_{\mathrm{BH}}/\dot{M}_{\mathrm{Edd}} > 0.1$, the bolometric luminosity can be expressed as $L_{\mathrm{bol}} = 0.1\dot{M}_{\mathrm{BH}}c^2$, where $\epsilon_{\mathrm{r}} = 0.1$ is the radiative efficiency, $\dot{M}_{\mathrm{BH}}$ is the accretion rate onto the BH, and c is the

speed of light. Radiatively inefficient AGN with $f_{\rm Edd} \leqslant 0.1$ have a bolometric luminosity defined as:

$$L_{\rm bol} = 0.1 L_{\rm Edd} (10 f_{\rm Edd})^2 = (10 f_{\rm Edd}) \epsilon_{\rm r} \dot{M}_{\rm BH} c^2. \quad (1)$$

We caution that we lowered the radiative efficiency assumed in the BH accretion and AGN feedback models of the Illustris and TNG simulations. Taking the same radiative efficiency of 0.1 in the post-processing analysis of the simulations is motivated by a better agreement with observational constraints on the faint end of the AGN luminosity function for Illustris and TNGs (Habouzit et al. 2022b). This choice does not affect the conclusions of the present paper. For consistency with previous simulation papers, we compute AGN hard X-ray luminosities by applying the bolometric correction of Hopkins, Richards & Hernquist (2007a). As tested in Habouzit et al. (2022b) the correction derived in Duras et al. (2020), which is very slightly lower than the one derived in Hopkins, Richards & Hernquist (2007a) in the low-mass galaxy regime, does not impact our results. For the purpose of this paper we also tested the correction of Shen et al. (2020), which returned similar results as that of Hopkins, Richards & Hernquist (2007a) for $M_\star \leqslant 10^{10.5}\,{\rm M}_\odot$.

2.3 Post-processing modeling of the X-ray binary population

We follow Schirra et al. (2020) to parameterize the X-ray emission ($2-10\,$keV band) of the XRB population of the simulated galaxies. XRBs consist essentially of a normal star and a compact object (either a neutron star or a BH) orbiting each other. Through this process, the material from the star is drawn towards the collapsed object giving rise to X-ray emission. The binary system makes up two categories depending on the mass of the star: low-mass XRBs (LMXBs) and high-mass XRBs (HMXBs). The former has a lifetime longer than 1 Gyr and their number in a galaxy is a function of the galaxy stellar mass, whereas the latter has a shorter lifetime of $< 100\,$Myr and as such, will be more numerous in star-forming galaxies.

The galaxy-wide XRB luminosity includes the contributions of both LMXBs and HMXBs, as:

$$\frac{L_{\rm XRB}}{\rm erg/s} = \alpha_{\rm LMXB}\,(1+z)^\gamma\,\frac{M_\star}{{\rm M}_\odot} + \beta_{\rm HMXB}\,(1+z)^\delta\,\frac{\rm SFR}{{\rm M}_\odot/{\rm yr}}, \quad (2)$$

with $M_\star$ the galaxy total stellar mass, SFR the galaxy star formation rate, and z the galaxy redshift. We employ the parameters of Lehmer et al. (2019): $\log_{10}\alpha_{\rm LMXB} = 29.15\,{\rm erg/s/M_\odot}$, $\log_{10}\beta_{\rm HMXB} = 39.73\,{\rm erg/s/(M_\odot/yr)}$, $\gamma = 2.0$, and $\delta = 1.3$. We note here that the empirical relations linking the X-ray luminosity of the galaxy XRB population to the galaxy stellar mass and SFR are derived from relatively massive and nearby galaxies. Here we choose to employ the most recent of such relations, but others have been derived from observations (e.g., Lehmer et al. 2010, 2016, 2019; Aird, Coil & Georgakakis 2017; Fornasini et al. 2018). Moreover, the galaxy-wide XRB emission could also depend on the metallicity and stellar ages of the galaxy (e.g., Fragos et al. 2013; Lehmer et al. 2016; Madau & Fragos 2017; Lehmer et al. 2021), but we neglect these possible dependencies. In this paper, we extend this empirical relation to galaxies with stellar mass $M_\star \leqslant 10^{10.5}\,{\rm M}_\odot$.

2.4 Post-processing modeling of the X-ray emission from the galaxy ISM

In the quest for X-ray emission, observers can receive from the sky three types of sources: hot gas from the interstellar medium (ISM), AGN and XRBs. The contributions from the latter two are predicted to be the most dominant in the hard X-ray band. The hot gas is predicted to be significant in the rest-frame energy $< 1.5\,$keV, whereas XRBs dominate the emission over the hot gas for higher energies of $> 1.5\,$keV (e.g., Lehmer et al. 2016). We compute the contribution of the ISM to the galaxy-wide total X-ray luminosity for all the simulations and show that the ISM contribution is indeed minor (see Section C).

We model the X-ray luminosity of the ISM by adopting the empirical relation of Mineo, Gilfanov & Sunyaev (2012) (in the soft 0.5-2 keV band):

$$\frac{L_{\rm ISM}}{\rm erg/s} = 8.3 \times 10^{38} \times \frac{\rm SFR}{{\rm M}_\odot/{\rm yr}}. \quad (3)$$

We then employ a k-correction to move from soft band to hard X-ray band , assuming a photon index of $\Gamma = 3$ (Mezcua et al. 2018). In this paper, we also use this modeling to estimate the AGN contribution from the galaxy-wide X-ray luminosity provided in some observational studies (see Appendix A).

3 RESULTS: BH OCCUPATION FRACTION

3.1 Local Universe

The probability of a given galaxy to host a BH is defined by the BH occupation fraction (OF), which is simply derived as the ratio between galaxies hosting a central massive BH and the total number of galaxies, in bins of galaxy stellar mass. We show the BH OF of all the simulations at $z = 0$ in Fig. 1. The BH OF increases with the galaxy stellar mass for all the simulations. While all of them have an occupation of unity in massive galaxies with $M_\star \geqslant 10^{11}\,{\rm M}_\odot$, they do not reach unity at the same galaxy stellar mass. The BH OF reaches unity at $M_\star = 10^{10.5}{\rm M}_\odot$ in EAGLE, TNG100 and TNG300, at $M_\star = 10^{11.3}{\rm M}_\odot$ in SIMBA and Horizon-AGN and finally, at $M_\star = 10^{11}{\rm M}_\odot$ in Illustris.

All the simulations except Horizon-AGN have a BH occupation higher than 90% for galaxies with $M_\star = 10^{10}\,{\rm M}_\odot$. The difference among simulations extends and increases down to the dwarf galaxy regime. The probability of a simulated galaxy with $M_\star = 10^9\,{\rm M}_\odot$ to host a BH is the lowest in Horizon-AGN with 7%, and ranges from 71% to 86% for the other simulations. The seeding of BHs in galaxies with ${\rm M}_\star \geqslant 10^{9.5}\,{\rm M}_\odot$ in SIMBA does not allow us to investigate the BH occupation in the dwarf regime with ${\rm M}_\star \leqslant 10^{9.5}\,{\rm M}_\odot$(shown as dashed line in Fig. 1). Galaxies of ${\rm M}_\star \geqslant 10^{9.5}\,{\rm M}_\odot$ in SIMBA are, however, reaching an OF of 90%, similar to most of the other simulations. Horizon-AGN, which is the only simulation avoiding a fixed halo/galaxy mass threshold for seeding BHs, has a different behavior: its BH OF is low (10%) in the dwarf regime, and increases progressively to reach 100% only in galaxies with $M_\star \geqslant 10^{11.3}\,{\rm M}_\odot$. To understand this feature, we investigate the redshift evolution of the BH OF in Section 3.3.

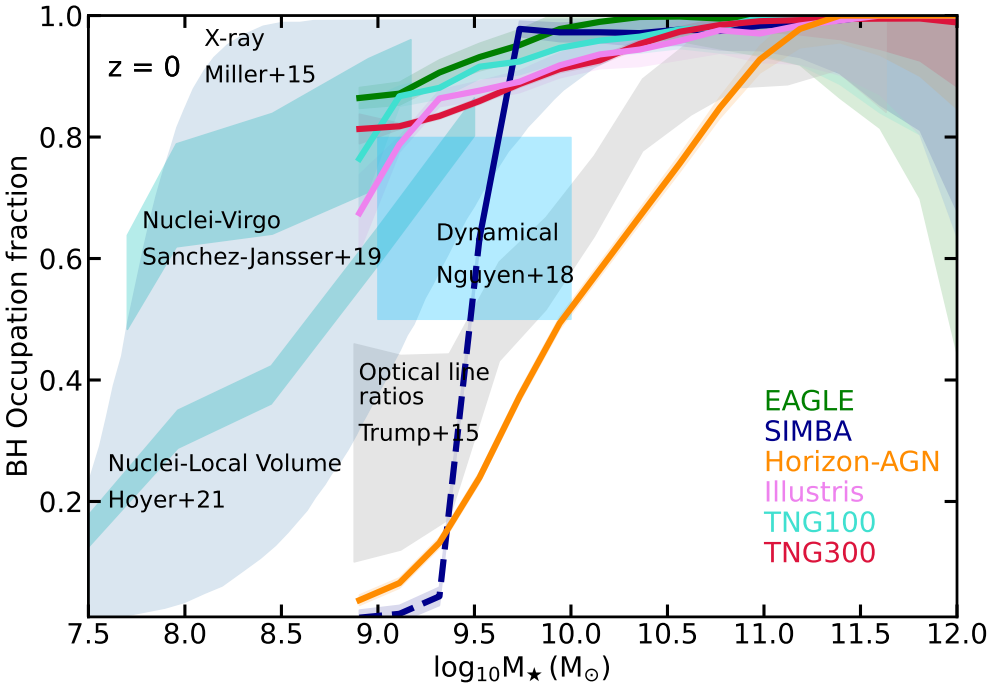

Figure 1. The BH occupation fraction for the local Universe ($z = 0$) predicted by the six simulations. Poisson error bars are shown as shaded regions, and are larger in more massive galaxies due to the presence of fewer galaxies. We show the fraction of SIMBA as a dashed line for $M_\star < 10^{9.5}\,\mathrm{M}_\odot$, as only more massive galaxies are seeded. We include the current observational constraints on the BH occupation fraction of Nguyen et al. (2018) with dynamical mass measurement of BHs in a few nearby nucleated galaxies, and from Miller et al. (2015) and Trump et al. (2015), both derived from a sample of AGN in low-mass galaxies. While not being a constraint on the BH occupation fraction, we also report the galaxy nucleation fraction in the galaxy cluster Virgo (Sánchez-Janssen et al. 2019) and in the Local Volume (Hoyer et al. 2021).

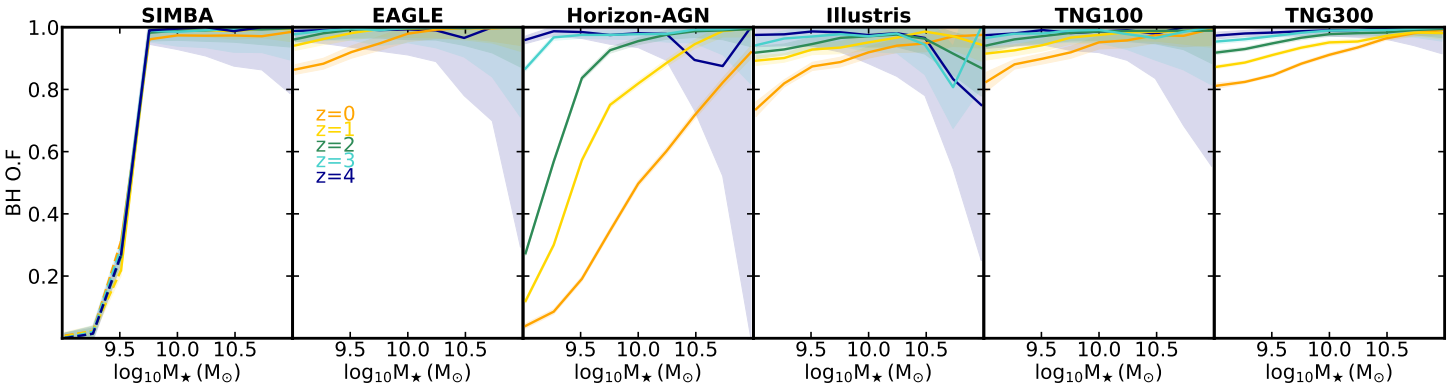

Figure 2. The BH occupation fraction at redshift $z = 0-4$ predicted by the all six simulations, over the range $10^9\,\mathrm{M}_\odot \leqslant M_\star \leqslant 10^{11}\,\mathrm{M}_\odot$ (including both central and satellite galaxies). Shaded areas represent Poisson error bars. All the simulations have a higher BH occupation fraction at higher redshift, at fixed stellar mass. With time, the fraction drops in the low-mass galaxy regime. This is mainly due to a larger number of satellite galaxies that do not host a BH.The decrease is stronger in the Horizon-AGN simulation.

3.2 Comparisons with current observational constraints in the local Universe

In Fig. 1, we show the current observational constraints on the BH OF at $z = 0$. Here, we only discuss the agreement of the simulations with the constraints. The implications and uncertainties are later reviewed in the Discussion section.

One of these constraints comes from BH dynamical mass measurements performed for five of the nearest low-mass (with $M_\star = 5 \times 10^8 - 10^{10}\,\mathrm{M}_\odot$) early-type galaxies M32, NGC 205, NGC 5102, NGC 5206 (Nguyen et al. 2018), and NGC 404 (Nguyen et al. 2017; Davis et al. 2020). While small in number statistics, this sample of five early-type galaxies (ETGs) is volume limited in the considered stellar mass range but only targets nucleated galaxies, and provides a BH occupation of 80% for these galaxies. The OF in ETGs was also investigated in Gallo et al. (2008) with X-ray source detections in 16% of their sample galaxies. The AGN detections were combined to the X-ray luminosity

function to derive a constraint on the BH occupation from their "clean" sample (1σ confidence) for galaxies within the range $M_\star = 10^{7.5} - 10^{12.5}\,\mathrm{M}_\odot$ (Miller et al. 2015).

We also consider the constraints from Trump et al. (2015) (SDSS data), for which the BH OF is fitted to the AGN OF, assuming that the population of BHs in galaxies would be composed of low-mass BHs (possible formed via the PopIII remnant model) and more massive BHs (formed via heavy seed models). These fits as a function of different galaxy quantities such as specific SFR (sSFR), galaxy mass-color, and concentration, indicate that BHs in low-mass galaxies would likely be undermassive in the $M_{\rm BH} - \sigma$ relation. Their BH OF is derived from the high-mass regime of this relation. Their fits from more than 200 000 galaxies indicate that BHs would be present in about 30 to 50% of low-mass galaxies.

All simulations except Horizon-AGN produce a BH OF consistent with the observational constraint derived by Miller et al. (2015), and are above the constraints derived for BH dynamical mass measurements of Nguyen et al. (2017, 2018). Horizon-AGN has a BH OF that is below the constraints of Miller et al. (2015); Nguyen et al. (2017, 2018). However, it is the only simulation to be in agreement with the constraints of Trump et al. (2015).

Finally, we also add the observational constraints of Sánchez-Janssen et al. (2019); Hoyer et al. (2021) (see also Carlsten et al. 2021) on the presence of nuclear star clusters in hundreds of galaxies in the nearby Virgo galaxy cluster and the Local Volume ($d \leqslant 12\,\mathrm{Mpc}$). While not directly constraining the BH OF, there is evidence for the presence of BHs within nuclear star clusters in the local Universe, while there exist also cases of nuclear star clusters devoid of BHs (Neumayer, Seth & Böker 2020). Sánchez-Janssen et al. (2019) investigate the presence of nuclear clusters in hundreds of galaxies in the center of Virgo. The nucleation fraction reaches $\sim 80-90\%$ for galaxies with $M_\star \sim 10^9\,\mathrm{M}_\odot$, and decreases towards less and more massive galaxies. For the Local Volume, Hoyer et al. (2021) find that the nucleation fraction reaches 80% in galaxies with $\sim 10^9\,\mathrm{M}_\odot$, up to slightly more than 80% in galaxies of $\sim 10^{10}\,\mathrm{M}_\odot$. We will discuss these constraints in Section 6.4.

3.3 Evolution of the BH occupation fraction with redshift

Fig. 2 shows that SIMBA achieves an OF of about unity at $\mathrm{M}_\star \geqslant 10^{9.5}\mathrm{M}_\odot$, and stays that way throughout the redshift range $z = 0 - 4$. This is a direct result of the seeding model: all galaxies with $M_\star \geqslant 10^{9.5}\,\mathrm{M}_\odot$ are seeded with a BH at any redshift. The other simulations have BH OF close to unity at high redshift ($z = 4$). These fractions all decrease with decreasing redshift at fixed stellar mass for $M_\star \leqslant 10^{10.5}\,\mathrm{M}_\odot$, although with different amplitudes. Taking $z = 2$ for example, the BH OF at $\mathrm{M}_\star = 10^{9.5}\mathrm{M}_\odot$ in EAGLE and Horizon-AGN are 97% and 78% respectively.

Eagle, Illustris and TNGs all have a halo mass threshold for the seeding of $\sim 10^{10}\,\mathrm{M}_\odot$. BH particles can spawn at any time if the condition for formation is fulfilled. The reason for the decreasing trend with decreasing redshift is that some low-mass satellite galaxies can lose their BHs due to BH repositioning when they are close to their central galaxies. This artificial process is more common at lower redshift, and is responsible for the lower OF at lower redshift.

The decrease of BH OF with decreasing redshift is the most pronounced in Horizon-AGN. This is a combination of several factors. In this simulation, there is an exclusion radius to form new BHs: with time new low-mass galaxies form and if they are close to a large galaxy which already has a BH they will not form their own BHs. Furthermore, no BH particles are formed after $z \sim 1.5$, therefore low-mass galaxies forming after $z \sim 1.5$ do not host a BH (see also Volonteri et al. 2016; Habouzit, Volonteri & Dubois 2017). Finally, seeding in Horizon-AGN takes place in cells whose local gas density is higher than the density to form stars, and as gas density decreases at low redshift, so does the probability of seeding.

3.4 Simulated BHs in the dwarf galaxy regime

While almost all the simulations produce a BH OF of $\geqslant 70\%$ in dwarf galaxies of $10^9\,\mathrm{M}_\odot$, the mass of the BHs that they contain vary from one simulation to another. This is due to a combination of the seeding mass, BH accretion and galaxy modeling (Li et al. 2019; Habouzit et al. 2021).

To investigate the BH population produced by the different simulations, we resort to the $M_{\rm BH} - M_\star$ diagram, shown in Fig. 3. Most simulations show that the $M_{\rm BH}$ scatter is higher in the galaxy low-mass regime ($M_\star \leqslant 10^{10.5}\,\mathrm{M}_\odot$) and decreases with increasing mass of both the BH and the host galaxy ($M_\star > 10^{10.5}\,\mathrm{M}_\odot$, shown in Fig. 2 of Habouzit et al. 2021). Besides having the same $M_{\rm BH}$ scatter trend with stellar mass, the simulations produce different BH mass ranges at fixed stellar mass. In the range $M_\star \leqslant 10^{10}\,\mathrm{M}_\odot$, BHs are on average more massive in Horizon-AGN, TNGs, and SIMBA, than they are in EAGLE and Illustris. These differences produce different shape of the $M_{\rm BH} - M_\star$ diagrams. Some simulations such as Horizon-AGN, Illustris, SIMBA and TNG100 (at low redshift) employ a modeling of SN feedback (either weak, or decoupled winds) that has a relatively weak effect on the growth of BHs. BHs can grow efficiently with their galaxies, resulting in a linear $M_{\rm BH} - M_\star$ relation. However, a simulation such as EAGLE regulates more strongly the growth of BHs with respect to the host galaxies in the low-mass galaxy regime, resulting in a plateau at the low-mass end of the $M_{\rm BH} - M_\star$ diagram. This is due to a combination of stronger SN feedback and on average lower accretion rates due to the EAGLE modified Bondi model (Rosas-Guevara et al. 2016).

To quantify the results presented above, we show in Fig. 4 the mass histograms of BHs embedded in galaxies with $M_\star \sim 10^9, 10^{9.5}, 10^{10}\,\mathrm{M}_\odot$ for all the simulations. The histograms are normalized by the number of BHs in each stellar mass bin. While EAGLE and Illustris have similar BH mass distributions for $\mathrm{M}_\star \sim 10^9\,\mathrm{M}_\odot$, with BHs mostly in the range $\mathrm{M}_{\rm BH} \sim 10^5 - 10^6\,\mathrm{M}_\odot$, Horizon-AGN and TNG have BHs in the range $\mathrm{M}_{\rm BH} \sim 10^6 - 10^7\,\mathrm{M}_\odot$. The Horizon-AGN BHs are able to grow efficiently from their seeding mass due to a relatively weak SN feedback, to produce $10^6 - 10^7\,\mathrm{M}_\odot$ BHs in dwarf galaxies. The high-mass BHs in galaxies of $10^{9-9.5}\,\mathrm{M}_\odot$ produced by some simulations appear to be more massive than the current observational constraints. In observations, BHs in these galaxies do not exceed

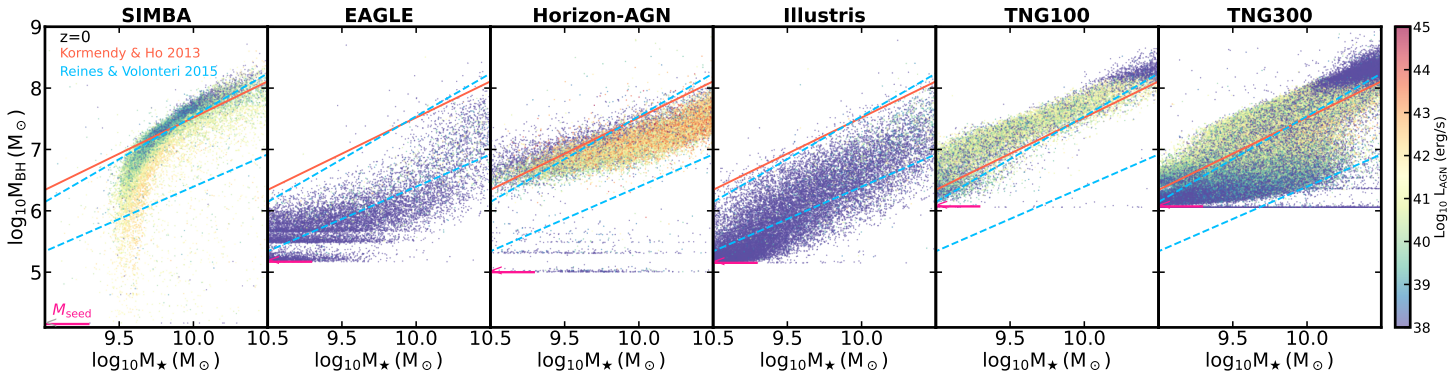

Figure 3. $M_{BH} - M_\star$ diagram for galaxies with $M_\star = 10^9 - 10^{10.5}\,M_\odot$ at $z = 0$, color coded as a function of the AGN hard X-ray luminosity (2-10 keV). The pink arrows point towards the seeding mass of the simulations. We overlay the $M_{BH} - M_\star$ scaling relations from Reines & Volonteri (2015) and Kormendy & Ho (2013) as a reference. The different $M_{BH} - M_\star$ diagrams of the simulations, and the diversity of AGN luminosities at fixed BH mass and galaxy total stellar mass indicate that simulations produce different AGN populations in the low-mass galaxy regime.

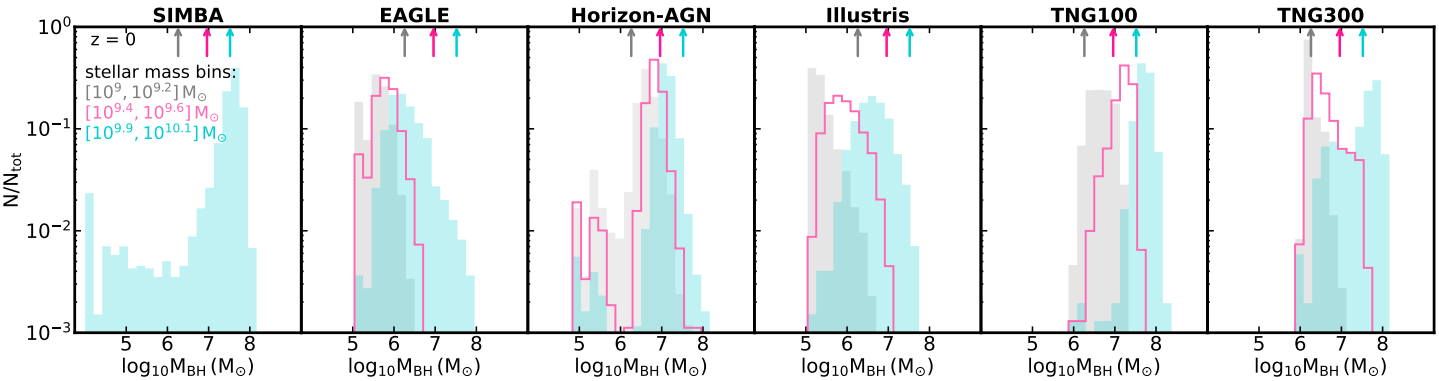

Figure 4. The normalised distribution of BH mass at $z = 0$ in three different stellar mass bins: $10^9 \leqslant M_\star/M_\odot \leqslant 10^{9.2}, 10^{9.4} \leqslant M_\star/M_\odot \leqslant 10^{9.6}$, and $10^{9.9} \leqslant M_\star/M_\odot \leqslant 10^{10.1}$. Arrows point to the mass of the most massive BHs in the observational sample of Reines & Volonteri (2015) derived in each of the three stellar bins. For SIMBA, we do not display the BH population for $M_\star < 10^{9.5}\,M_\odot$ since the simulation starts seeding BHs at the aforementioned value. The mass of BHs in low-mass galaxies strongly depends on the seeding mass in some simulations with strong SN feedback and/or low BH accretion rates, and does not in simulations with weak SN feedback, and/or efficient BH accretion, and/or accretion that does not strongly depend on BH mass. For all the stellar mass bins presented here, BH accretion is dominant over BH mergers.

$M_{BH} \sim 10^{6.5}\,M_\odot$ (Reines & Volonteri 2015; Greene, Strader & Ho 2020; Neumayer, Seth & Böker 2020).

The BH mass distribution in more massive galaxies with $M_\star \sim 10^{10}\,M_\odot$ also varies from one simulation to another. Such galaxies host BHs of $M_{BH} \sim 10^5 - 10^8\,M_\odot$ on average across simulations. In EAGLE and Illustris, BHs of $M_{BH} \sim 10^5 - 10^8\,M_\odot$ are found in these galaxies, with distributions peaking at $M_{BH} \sim 10^{6-7}\,M_\odot$. Fig. 4 shows that the distribution in Horizon-AGN is centered on massive BHs with $M_{BH} \sim 10^7\,M_\odot$, while even more massive BHs with $M_{BH} \sim 10^7 - 10^8\,M_\odot$ are found in SIMBA and TNG100. The BHs in SIMBA are able to grow efficiently from their seeding mass of $M_{BH} \sim 10^4\,M_\odot$ (lowest of all simulations) to be the most massive in galaxies with $M_\star \sim 10^{10}\,M_\odot$, among all the simulations. The $M_{BH} - M_\star$ relation in SIMBA is independent of seed mass, because gravitational torque accretion has a very weak dependence on M_{BH} (Anglés-Alcázar, Özel & Davé 2013; Anglés-Alcázar et al. 2015, 2017a). SIMBA allows BHs in the torque-limited accretion mode to accrete up to 3× the Eddington limit, while the other simulations are capped at Eddington. This can have some effect on the early growth of BHs (Davé et al. 2019).

4 RESULTS: AGN OCCUPATION FRACTION

Using the simulations as a laboratory to study the AGN OF in the low-mass galaxies may aid in revealing details on the BH population and their host galaxies. We define the AGN OF as the ratio between the number of AGN and the total number of galaxies in bins of galaxy stellar mass.

4.1 Local Universe

In Fig. 5 (top panels), we show the AGN OF for three different AGN luminosity ranges in hard X-ray (2-10 keV): $L_{AGN} = 10^{38} - 10^{40}$, $10^{40} - 10^{42}$ and $10^{42} - 10^{44}$ erg/s[2]. The

[2] The AGN luminosity distributions of the simulations at $z = 0$ are shown in Appendix B (Fig. B1).

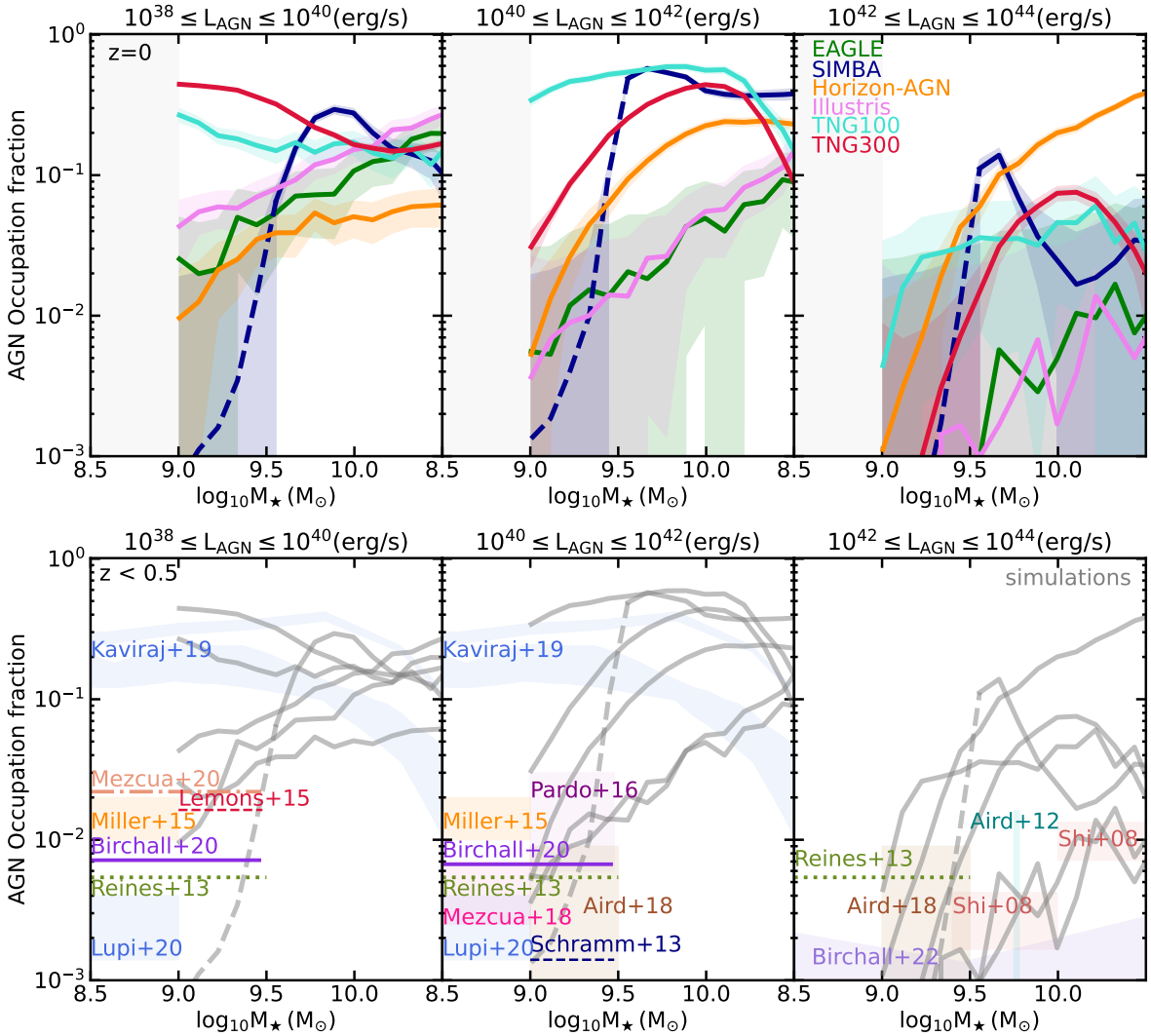

Figure 5. *Top figure*: AGN occupation fraction produced by the simulations in the low-mass regime ($10^9 \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{10.5}$) at $z = 0$. The different panels show three different AGN luminosity ranges in hard X-ray (2-10 keV): $10^{38} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{40}$, $10^{40} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{42}$ and $10^{42} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{44}$. The shaded areas represent Poisson errors. The AGN fractions in the simulations span several orders of magnitude. *Bottom figure*: Current observational constraints on the AGN occupation fraction for low-mass galaxies ($10^{9.5} \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{10.5}$) for $z < 0.5$. We caution here that not all these constraints include a correction to account for incompleteness of the observational samples. We overlay the AGN occupation fraction derived from the simulations, at $z = 0$, in grey.

shaded areas represent the uncertainties assuming a Poisson distribution. For a given simulation, the fraction depends on the number of galaxies hosting a BH (as described in the previous section), and on the ability of galaxies to feed the existing BHs. The BH OF of Horizon-AGN is low in low-mass galaxies, and this is reflected in the AGN fraction.

We first analyze the evolution of the AGN fractions with AGN luminosity. Low-mass galaxies are on average more likely to host a faint AGN with $L_{\rm AGN} = 10^{38} - 10^{42}$ erg/s (two left panels) than a bright one with $L_{\rm AGN} = 10^{42} - 10^{44}$ erg/s (right panel). In general, the presence of faint AGN in dwarf galaxies is due to a combination of the low-mass BHs in these galaxies (Fig. 4), but also of the effect of SN feedback preventing BH growth or the modeling of BH accretion (e.g., in EAGLE). We note a different behavior in TNG100, for which the AGN fractions are higher when considering AGN with $L_{\rm AGN} = 10^{40} - 10^{42}$ erg/s than fainter or brighter AGN. At $z = 0$, the higher fraction of AGN with $L_{\rm AGN} \leqslant 10^{42}$ erg/s in $10^9 - 10^{10}\,\mathrm{M}_\odot$ galaxies in TNG100 compared to the other simulations comes from the combination of the high seeding mass of $\sim 10^6\,\mathrm{M}_\odot$ and a relatively weaker SN feedback[3]. The high seeding mass favors higher accretion rates (if gas is available) in the Bondi formalism because $\dot{M}_{\rm BH} \propto M_{\rm BH}^2$. The BH accretion in SIMBA is governed by a gravitational torque limited model for accretion from cold gas ($T < 10^5$ K), and Bondi model in the presence of hot gas ($T > 10^5$ K, Davé et al. 2019). Relatively low-mass BHs of $M_{\rm BH} \sim 10^6\,\mathrm{M}_\odot$ have high Eddington ratios $\log_{10} f_{\rm Edd}$ in the range (-2,0) (Fig. 5 of Thomas et al. 2019). The presence of cold gas in the host galax-

[3] The strength of SN feedback in TNG changes with metallicity, and weakens with decreasing redshift (Pillepich et al. 2018a).

ies of these BHs, and their high gas fraction, provide the BHs with the additional torque-limited accretion channel. This explains the relatively high AGN OF for AGN with $L_{\rm AGN} = 10^{42} - 10^{44}$ erg/s located in $\sim 10^{9.5}\,{\rm M}_\odot$ galaxies.

Regardless of the luminosity range, Fig. 5 (top panels) shows that the AGN OF evolves with stellar mass, and varies from 0.001 to 0.6 (overall estimate derived from all the simulations together) when considering all AGN with hard X-ray luminosities above 10^{38} erg/s.

Focusing on the stellar mass range $M_\star = 10^9 - 10^{10.5}\,{\rm M}_\odot$, more massive galaxies have in general a higher probability of hosting an AGN, for any AGN luminosity. This is not the case for the SIMBA and TNG simulations. In the case of the TNGs, this is explained by the efficient low-accretion kinetic mode of the AGN feedback taking place. The self-regulation makes the AGN weaker, shifting them to the $L_{\rm AGN} = 10^{38} - 10^{40}$ erg/s panel. To conclude, all simulations predict a low intrinsic fraction of AGN with $10^{42} \leqslant L_{\rm AGN}/\,({\rm erg/s}) \leqslant 10^{44}$ in local low-mass galaxies with $M_\star \leqslant 10^{10.5}\,{\rm M}_\odot$ (below 40%), while half of the simulations predict an even lower fraction (below 10%).

4.2 Comparisons with observational constraints

We now turn to compare the AGN occupation fractions in simulations and observations, shown in Fig. 5 (bottom panels). To do so, we first compile an exhaustive list of the existing constraints from the literature (Shi et al. 2008; Aird et al. 2012; Schramm & Silverman 2013; Reines, Greene & Geha 2013; Miller et al. 2015; Lemons et al. 2015; Pardo et al. 2016; Aird, Coil & Georgakakis 2018; Mezcua et al. 2018; Kaviraj, Martin & Silk 2019; Mezcua & Domínguez Sánchez 2020; Birchall, Watson & Aird 2020; Birchall et al. 2022) that we report in Appendix A, along with a detailed discussion on how results are treated and homogenized. We use a cut in hard X-ray luminosity to derive the occupation fractions in the simulations. In some cases, however, the observational constraints are not for the exact same AGN hard X-ray luminosity as what we employ for the simulations. Other wavelengths and selection methods (e.g., emission line ratios, BPT diagrams) are sometimes used to identify AGN. The results of Reines, Greene & Geha (2013) are reported in all panels because they can not be connected to hard X-ray luminosities, and the results of Miller et al. (2015) in two panels because they can not be divided further. We sometimes extrapolate the results of some of these studies to derive an OF (when not already present in the respective papers). Finally, only some of these works employ a correction for completeness (often based on X-ray deep surveys, Aird et al. 2012; Miller et al. 2015; Pardo et al. 2016; Aird, Coil & Georgakakis 2018; Mezcua et al. 2018; Birchall, Watson & Aird 2020); this is a key aspect when computing an observational constraint on the AGN OF, and we detail this in Appendix A. While our comparison between the simulations and the observations presented in Fig. 5 carries uncertainties (because of e.g., observation completeness, observations in different wavelengths), it provides us with qualitative results.

Most constraints indicate an AGN OF in dwarf galaxies, across all luminosity ranges, of $\leqslant 0.03$, with significant variations over more than an order of magnitude. Kaviraj, Martin & Silk (2019) find an enhanced presence of AGN in low-mass galaxies, with an AGN OF $\geqslant 0.01$ for relatively faint AGN. Detecting the presence of AGN in dwarf galaxies is particularly difficult. Star formation can be a major contaminant at many wavelengths and is also aperture dependent. Recently, Lupi, Sbarrato & Carniani (2020) carry out a second analysis of the same sample, considering contamination from close sources, and contamination from the stellar galaxy component, and they find a much lower AGN fraction of about $0.0014 - 0.004$.

For relatively bright AGN ($L_{\rm AGN} > 10^{40}$ erg/s, top and middle panels of Fig. 5), most simulations are in broad agreement with observations. TNG100, TNG300 and SIMBA have overall higher AGN OF. In SIMBA this enhanced activity is because BHs need to accrete efficiently to reach the $M_{\rm BH} - M_\star$ scaling relation, by design of the simulation. At the faint end, $L_{\rm AGN} = 10^{38} - 10^{40}$ erg/s, we find a good agreement between EAGLE, Horizon-AGN and observations. Illustris and TNGs produce about one order of magnitude more faint AGN in dwarf galaxies because of the relatively weak SN feedback employed at low redshift and/or the high seeding mass (Pillepich et al. 2018a; Habouzit et al. 2021).

This analysis, however, carries large uncertainties from both the observational and theoretical side. In observations, completeness is not always stated in papers or easily inferred. In some simulations, BHs in low-mass galaxies are more massive compared to observations (Fig. 3, Fig. 4), and this could lead to more and/or brighter AGN, thus artificially enhancing the AGN OF. It is also important to mention that we did not correct the simulations for possible gas and dust obscuration of the AGN. The AGN OF derived from the simulations can thus be considered as upper limits when compared to observational constraints. We discuss this further in the Section 6.

4.3 Evolution of the AGN occupation fraction with redshift

4.3.1 Evolution in the simulations

The AGN OF varies significantly with redshift for all the simulations, often over more than one order of magnitude. This is shown in Fig. 6 for three AGN hard X-ray luminosity ranges.

In SIMBA, Horizon-AGN, and TNGs, the number of galaxies with $M_\star = 10^{9.5} - 10^{10.5}\,{\rm M}_\odot$ powering AGN with $L_{\rm AGN} = 10^{38} - 10^{40}$, and $10^{40} - 10^{42}$ erg/s (top and middle rows of Fig. 6) increases with time, from $z = 4$ to $z = 0$. As such, these AGN are more preponderant in the nearby Universe. The evolution with time in Illustris and EAGLE is not exactly the same: the fractions increase with time down to $z \sim 1$, but then decreases down to $z = 0$.

The number of galaxies producing brighter AGN with $L_{\rm AGN} = 10^{42} - 10^{44}$ erg/s decreases with time for all the simulations, and even more so in the redshift range $z = 1 - 0$. The strong decrease of these AGN at $z \leqslant 1$ is due to cosmic starvation: the quantity of gas available to feed the BHs decreases with time. At $z \geqslant 1$, instead, a significant fraction of low-mass galaxies may be harboring a relatively bright AGN, according to these simulations. For example, the AGN fraction ($L_{\rm AGN} \geqslant 10^{42}$ erg/s) in low-mass galaxies increases by at least $\sim 30\%$ from $z = 0$ to $z = 1$, in all simulations

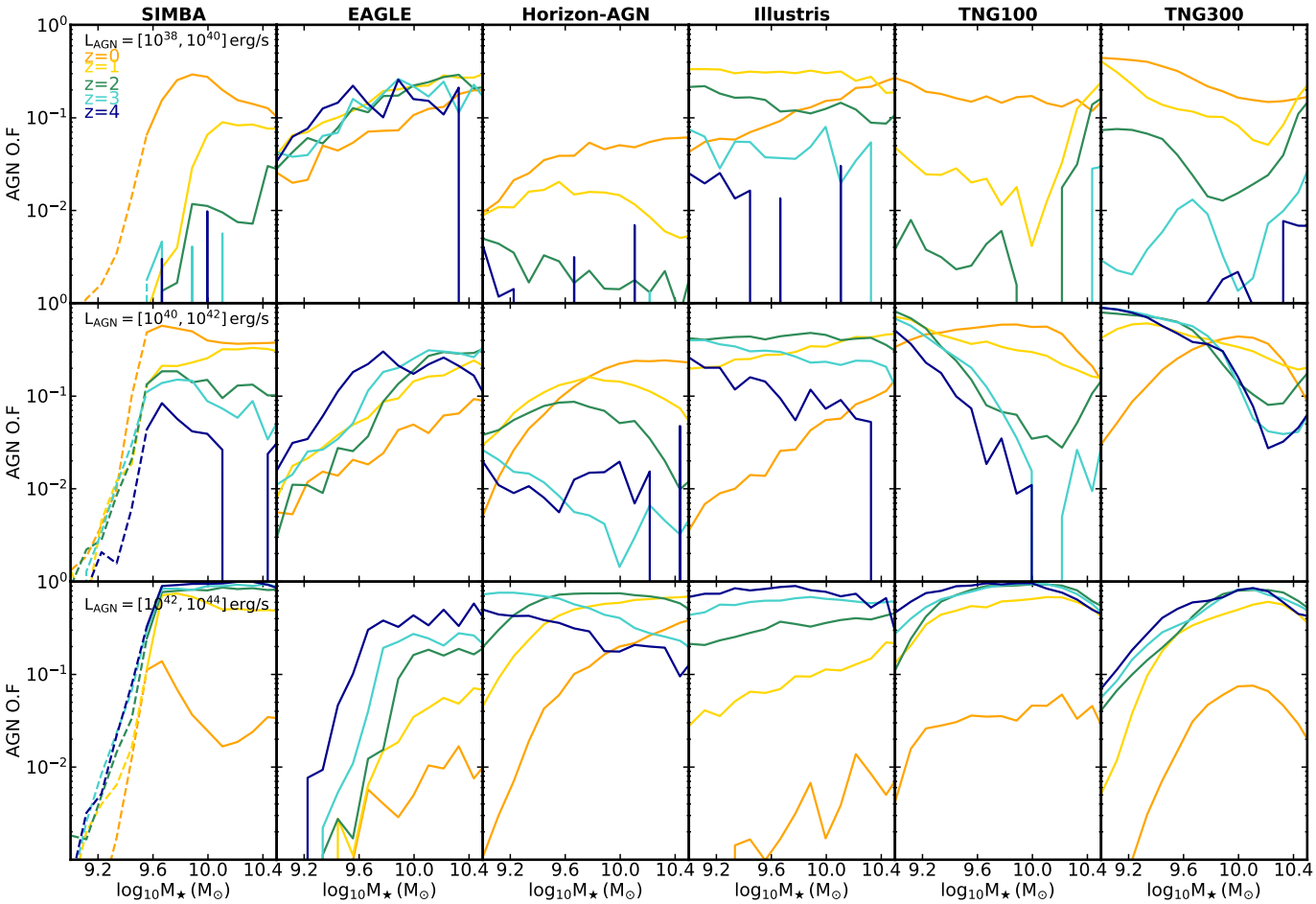

Figure 6. The time evolution of the AGN occupation fraction for $z = 0 - 4$ in low-mass galaxies $10^9 \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{10.5}$. AGN are selected on the basis of three different X-ray luminosity ranges: $10^{38} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{40}$ (top row), $10^{40} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{42}$ (middle row) and $10^{42} \leqslant L_{\rm AGN}/(\mathrm{erg/s}) \leqslant 10^{44}$ (bottom row). For visibility, we do not display here the Poisson errors (but see Fig 5 for an example at $z = 0$).

except EAGLE for galaxies of $\leqslant 10^{9.5}\,\mathrm{M}_\odot$. Pushing observations to $z = 1$ could allow us to build tighter constraints on the BH OF by detecting more AGN in low-mass galaxies with $M_\star = 10^9 - 10^{10.5}\,\mathrm{M}_\odot$.

4.3.2 Comparison with observations

In observations, whether the AGN OF increases with redshift is still unclear. This is partly due to observations being often biased towards more luminous AGN beyond the local Universe. Shi et al. (2008) find that the AGN OF slightly increases in $\sim 10^{10}\,\mathrm{M}_\odot$ galaxies from $0.1 < z < 0.4$ to $0.4 < z < 0.7$. Aird et al. (2012) also find that the AGN fraction, in fixed hard X-ray luminosity ranges (from $L_{\rm AGN} \sim 10^{42}$ erg/s to 10^{44} erg/s), increases from $0.2 < z < 0.6$ to $0.6 < z < 1$ in galaxies with $M_\star \sim 10^{10.25}\,\mathrm{M}_\odot$. Finally, (Birchall et al. 2022) also recently find a slight increase in the AGN fraction from $z = 0$ to $z = 0.3$ at fixed galaxy stellar mass (in the range $M_\star = 10^{10-11}\,\mathrm{M}_\odot$). We find the same trend in most of the simulations, a major difference being that the time evolution found in observations is smaller than the evolution in the simulations between $z = 0$ and $z = 1$.

On the other hand, several studies have found a decrease of the AGN fraction with redshift. Mezcua et al. (2018) report a decrease for AGN with $L_{\rm x} \geqslant 10^{42.4}$ erg/s in lower mass galaxies with $M_\star = 10^9 - 10^{9.5}\,\mathrm{M}_\odot$ at $z \sim 0.4$, and a slight increase at $z \sim 0.6$ that is still lower than the AGN occupation they find at $z \sim 0.1 - 0.2$. Aird, Coil & Georgakakis (2018) find a decrease of the AGN OF with redshift for AGN with $L_{\rm x} \leqslant 10^{42.4}$ erg/s and an increase for brighter AGN.

As said in the previous section, the different subgrid physics employed in the simulations leads to large discrepancies in the AGN fractions at $z = 1$, and to different evolution from $z = 1$ to $z = 0$. This makes $z = 1$ a key regime to further investigate the AGN OF in the observations, to possibly help disentangle the simulation modeling. The fact that observational studies have not reached yet a consensus on whether the OF increases towards higher redshift, also supports $z = 1$ as a crucial regime to constrain. New X-ray surveys from high sensitivity missions such as Athena, or the AXIS and Lynx NASA concept missions, will lead to the discovery of fainter AGN beyond the local Universe. Getting tighter constraints on the AGN OF also requires to overcome detection limits for the host low-mass galaxies

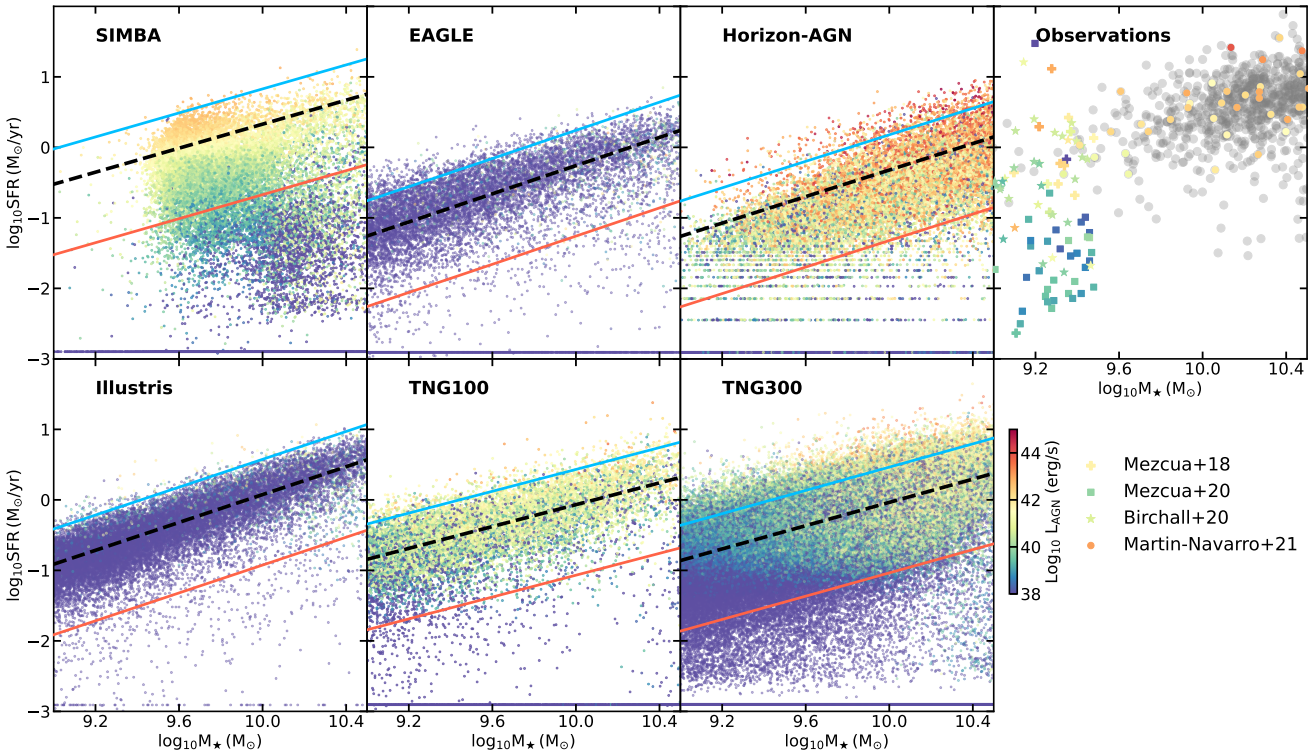

Figure 7. SFR-$M_\star$ diagram at $z = 0$ for the six cosmological simulations. Galaxies with SFR lower than 10^{-3} $M_\odot$/yr are set to this value. Galaxies are color coded by the hard X-ray luminosity of their AGN. We set all AGN luminosity values lower than $L_{\rm AGN} = 10^{38}$ ergs/s (including inactive BHs) to the aforementioned value. We show the star-forming main-sequence of each simulation with a black dashed line. Starburst galaxies are located above the blue lines, and quiescent galaxies below the red ones. In the right panel, we show several recent samples of galaxies from Mezcua & Domínguez Sánchez (2020); Birchall, Watson & Aird (2020), and Martín-Navarro, Shankar & Mezcua (2021). Observed galaxies are color coded (when possible, otherwise in grey) according to the luminosity of their AGN, as for the simulations. For a fair comparison, we have converted the AGN luminosity of several of these samples to hard (2-10 keV) X-ray luminosity.

(e.g., with JWST given a wide enough survey), and to match these galaxies to the AGN surveys.

5 RESULTS: PROPERTIES OF AGN HOST DWARF GALAXIES

Whether an AGN can be successfully detected does not only depend on its luminosity, but also on the properties of its host galaxy. For example, starburst activity in dwarf galaxies can be responsible for a high IR emission making IR detections of AGN challenging (e.g., Lupi, Sbarrato & Carniani 2020), but is also responsible for X-ray emission from binaries which can contaminate the AGN emission (Mezcua et al. 2018). To understand how many simulated AGN could be detected in observations, we investigate the mass and SFR properties of the AGN host galaxies.

5.1 Definition of the star-forming main-sequence

We start by computing the star-forming main sequence (MS, black dashed line in Fig. 7) for all the simulations at $z = 0$, by fitting their main SFR-$M_\star$ relation in the stellar mass range $M_\star = 10^9 - 10^{10}$ $M_\odot$ that is not affected strongly by AGN feedback quenching. The star-forming main-sequence is defined as a power law:

$$\log_{10} \frac{\rm SFR_{MS}}{\rm M_\odot/yr} = \alpha + \beta \log_{10} \left(\frac{M_\star}{10^{10}\,{\rm M_\odot}} \right). \quad (4)$$

Our definition is redshift and sample dependent, which is convenient since the simulations do not produce the exact same MS. This also allows one to compare with observational samples in a consistent way. For SIMBA, EAGLE, Horizon-AGN, Illustris, TNG100 and TNG300, we find the slope $\beta = 0.85, 0.99, 0.94, 0.98, 0.77, 0.82$ and the intercept $\alpha = 0.32, -0.25, -0.32, 0.08, -0.07, -0.04$, respectively. We define starburst galaxies as having SFR larger than 0.5 dex above the MS (above the blue line in Fig. 7), and quenched galaxies as galaxies 1 dex below the MS[4] (below the red line). At $z = 0$ for $\sim 10^9$ $M_\odot$ galaxies, the portion of starburst galaxies is about 5% and that of quenched galaxies between 12% and 33%. The large fraction of low-mass galaxies with reduced SFR mostly results from the quenching of satellite galaxies (Donnari et al. 2021; Dickey et al. 2021). Main-sequence galaxies are the most predominant with values ranging between $61 - 86\%$ across the simulations.

[4] The definition of starburst and quenched galaxies in a given simulation is with respect to the MS of this simulation. Thus, a starburst galaxy in a given simulation may not be a starburst in another simulation (even with the same SFR).

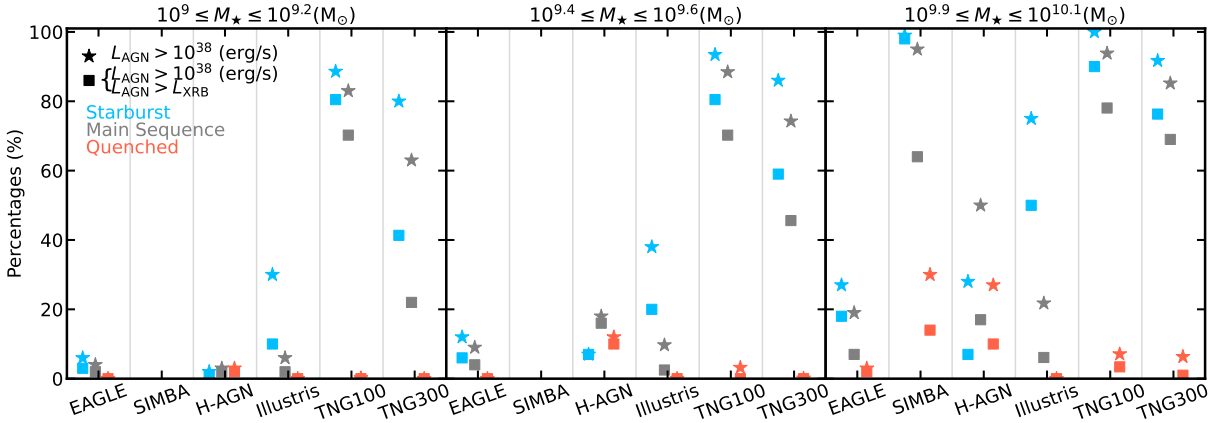

Figure 8. Percentages of galaxies at $z = 0$ with hard X-ray luminosity $L_{\mathrm{AGN}} > 10^{38}$ (erg/s) (shown as stars) and percentages of those galaxies which have a brighter AGN than the galaxy-wide XRB population, i.e., $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$ (shown as squares). We derive these percentages for different galaxy types color coded as: starburst (blue), main sequence (grey) and quenched (red). While only a small fraction of quenched galaxies host an AGN in most of the simulations, the fraction among starburst and main-sequence is more important. Starburst galaxies are not always the galaxies hosting the largest number of AGN (see Horizon-AGN). The fraction of galaxies for which the AGN outshine the galaxy-wide XRB population is significantly lower for all galaxy types and simulations.

5.2 Star-forming properties of low-mass galaxies hosting AGN

In Fig. 7, we show the SFR-$M_\star$ plane of the simulations. Each galaxy is color coded by its AGN hard X-ray luminosity. The correlation between AGN luminosity and the location of galaxies in the SFR-$M_\star$ diagram vary from one simulation to another. We find that, on average, brighter AGN are located in galaxies with higher SFR for all simulations at $z = 0$. The trend with SFR is clearly visible in SIMBA, Horizon-AGN and TNGs. In these simulations, a galaxy actively forming new stars is also on average feeding its central BH efficiently. Such a positive correlation can be expected due to a common gas supply for star formation and BH accretion (particularly for some simulation subgrid physics), but high time variability of BH accretion rate and its small spatial scales complicates a direct one-to-one connection to global star formation (e.g., Hickox 2014; Anglés-Alcázar et al. 2021). EAGLE and Illustris produce populations of faint AGN and inactive BHs, which makes any trend between SFR, $M_\star$, and L_{AGN} difficult to identify. This is due to i) the fact that BH accretion rates are computed from the BH gas cells in Illustris, which makes the rates stochastic and not well correlated with the global SFR of the host galaxies, ii) the lower accretion rates produced by the accretion modeling of EAGLE compared to other simulations, and its efficient SN feedback.

We now turn to quantify the number of galaxies[5] of different types hosting an AGN. We show the percentages of starburst, main-sequence, and quiescent galaxies hosting an AGN with $L_{\mathrm{AGN}} \geqslant 10^{38}$ erg/s in Fig. 8 (star symbols). We report the percentages for $L_{\mathrm{AGN}} \geqslant 10^{38}$ and $\geqslant 10^{40}$ erg/s in Table D1.

Most of the simulations do not produce any AGN (of $L_{\mathrm{AGN}} \geqslant 10^{38}$ or 10^{40} erg/s) in quenched galaxies with $M_\star \sim 10^{9}\,\mathrm{M}_\odot$. Discrepancies among the simulations arise in main-sequence galaxies. More than 50% of the TNG100 main-sequence galaxies have an AGN with $L_{\mathrm{AGN}} \geqslant 10^{40}$ erg/s, while this number drops below 5% for all the other simulations. We identify the same trends for more massive galaxies with $M_\star \sim 10^{9.5}\,\mathrm{M}_\odot$ and $M_\star \sim 10^{10}\,\mathrm{M}_\odot$, although with on average slightly higher fraction of AGN.

Interestingly, starburst galaxies are not always the ones powering more AGN with $L_{\mathrm{AGN}} \geqslant 10^{40}$ erg/s. For example, in Horizon-AGN main-sequence galaxies are on average more successfully feeding their AGN. Horizon-AGN is also the only simulation producing AGN in more than 10% of $\sim 10^{10}\,\mathrm{M}_\odot$ quenched galaxies.

Overall, there is no consensus in the simulations presented here on the fraction of AGN embedded in different types of galaxies based on their SFR. While a simulation like Horizon-AGN is able to power AGN in starburst, main-sequence, and a few quiescent galaxies, another simulation such as EAGLE will only power very few AGN in the main-sequence and starburst galaxies, and none in quiescent galaxies.

5.3 Comparison with observations

In the top right panel of Fig. 7, we compile some of the most recent observations in nearby or low-redshift low-mass galaxies. Because not all samples were obtained following the same observational techniques and selections, our study aims at qualitatively comparing in which low-mass galaxies AGN are powered both in observations and simulations.

In Fig. 7, we present the AGN samples of Mezcua et al. (2018); Mezcua & Domínguez Sánchez (2020); Birchall, Watson & Aird (2020). We also searched for X-ray counterparts to the AGN of Martín-Navarro, Shankar & Mezcua (2021). Technical details of these samples, and how we extract the hard X-ray luminosity of the AGN, are presented in Appendix E. When taken together, the observational samples show that on average galaxies with higher SFR host brighter AGN in $\sim 10^{9.5}\,\mathrm{M}_\odot$ galaxies, a feature that we also find in SIMBA, Horizon-AGN, and TNGs. The luminosities

[5] We do not differentiate between isolated and non-isolated galaxies, nor between central and satellite galaxies.

of the AGN are also in agreement with those produced in these simulations.

Galaxies hosting the brightest AGN in observations can have high SFR ($\log_{10}\mathrm{SFR}/(\mathrm{M}_\odot/\mathrm{yr}) \geqslant 0$ for $M_\star = 10^9 - 10^{10.5}\,\mathrm{M}_\odot$), with the most star-forming galaxies being absent from all the simulations studied here. Conversely, the lowest mass, least star-forming galaxies host on average very faint AGN, although some can reach $L_{\mathrm{AGN}} \sim 10^{42}\,\mathrm{erg/s}$. For galaxies with $M_\star < 10^{9.6}\,\mathrm{M}_\odot$, when using the TNG definition of quenched galaxies, we find that in observations (Birchall, Watson & Aird 2020; Mezcua & Domínguez Sánchez 2020; Mezcua et al. 2018) the brightest AGN in quenched galaxies have luminosities $L_{\mathrm{AGN}} < 10^{42.4}\,\mathrm{erg/s}$. Almost no simulation shows AGN in quenched dwarf galaxies in the same mass range. SIMBA and Horizon-AGN, however, defy this with e.g., a $\sim 3 - 12\%$ probability of finding very faint AGN with $L_{\mathrm{AGN}} \geqslant 10^{38}\,\mathrm{erg/s}$ in the galaxy mass range $M_\star = 10^9 - 10^{9.4}\,\mathrm{M}_\odot$. Although not shown in Fig. 7, Dickey et al. (2019) suggest that it is possible to find AGN in quenched galaxies[6]. Their work identifies 16 potential AGN hosts, out of a sample of 20 quenched galaxies with $M_\star = 10^{8.5} - 10^{9.5}\,\mathrm{M}_\odot$.

5.4 XRBs and AGN

The choice of hard X-ray emission is favored in the search of AGN because, in this band, obscuration is significantly reduced. However, AGN in low-mass galaxies face the challenge of being outshined by XRBs, especially in star-forming galaxies where HMXBs can be numerous (e.g., Lehmer et al. 2019). In the following, we quantify the number of galaxies in which XRBs could prevent the detection of AGN. To that purpose, Fig. 8 (square symbols, see also Table D1) shows the fractions of different type galaxies with $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$ for all the simulations.

We find that AGN with $L_{\mathrm{AGN}} \geqslant 10^{40}\,\mathrm{erg/s}$ do not suffer significantly from XRB contamination in all the simulations. However, the fraction of galaxies with an AGN of $L_{\mathrm{AGN}} \geqslant 10^{38}\,\mathrm{erg/s}$ brighter than the galaxy-wide XRB luminosity ($L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$) varies strongly from one simulation to another, for all the galaxy stellar mass bins and galaxy SFR types. For starburst galaxies in some simulations, only about half of the AGN with $L_{\mathrm{AGN}} \geqslant 10^{38}\,\mathrm{erg/s}$ are brighter than the host galaxy XRB population (SIMBA, Horizon-AGN), while in other simulations (e.g., SIMBA for galaxies with $\sim 10^{10}\,\mathrm{M}_\odot$ or TNG100) most of the AGN are brighter than the XRBs (i.e., close star and square symbols in Fig. 8). This depends on whether the simulations produce a global faint (e.g., EAGLE) or bright population of AGN (e.g., TNG100), and on the correlation between AGN luminosity and SFR. On average, the percentage of AGN with $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$ among the $L_{\mathrm{AGN}} \geqslant 10^{38}\,\mathrm{erg/s}$ AGN decreases from starburst to main-sequence, and to quiescent galaxies, for all galaxy masses. This is a direct consequence of AGN being generally fainter in these galaxies than in starburst galaxies.

We further explore the $z = 1$ regime. We find that the fraction of AGN able to outshine the galaxy-wide XRBs population increases by $\sim 10-30\%$ across the simulations. This is not true for EAGLE where the probability of finding an AGN brighter than an XRB at $z = 1$ is still relatively the same as that at $z = 0$. According to most of the simulations, the detections of AGN in X-ray could be less challenged by XRBs at $z = 1$ than at $z = 0$. The limitation would come from our instrumental power to detect relatively faint AGN beyond the local Universe.

[6] Dickey et al. (2019) use a definition of quenched galaxies different from ours, and based on the $H\alpha$ equivalent width and the Dn(4000) index.

6 DISCUSSION

6.1 Impact of the subgrid physics

Most cosmological simulations do not follow the current theoretical models of BH formation (e.g., PopIII remnant, compact stellar cluster, and direct collapse models), and instead seed massive halos or galaxies with a fixed mass BH. The seeding of some simulations leads to the presence of BHs exceeding $10^{6.5}\,\mathrm{M}_\odot$ in galaxies with stellar mass of $10^9 - 10^{9.5}\,\mathrm{M}_\odot$, while less massive BHs are found in observations. The decrease of the BH OF with time in almost all the simulations is mostly due to satellite galaxies loosing their BHs, and as such it is not a direct outcome of BH formation efficiency at high redshift. This means that the broad agreement at $z = 0$ between the simulations and observations in low-mass galaxies does not validate the seeding models employed in the large-scale cosmological simulations.

These large-scale simulations only form one BH per galaxy or halo, and some of them also re-position BH particles to their host galaxy's centers based on different criteria. BH dynamics is thus not followed on small scales, which can significantly affect the OF and location of low-mass BHs in galaxies (e.g. Ma et al. 2021), and BHs merge as soon as they are within the resolution distance of $\sim 1\,\mathrm{kpc}$ following galaxy mergers. In most of the simulations, these aspects result in the presence of a single BH per galaxy and the absence of off-center BHs, despite the latter being observed in local dwarf galaxies (Reines et al. 2020; Mezcua & Domínguez Sánchez 2020). BHs are more likely to accrete efficiently if located in the potential well of their host galaxies due to being embedded in their central gas reservoirs. Most simulations therefore likely provide an *optimistic* view on BHs' ability to accrete gas in dwarf galaxies.

Moreover, BH accretion in large-scale cosmological simulations is an unresolved physical process, and is modeled as subgrid physics. In most of the simulations (except SIMBA), the accretion rate onto a BH is always proportional to M_{BH}^2 (Bondi formalism). At fixed gas reservoir and fixed galaxy subgrid modeling, a more massive BH will accrete more than a less massive one. Since the BH masses are higher than in current observations in several simulations (Horizon-AGN, TNGs, SIMBA), then their AGN luminosities could be enhanced with respect to observations. It could explain the higher AGN luminosity function identified in Habouzit et al. (2022a, their Fig. 5, top panels) for AGN with $L \leqslant 10^{43.5}\,\mathrm{erg/s}$ for the TNG simulations, and with $L \geqslant 10^{43.5}\,\mathrm{erg/s}$ for Horizon-AGN. An agreement with the observed AGN OF would imply that observed and simulated BHs have very different accretion properties: to reach the same luminosity observed BHs (which are lighter than

simulated BHs) must have higher accretion rates. Depending on the resolution of the simulations, and how the accretion is computed (including the possible addition of a boost factor), the rates can be both under or over-estimated (Negri & Volonteri 2017; Anglés-Alcázar et al. 2021).

TNG100 produces the highest number of AGN (with $L_{\rm AGN} \geqslant 10^{40}$ erg/s) in low-mass galaxies of all the simulations, and this is mainly due to the high seed mass of $10^6\,{\rm M}_\odot$ and a less effective SN feedback at low redshift. In order to reduce the AGN OF below 0.03, in line with the bulk of observational constraints, about 90% of the TNG100 AGN should be obscured and missed in observations of dwarf galaxies. Alternatively, a large fraction of the AGN should not be located in high gas density regions and be off-centred, so that accretion is reduced. From the numerical point of view, more efficient SN feedback would also reduce the AGN fraction. On the other hand, good agreement between the simulations and the current constraints would mean that no AGN is obscured in low-mass galaxies, or that on average too few BHs are powering AGN in the simulation.

In this paper, we used the same radiative efficiency of 0.1 for all the simulations. Using a radiative efficiency of 0.2 (as employed for the BH accretion and AGN feedback model in the Illustris and TNG simulations) on average predicts a higher AGN fraction for $L_{\rm AGN} \geqslant 10^{40}$ erg/s (Fig. 5), resulting in more discrepancies between observations and simulations within that luminosity range.

6.2 Impact of the simulation resolution

The resolution of the simulations affect both the properties of the BHs and AGN, and consequently the AGN OF. One way to assess the impact of the resolution is comparing the sibling TNG simulations, whose subgrid model parameters are fixed and do not depend on the simulation resolution (Pillepich et al. 2018c). The less resolved gas content surrounding the BHs in TNG300 leads to lower accretion rates, partly due to a more impactful SN feedback (particularly at high redshift, and in low-mass galaxies). This can be seen in Fig. 3 and Fig. 4, where the $M_{\rm BH}$ scatter at fixed stellar mass is larger in TNG300 than TNG100, due to more low-mass BHs. These more numerous low-mass BHs do not accrete gas efficiently in TNG300, as shown by the AGN luminosity color code in Fig. 3. This reflects on the AGN OF: TNG300 has fewer bright AGN $\geqslant 10^{40}$ erg/s, whereas TNG100 more faint AGN (see Fig. 5).

6.3 Faint AGN and AGN obscuration

For relatively faint AGN in dwarf galaxies, observations are best conducted in galaxies with low SFR, for which the contamination from HMXBs should be smaller (as HMXB are short-lived). As we demonstrated here, the contamination from XRBs starts being important mostly for faint AGN with $L_{\rm x} \leqslant 10^{40}$ erg/s. Those AGN are powered by low-mass BHs in the simulations. In this paper, we have not modeled the obscuration of the AGN when computing their X-ray luminosities. As X-ray photons penetrate more easily through the gas content, X-ray detections are often used for AGN and are less biased toward more luminous AGN. Chen et al. (2017) report the analysis of a sample of 10 low-mass AGN. In the dwarf regime ($M_\star \leqslant 3\times10^9\,{\rm M}_\odot$), the authors find that 20% of the AGN do not show AGN-like optical narrow emission lines, and thus could be obscured in these wavelengths. Still low-mass galaxies could harbor a large fraction of heavily obscured AGN in X-ray, as hinted by X-ray stacking studied beyond the local Universe (Xue et al. 2012; Mezcua et al. 2016). AGN obscuration in a large fraction of low-mass galaxies could impact the comparison between simulations and observations.

6.4 Nuclear star clusters and BHs in low-mass galaxies

In Fig. 1, we show observational constraints from Sánchez-Janssen et al. (2019); Hoyer et al. (2021) on the nucleation fraction in nearby galaxies. While observing inactive BHs or faint AGN in local low-mass galaxies is extremely difficult, detecting nuclear star clusters in galaxies is possible today from the Local Group ($d \leqslant 3\,$Mpc) up to the Coma galaxy cluster (~ 100 Mpc).

The presence of nuclear star clusters in galaxies can facilitate the detection and characterization of the possible central low-mass BHs through dynamical measurements (e.g., enhanced velocity dispersion, Nguyen et al. 2018). A nuclear cluster can also theoretically enhance the number of tidal disruption events, up to a couple of orders of magnitude for galaxies with $M_\star \leqslant 3\times10^{10}\,{\rm M}_\odot$ (Pfister et al. 2020). Nucleated galaxies are thus good locations to investigate the presence of massive BHs. A few cases of the co-existence of BHs and nuclear clusters have been observed besides the Milky Way (see Table 3 in Neumayer, Seth & Böker (2020) and references therein). Whether BHs and nuclear clusters co-exist in the center of all local low-mass galaxies or only a fraction of them is still unclear. This currently prevents the extrapolation from the constraints on the nucleation fraction to a constraint on the BH OF.

In nearby nucleated galaxies, BHs tend to be overmassive compared to the nuclear clusters in galaxies with $M_\star \geqslant 10^{10}\,{\rm M}_\odot$ and undermassive in lower-mass galaxies as recently compiled in Neumayer, Seth & Böker (2020, their Fig. 13). In galaxies with $M_\star \leqslant 10^9\,{\rm M}_\odot$, the stellar population of nuclear clusters is commonly old and metal poor. The migration of globular clusters (usually with old stellar populations) towards the center of galaxies could explain such metallicity and stellar ages. The star clusters in more massive galaxies with $M_\star \geqslant 10^9\,{\rm M}_\odot$ have younger stellar populations, likely due to in-situ star formation. In observations, the nucleation fraction is generally found to be decreasing for more massive galaxies. For these aforementioned galaxies, the presence of a BH could disrupt the nuclear clusters (if present) and explain their absence (Antonini, Barausse & Silk 2015).

Understanding the possible co-existence of BHs and nuclear star clusters in cosmological simulations would require higher spatial resolution in the center of galaxies; there have been developments in this direction recently (Beckmann, Devriendt & Slyz 2019; Anglés-Alcázar et al. 2021). This would likely not be enough as some theories of nuclear cluster formation involve the migration of globular clusters to galaxy centers. In that case, high resolution in the entire galaxies would be needed to capture the assembly of the clusters;

this is currently not possible with cosmological simulations of $\sim 100\,\mathrm{cMpc}$ side length.

7 CONCLUSION

In this paper, we performed a systematic analysis of the BH and AGN population in dwarf and low-mass galaxies with $M_\star = 10^9 - 10^{10.5}\,\mathrm{M}_\odot$ in six large-scale cosmological simulations of volume $(\geqslant 100\,\mathrm{cMpc})^3$. We summarize our findings below.

- The modeling of some simulations (seeding and/or galaxy physics) produces a population of BHs too massive in low-mass galaxies with respect to current local observations. For example, the BHs of $\geqslant 10^{6.5}\,\mathrm{M}_\odot$ in simulated galaxies with $\leqslant 10^{9.5}\,\mathrm{M}_\odot$ (Fig. 3, Fig. 4).
- In simulations, the seeding of BHs in massive halos of $M_\mathrm{h} \sim 10^{10}\,\mathrm{M}_\odot$ or galaxies of $M_\star \geqslant 10^{9.5}\,\mathrm{M}_\odot$, or based on local gas properties, leads to a BH OF of almost unity in low-mass galaxies at high redshift ($z \geqslant 3$, Fig. 2).
- At fixed stellar mass, the OF decreases with time in all simulations except in SIMBA (Fig. 2). The decrease is driven by a larger number of satellite galaxies artificially losing their BHs. As such, the BH OF at $z = 0$ is not inherited from the intrinsic OF at high redshift. Comparison with observations at $z = 0$ thus do not inform us on the accuracy of the seeding subgrid models.
- At $z = 0$, almost all the simulations have a high BH OF in the range $70-90\%$ in galaxies with $M_\star \sim 10^9\,\mathrm{M}_\odot$ (Fig. 1). This is in broad agreement with some large observational constraints (coming from BH dynamical mass measurement and from extrapolation from AGN samples). Horizon-AGN has an OF of $\sim 10\%$, also in broad agreement with some constraints.
- The AGN OF in simulations spans from less than an order of magnitude in galaxies with $M_\star \sim 10^{10}\,\mathrm{M}_\odot$ to two (or more) orders of magnitude in galaxies with $M_\star \sim 10^9\,\mathrm{M}_\odot$ (Fig. 5).
- Current low-redshift observational constraints on the AGN OF span more than an order of magnitude (see compilation in Appendix A, Fig. 5). Compared to observations, some simulations power too many AGN in low-mass galaxies, except if e.g. almost all AGN in dwarfs are obscured. This could be due to a more massive BH population in low-mass galaxies in these simulations (Fig. 3, Fig. 4).
- There is no consensus among the simulations on whether the fraction of AGN in low-mass galaxies is higher at higher redshift (e.g., $z \geqslant 1$). If so, it could allow us to put tighter constraints on the BH OF by observing more AGN (thus reducing the extrapolation from the AGN OF to BH OF). The combination of new X-ray facilities such as Athena, and new galaxy surveys (e.g., JWST) will be pivotal.
- The correlation between AGN luminosity and SFR found in observations of low-mass galaxies is also produced in some simulations, but not all. At fixed galaxy stellar mass, starburst galaxies often host more and brighter AGN than main-sequence and quiescent galaxies. This does not help in constraining the AGN OF, as emission from star-forming regions prevent the AGN detections in several wavelength bands.
- We find that AGN with $L_\mathrm{AGN} \geqslant 10^{40}\,\mathrm{erg/s}$ almost always outshine the galaxy-wide XRB population (in starburst, main-sequence, and quiescent galaxies of all the simulations). As our ability to detect fainter X-ray AGN with $L_\mathrm{AGN} \geqslant 10^{38}\,\mathrm{erg/s}$ will improve with time, their confirmation, however, could be significantly reduced due to the contamination from XRBs.

The BH and AGN occupation fractions have implications for the evolution of low-mass galaxies in the simulations. For example, the small BH OF in low-mass galaxies at $z \leqslant 2$ in Horizon-AGN prevents these galaxies from any energy release via AGN feedback. The role of AGN feedback in low-mass galaxies is still unclear today (e.g., Dashyan et al. 2018; Koudmani et al. 2019), but evidence for feedback has been observed (e.g., Mezcua, Suh & Civano 2019; Manzano-King, Canalizo & Sales 2019). Theoretically, the effect could be small as the BHs have low masses and could have reduced accretion rates due to SN feedback (Dubois et al. 2015; Habouzit, Volonteri & Dubois 2017; Anglés-Alcázar et al. 2017b), but AGN feedback could also provide a second channel to clear out the low-mass galaxies from their gas content and enhance the effect of SN feedback. On the other hand, if too many (and/or too bright) AGN are produced in the simulations compared to observations, they can potentially release too much energy in the host galaxy history, which can e.g., alter the properties of the gas, and galaxy gas and stellar contents. In the hierarchical build-up of galaxies, this can potentially affect the initial stages of many galaxies, and thus needs to be considered and further investigated in the large-scale cosmological simulations.

ACKNOWLEDGMENTS

We are grateful to the anonymous referee for their insightful comments and suggestions that improved this work. Authors would also like to thank G. Mamon, S. Kaviraj, and S. Genel for fruitful discussions. HH thanks the Institut d'Astrophysique de Paris for hosting her. M.M. acknowledges support from the Ramon y Cajal fellowship (RYC2019-027670-I). DAA was supported in part by NSF grants AST-2009687 and AST-2108944, and by the Flatiron Institute, which is supported by the Simons Foundation.

DATA AVAILABILITY STATEMENT

The data from the Illustris and the TNG100 simulations can be found on their respective websites: https://www.illustris-project.org, https://www.tng-project.org. The data from the EAGLE simulation can be obtained upon request to the EAGLE team at their website: http://icc.dur.ac.uk/Eagle/. The data from the SIMBA simulation can be found on the website: http://simba.roe.ac.uk/. The Horizon-AGN simulation is not public, but some catalogs are available at: https://www.horizon-simulation.org/data.html.

REFERENCES

Agarwal B., Dalla Vecchia C., Johnson J. L., Khochfar S., Paardekooper J.-P., 2014, MNRAS, 443, 648

Aird J., Coil A. L., Georgakakis A., 2017, MNRAS, 465, 3390
Aird J., Coil A. L., Georgakakis A., 2018, MNRAS, 474, 1225
Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892
Aird J. et al., 2012, ApJ, 746, 90
Alister Seguel P. J., Schleicher D. R. G., Boekholt T. C. N., Fellhauer M., Klessen R. S., 2020, MNRAS, 493, 2352
Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017a, MNRAS, 464, 2840
Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017b, MNRAS, 472, L109
Anglés-Alcázar D., Özel F., Davé R., 2013, ApJ, 770, 5
Anglés-Alcázar D., Özel F., Davé R., Katz N., Kollmeier J. A., Oppenheimer B. D., 2015, ApJ, 800, 127
Anglés-Alcázar D. et al., 2021, ApJ, 917, 53
Antonini F., Barausse E., Silk J., 2015, The Astrophysical Journal, 812, 72
Baldassare V. F., Reines A. E., Gallo E., Greene J. E., 2015, ApJ, 809, L14
Baldassare V. F., Reines A. E., Gallo E., Greene J. E., 2017, ApJ, 836, 20
Baldassare V. F. et al., 2016, The Astrophysical Journal, 829, 57
Beckmann R. S., Devriendt J., Slyz A., 2019, MNRAS, 483, 3488
Begelman M. C., Volonteri M., Rees M. J., 2006a, MNRAS, 370, 289
Begelman M. C., Volonteri M., Rees M. J., 2006b, MNRAS, 370, 289
Bellovary J., Volonteri M., Governato F., Shen S., Quinn T., Wadsley J., 2011, ApJ, 742, 13
Birchall K. L., Watson M. G., Aird J., 2020, MNRAS, 492, 2268
Birchall K. L., Watson M. G., Aird J., Starling R. L. C., 2022, MNRAS, 510, 4556
Buchner J. et al., 2015, ApJ, 802, 89
Carlsten S. G., Greene J. E., Beaton R. L., Greco J. P., 2021, arXiv e-prints, arXiv:2105.03440
Chen C.-T. J. et al., 2017, The Astrophysical Journal, 837, 145
Chilingarian I. V., Katkov I. Y., Zolotukhin I. Y., Grishin K. A., Beletsky Y., Boutsia K., Osip D. J., 2018, ApJ, 863, 1
Chon S., Hosokawa T., Omukai K., 2021, MNRAS, 502, 700
Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91
Crain R. A. et al., 2015, MNRAS, 450, 1937
Dashyan G., Silk J., Mamon G. A., Dubois Y., Hartwig T., 2018, MNRAS, 473, 5698
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Davis T. A. et al., 2020, MNRAS, 496, 4061
Devecchi B., Volonteri M., 2009, The Astrophysical Journal, 694, 302
Dickey C. M., Geha M., Wetzel A., El-Badry K., 2019, The Astrophysical Journal, 884, 180
Dickey C. M. et al., 2021, ApJ, 915, 53
Dijkstra M., Haiman Z., Mesinger A., Wyithe J. S. B., 2008, MNRAS, 391, 1961
Dong X.-B., Ho L. C., Yuan W., Wang T.-G., Fan X., Zhou H., Jiang N., 2012, ApJ, 755, 167
Donnari M. et al., 2021, MNRAS, 500, 4004
Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948
Dubois Y. et al., 2014, MNRAS, 444, 1453
Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., Teyssier R., 2015, MNRAS, 452, 1502
Duras F. et al., 2020, A&A, 636, A73
Evans I. N. et al., 2010, ApJS, 189, 37
Filippenko A. V., Sargent W. L. W., 1989, ApJ, 342, L11
Fornasini F. M., Civano F., Fabbiano G., Elvis M., Marchesi S., Miyaji T., Zezas A., 2018, ApJ, 865, 43
Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013, ApJ, 776, L31
Gallo E., Treu T., Jacob J., Woo J.-H., Marshall P. J., Antonucci R., 2008, ApJ, 680, 154
Gallo E., Treu T., Marshall P. J., Woo J.-H., Leipski C., Antonucci R., 2010, ApJ, 714, 25
Genel S. et al., 2014, MNRAS, 445, 175
Greene J. E., 2012, Nature Communications, 3, 1304
Greene J. E., Ho L. C., 2004, ApJ, 610, 722
Greene J. E., Ho L. C., Barth A. J., 2008, ApJ, 688, 159
Greene J. E., Strader J., Ho L. C., 2019, arXiv e-prints, arXiv:1911.09678
Greene J. E., Strader J., Ho L. C., 2020, ARA&A, 58, 257
Gültekin K. et al., 2009, ApJ, 698, 198
Habouzit M. et al., 2021, MNRAS, 503, 1940
Habouzit M. et al., 2022a, MNRAS, 511, 3751
Habouzit M. et al., 2022b, MNRAS, 509, 3015
Habouzit M., Volonteri M., Dubois Y., 2017, MNRAS, 468, 3935
Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529
Hickox R. C., 2014, AAS/High Energy Astrophysics Division# 14, 14, 400
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hopkins P. F., Richards G. T., Hernquist L., 2007a, ApJ, 654, 731
Hopkins P. F., Richards G. T., Hernquist L., 2007b, ApJ, 654, 731
Hoyer N., Neumayer N., Georgiev I. Y., Seth A. C., Greene J. E., 2021, arXiv preprint arXiv:2107.05313
Johnson J. L., Bromm V., 2007a, MNRAS, 374, 1557
Johnson J. L., Bromm V., 2007b, MNRAS, 374, 1557
Kauffmann G. et al., 2003, Monthly Notices of the Royal Astronomical Society, 346, 1055
Kaviraj S., Martin G., Silk J., 2019, MNRAS, 489, L12
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Koudmani S., Henden N. A., Sijacki D., 2021, MNRAS, 503, 3568
Koudmani S., Sijacki D., Bourne M. A., Smith M. C., 2019, MNRAS, 484, 2047
Latimer C. J., Reines A. E., Bogdan A., Kraft R., 2021, arXiv e-prints, arXiv:2111.10388
Lehmer B. D., Alexander D. M., Bauer F. E., Brandt W. N., Goulding A. D., Jenkins L. P., Ptak A., Roberts T. P., 2010, ApJ, 724, 559
Lehmer B. D. et al., 2016, ApJ, 825, 7
Lehmer B. D. et al., 2021, ApJ, 907, 17
Lehmer B. D. et al., 2019, ApJS, 243, 3

Lemons S. M., Reines A. E., Plotkin R. M., Gallo E., Greene J. E., 2015, The Astrophysical Journal, 805, 12
Li Y. et al., 2019, arXiv e-prints, arXiv:1910.00017
Liu H.-Y., Yuan W., Dong X.-B., Zhou H., Liu W.-J., 2018, ApJS, 235, 40
Lupi A., Sbarrato T., Carniani S., 2020, MNRAS, 492, 2528
Ma L., Hopkins P. F., Ma X., Anglés-Alcázar D., Faucher-Giguère C.-A., Kelley L. Z., 2021, MNRAS, 508, 1973
Madau P., Fragos T., 2017, ApJ, 840, 39
Madau P., Rees M. J., 2001, ApJL, 551, L27
Manzano-King C. M., Canalizo G., Sales L. V., 2019, ApJ, 884, 54
Marinacci F. et al., 2018, MNRAS, 480, 5113
Marleau F. R., Clancy D., Habas R., Bianconi M., 2017, A&A, 602, A28
Martín-Navarro I., Shankar F., Mezcua M., 2021, MNRAS
McAlpine S., Bower R. G., Harrison C. M., Crain R. A., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 468, 3395
McAlpine S., Bower R. G., Rosario D. J., Crain R. A., Schaye J., Theuns T., 2018, MNRAS, 481, 3118
McAlpine S. et al., 2016, Astronomy and Computing, 15, 72
Mezcua M., 2017, International Journal of Modern Physics D, 26, 1730021
Mezcua M., Civano F., Fabbiano G., Miyaji T., Marchesi S., 2016, ApJ, 817, 20
Mezcua M., Civano F., Marchesi S., Suh H., Fabbiano G., Volonteri M., 2018, MNRAS, 478, 2576
Mezcua M., Domínguez Sánchez H., 2020, ApJ, 898, L30
Mezcua M., Sánchez H. D., 2020, The Astrophysical Journal Letters, 898, L30
Mezcua M., Suh H., Civano F., 2019, MNRAS, 488, 685
Miller B. P., Gallo E., Greene J. E., Kelly B. C., Treu T., Woo J.-H., Baldassare V., 2015, ApJ, 799, 98
Mineo S., Gilfanov M., Sunyaev R., 2012, MNRAS, 419, 2095
Moran E. C., Shahinyan K., Sugarman H. R., Vélez D. O., Eracleous M., 2014, The Astronomical Journal, 148, 136
Mullaney J. R. et al., 2012, ApJ, 753, L30
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Negri A., Volonteri M., 2017, MNRAS, 467, 3475
Nelson D. et al., 2015, Astronomy and Computing, 13, 12
Nelson D. et al., 2018, MNRAS, 475, 624
Neumayer N., Seth A., Böker T., 2020, A&AR, 28, 4
Nguyen D. D. et al., 2017, ApJ, 836, 237
Nguyen D. D. et al., 2019, ApJ, 872, 104
Nguyen D. D. et al., 2018, ApJ, 858, 118
Omukai K., Schneider R., Haiman Z., 2008, ApJ, 686, 801
Pacucci F., Mezcua M., Regan J. A., 2021, ApJ, 920, 134
Pardo K. et al., 2016, ApJ, 831, 203
Pardo K. et al., 2016, The Astrophysical Journal, 831, 203
Pfister H., Volonteri M., Dai J. L., Colpi M., 2020, MNRAS, 497, 2276
Pillepich A. et al., 2018a, MNRAS, 475, 648
Pillepich A. et al., 2018b, MNRAS, 473, 4077
Pillepich A. et al., 2018c, MNRAS, 473, 4077
Regan J. A., Haehnelt M. G., 2009, MNRAS, 393, 858
Reines A. E., Comastri A., 2016, PASA, 33, e054
Reines A. E., Condon J. J., Darling J., Greene J. E., 2020, ApJ, 888, 36
Reines A. E., Greene J. E., Geha M., 2013, The Astrophysical Journal, 775, 116
Reines A. E., Sivakoff G. R., Johnson K. E., Brogan C. L., 2011, Nature, 470, 66
Reines A. E., Volonteri M., 2015, The Astrophysical Journal, 813, 82
Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, MNRAS, 462, 190
Rosas-Guevara Y. M. et al., 2015, MNRAS, 454, 1038
Sánchez-Janssen R. et al., 2019, ApJ, 878, 18
Sartori L. F., Schawinski K., Treister E., Trakhtenbrot B., Koss M., Shirazi M., Oh K., 2015, MNRAS, 454, 3722
Satyapal S., Secrest N. J., McAlpine W., Ellison S. L., Fischer J., Rosenberg J. L., 2014, ApJ, 784, 113
Schaye J. et al., 2015, MNRAS, 446, 521
Schirra A. P. et al., 2020, arXiv e-prints, arXiv:2011.02501
Schramm M., Silverman J. D., 2013, The Astrophysical Journal, 767, 13
She R., Ho L. C., Feng H., 2017, ApJ, 842, 131
Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, MNRAS, 495, 3252
Shi Y., Rieke G., Donley J., Cooper M., Willmer C., Kirby E., 2008, ApJ, 688, 794
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Springel V. et al., 2018, MNRAS, 475, 676
Thomas N., Davé R., Anglés-Alcázar D., Jarvis M., 2019, MNRAS, 487, 5764
Thomas N., Davé R., Jarvis M. J., Anglés-Alcázar D., 2021, MNRAS, 503, 3492
Trump J. R. et al., 2015, The Astrophysical Journal, 811, 26
van Wassenhove S., Volonteri M., Walker M. G., Gair J. R., 2010, MNRAS, 408, 1139
Visbal E., Haiman Z., Bryan G. L., 2014, MNRAS, 442, L100
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, MNRAS, 460, 2979
Volonteri M., Haardt F., Madau P., 2003a, ApJ, 582, 559
Volonteri M., Haardt F., Madau P., 2003b, ApJ, 582, 559
Volonteri M., Lodato G., Natarajan P., 2008a, MNRAS, 383, 1079
Volonteri M., Lodato G., Natarajan P., 2008b, MNRAS, 383, 1079
Weinberger R. et al., 2017, MNRAS, 465, 3291
Weinberger R. et al., 2018, MNRAS, 479, 4056
Wise J. H., Regan J. A., O'Shea B. W., Norman M. L., Downes T. P., Xu H., 2019, Nature, 566, 85
Woo J.-H., Cho H., Gallo E., Hodges-Kluck E., Le H. A. N., Shin J., Son D., Horst J. C., 2019, Nature Astronomy, 3, 755
Xue Y. et al., 2012, The Astrophysical Journal, 758, 129

APPENDIX A: CURRENT OBSERVATIONAL CONSTRAINTS ON THE AGN OCCUPATION FRACTION

We provide in this section an exhaustive list of existing constraints on the AGN OF in the low-redshift Universe. These observational constraints are used in Section 4 and reported in Fig. 5.

The early work of Kauffmann et al. (2003) derive the AGN OF from the SDSS survey, finding a fraction below 10% for galaxies with $M_\star \leqslant 10^{10}\,\mathrm{M}_\odot$, which peaks at about 25% in galaxies with $M_\star \sim 10^{11}\,\mathrm{M}_\odot$.

The constraints of Shi et al. (2008) use a luminosity threshold of $L_{\mathrm{x,\,AGN,\,2-8\,keV}} \geqslant 10^{42}\,\mathrm{erg/s}$, and galaxies with $M_\star = 5\times10^{9} - 2\times10^{10}\,\mathrm{M}_\odot$. The AGN of this sample have 2-8 keV luminosities mostly in the range $L_{\mathrm{x,\,AGN,\,2-8\,keV}} = 10^{42} - 10^{43.5}\,\mathrm{erg/s}$. This can be converted to the $2-10\,\mathrm{keV}$ band using $L_{2-8\,\mathrm{keV}}/L_{2-10\,\mathrm{keV}} = 0.86$, assuming a power law photon index of 1.0 (Shi et al. 2008). As such, we get: $L_{\mathrm{x,\,AGN,\,2-10\,keV}} = 10^{41.9} - 10^{43.4}\,\mathrm{erg/s}$.

Aird et al. (2012) identify more than 200 AGN with $L_{\mathrm{x,\,AGN,\,2-10keV}} \geqslant 10^{42}\,\mathrm{erg/s}$ in more than 25000 galaxies at $z < 1$. For galaxies with $M_\star \sim 10^{9.75}\,\mathrm{M}_\odot$ and $0.2 < z < 0.6$, the occupation fractions are within the ranges OF$= 0.3-1.6, 0.17-0.5, 0.01-0.17\,\%$ for X-ray (2-10 keV) AGN luminosities of $L_\mathrm{x} = 10^{42}-10^{42.5}, 10^{42.5}-10^{43}, 10^{43}-10^{43.5}\,\mathrm{erg/s}$, respectively. For $0.6 < z < 1$ and galaxies with $M_\star \sim 10^{10.3}\,\mathrm{M}_\odot$, the authors find AGN occupation fractions of OF $= 0.6-1.6, 0.3-1.0, 0.2-0.3\,\%$ for $L_\mathrm{x} = 10^{42}-10^{42.5}, 10^{42.5}-10^{43}, 10^{43}-10^{43.5}\,\mathrm{erg/s}$, respectively. Aird et al. (2012) find that at fixed AGN luminosity the AGN fraction decreases toward less massive galaxies in the range $M_\star = 10^{10}-10^{12}\,\mathrm{M}_\odot$. If this trend extends down to even lower mass galaxies, the AGN OF in galaxies of $M_\star < 10^{9.75}\,\mathrm{M}_\odot$ could be lower than the AGN fractions listed above.

Schramm & Silverman (2013) report X-ray detections in more than 5000 galaxies with $M_\star \leqslant 3\times10^{9}\,\mathrm{M}_\odot$ at $z < 1$, among which 3 detections out of 2100 galaxies at $z < 0.3$. The latter provides a fraction of OF $= 0.14\%$ at $z < 0.3$. These fractions are for a galaxy X-ray luminosity of $L_{\mathrm{x,\,gal,\,0.5-8\,keV}} \leqslant 10^{44}\,\mathrm{erg/s}$. The three detections are in galaxies with $L_{\mathrm{x,\,gal,\,0.5-8\,keV}} = 10^{40.3}, 10^{40.3}, 10^{41.1}\,\mathrm{erg/s}$. The SFR and $M_\star$ of these three galaxies are given in Schramm & Silverman (2013), which allows us to estimate the AGN luminosities by subtracting the contribution of the XRBs and the ISM.

The constraints of Reines, Greene & Geha (2013) include $\sim$ 150 galaxies that have been identified with BH accretion signatures out of 25974 galaxies, i.e. 136 galaxies with photoionization signatures of BH accretion (and for some of them broad $H\alpha$ emission) and 16 galaxies with narrow-line ratios which are consistent with AGN or star formation. Based on the follow-ups carried by Baldassare et al. (2016) on the 16 candidates, 11 galaxies had a broad H_α line that vanished and were thus discarded as AGN, 2 galaxies were proven to be AGN, and 3 galaxies were still ambiguous and still qualified as AGN candidates. From the 140 remaining galaxies, we derive an AGN fraction of OF $= 0.53\%$ for galaxies with $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$ in the local Universe. From the X-ray follow up of Baldassare et al. (2017), the hard X-ray luminosity of these AGN are within the range $L_{\mathrm{x,\,AGN,\,2-10\,keV}} = 10^{39.76} - 10^{41.84}\,\mathrm{erg/s}$.

Satyapal et al. (2014) investigate the presence of AGN in 13862 low-redshift bulgeless galaxies from SDSS with extremely red mid-IR colors, possible signature of an AGN. From 25 to 300 AGN candidates are extracted with the color selection; they do not have AGN signature in the optical. The AGN fraction is about $\leqslant 40\%$ in bulgeless galaxies with $M_\star = 10^{8} - 10^{10}\,\mathrm{M}_\odot$.

Moran et al. (2014) finds an OF of 2.7% for galaxies with $M_\star = 4\times10^{9} - 10^{10}\,\mathrm{M}_\odot$. The study uses a distance-limited sample of galaxies from SDSS, in which 28 AGN are identified based on their optical emission lines.

Lemons et al. (2015) identify 19 galaxies with X-ray detections out of about 491 galaxies with $M_\star \leqslant 3\times10^{9}\,\mathrm{M}_\odot$ (here we used the estimate of the total number of galaxies from Pardo et al. 2016). In the sample, 10 galaxies have signatures of a central AGN, and 8 galaxies have a higher X-ray emission than the predicted galaxy-wide emission from XRBs. From the latter, we estimate an AGN fraction of OF $= 1.63\%$. X-ray sources have $2-10\,\mathrm{keV}$ luminosities of $L_{\mathrm{x,\,source,\,2-10\,keV}} \leqslant 10^{40.1}\,\mathrm{erg/s}$.

Miller et al. (2015) find that $1-2\%$ of galaxies with $M_\star \leqslant 10^{9}\,\mathrm{M}_\odot$ have nuclear X-ray emission. We consider the result as an upper limit, since XRBs could contaminate the emission. We also report these results in all the panels of Fig. 5, since no X-ray luminosities are given in this study.

Sartori et al. (2015) identify 336 AGN candidates in a sample of $\sim$ 48000 SDSS dwarf galaxies with $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$ at $z < 0.1$. AGN are selected based on the BPT diagram, but also if they exhibit optical narrow-line signatures of an accreting BH (He II selected AGN), or based on galaxy mid-IR colors consistent with the presence of an AGN. The AGN OF is within the range $\sim 0.02-0.5\%$ for galaxies with $M_\star \sim 10^{8.5}\,\mathrm{M}_\odot$, and of $0.1-0.3\%$ for galaxies with $M_\star \sim 10^{9.5}\,\mathrm{M}_\odot$ and for these three selection methods.

Pardo et al. (2016) derive an occupation fraction of OF $= 0.6-3\%$ for galaxies with $M_\star = 10^{9}-3\times10^{9}\,\mathrm{M}_\odot$, and $L_{\mathrm{x,\,source,0.5-7\,keV}} = 10^{39.9} - 10^{42.1}\,\mathrm{erg/s}$. From the hardness ratios, the authors found a low probability of contamination from X-ray emitted by HMXBs and/or from the hot gas associated with SN remnants. We report these values in Fig. 5, although the band used in Pardo et al. (2016) is not hard 2-10 keV X-ray as used in this paper.

Furthermore, we include the constraints of Mezcua et al. (2016) which investigate the presence of AGN in dwarf starburst and spiral galaxies up to $z = 1.5$. They use stacking X-ray counts from more than 7000 of galaxies in the range $0 < z < 0.3$, and about 10000 galaxies and 3000 galaxies in the range $0.7 < z < 1.0$, and $1 < z < 1.5$, respectively. The fraction of X-ray excess after removing the possible contributions from XRBs and hot ISM is $\sim 70\%$ in the redshift bins below $z < 0.7$, and is $\sim 50\%$ in the bins of higher redshifts. From stacking analysis, we can not assess whether the X-ray excess is due to a few galaxies with powerful AGN, or to many galaxies with more moderate AGN. However, the results of Mezcua et al. (2016) suggest that AGN could be common in low-mass galaxies up to $z = 1.5$.

Marleau et al. (2017) investigate the presence of AGN in dwarf galaxies (62) with stellar masses from 10^{6} to $10^{9}\,\mathrm{M}_\odot$. Their sample consists of 303 candidate AGN, selected based on IR colors. Using the MPA/JHU catalogue, Marleau et al.

(2017) derives an AGN fraction that does not exceed 5% across the stellar mass range $10^6 \leqslant M_\star/\mathrm{M}_\odot \leqslant 10^{11}$. The fraction slightly increases as a function of decreasing stellar mass for $M_\star \leqslant 10^9\,\mathrm{M}_\odot$.

Mezcua et al. (2018) constrains the AGN fraction in galaxies with $M_\star = 10^9 - 3\times10^9\,\mathrm{M}_\odot$. The fraction of AGN in dwarf galaxies is OF $\sim 0.43\%$ for $L_{\mathrm{x,\,AGN,\,0.5-10\,keV}} = 10^{41.5} - 10^{42.4}$ erg/s and at $z < 0.3$, and OF $\sim 0.09\%$ for $L_{\mathrm{x,\,AGN,\,0.5-10\,keV}} \geqslant 10^{42.4}$ erg/s and at $z < 0.3$. At higher redshift of $z \sim 0.6$, the AGN OF is lower with OF $\sim 0.04\%$ for $L_{\mathrm{x,\,AGN,\,0.5-10\,keV}} \geqslant 10^{42.4}$ erg/s.

Aird, Coil & Georgakakis (2018) find an AGN fraction of OF $\sim 0.1-0.9\%$ in galaxies with $M_\star = 10^{8.5} - 10^{9.5}\,\mathrm{M}_\odot$. This fraction is defined as the number of AGN with an Eddington ratio $f_{\mathrm{Edd}} > 0.01$. The exact definition in Aird, Coil & Georgakakis (2018) is based on AGN with $k_{\mathrm{bol}} \times L_{\mathrm{x,\,AGN,\,2-10\,keV}}/(1.38\times10^{38}\,\mathrm{erg/s}\times0.002\times M_\star/\mathrm{M}_\odot) > 0.01$, with this quantity being similar to f_{Edd} for their choice of parameters. This limit can be converted into a limit in 2-10 keV luminosity of $L_{\mathrm{x,\,AGN,\,2-10\,keV}} \geqslant 10^{41} - 10^{41.5}$ erg/s for the range $M_\star = 10^9 - 10^{9.5}\,\mathrm{M}_\odot$.

Chilingarian et al. (2018) provides a sample of 305 AGN candidates with Type 1 signatures in the range $M_{\mathrm{BH}} = 3\times10^4 - 2\times10^5\,\mathrm{M}_\odot$ in galaxies with $M_{\mathrm{bulge}} = 10^7 - 10^{12}\,\mathrm{M}_\odot$, out of which 10 sources were confirmed as AGN. The X-ray luminosity of the AGN varies in $L_{\mathrm{x,\,AGN}} = 10^{40} - 3\times10^{42}$ erg/s, for different X-ray bands.

Dickey et al. (2019) present constraints on the AGN fraction in a small sample of galaxies, including 20 quiescent galaxies and 7 star-forming galaxies. Interestingly, many of the quiescent galaxies show signatures of AGN activity (from optical emission line ratios), and leads to an occupation fraction of OF = 80% in those galaxies. A fraction of OF = 28.5% is found in the star-forming galaxies. Given the small sample studied in Dickey et al. (2019), we do not add these constraints to Fig. 5.

Kaviraj, Martin & Silk (2019) recently derived the fraction of AGN in a wide range of galaxy stellar mass ($M_\star = 10^7 - 10^{11.5}\,\mathrm{M}_\odot$), using diagnostics based on IR colors. The AGN fraction varies in the range OF $= 20-40\%$ for their fiducial model and $M_\star = 10^9 - 10^{10.5}\,\mathrm{M}_\odot$, and OF $= 10-25\%$ in the same galaxy mass range when adopting other color criteria that have been used in the literature. An estimate of the hard X-ray luminosity of these AGN is $L_{\mathrm{AGN}} \leqslant 10^{42}$ erg/s (S. Kaviraj, private communication).

Lupi, Sbarrato & Carniani (2020) recently used the same sample of Kaviraj, Martin & Silk (2019) to find that 7 to 20 candidates could be AGN, over 5000 galaxies with $M_\star = 10^8 - 10^9\,\mathrm{M}_\odot$. This corresponds to an OF of $0.14-0.4\,\%$ for AGN with estimated bolometric luminosities of $L_{\mathrm{bol,\,AGN}} \leqslant 3\times10^{44}$ erg/s. This is lower than the $\geqslant 10\,\%$ fractions found in Kaviraj, Martin & Silk (2019). Lupi, Sbarrato & Carniani (2020) convoluted both optical and mid-IR data and took into account the contamination from close sources, the relatively low resolution of the WISE mid-IR instrument that was used (with respect to for example optical surveys), and the contamination from the galaxies. The difference among the two studies demonstrates the difficulty of using mid-IR selection to identify AGN in these galaxies.

Birchall, Watson & Aird (2020) investigate the fraction of AGN (detected in X-ray, 2-12 keV) in different stellar mass ranges of galaxies in the local Universe ($z \leqslant 0.25$). For $M_\star = 10^7 - 10^9\,\mathrm{M}_\odot$, the AGN fraction is OF $\sim 0.19\%$ for $L_{\mathrm{x,\,AGN,\,2-12\,keV}} = 10^{41.5} - 10^{42.4}$ erg/s, and OF $\sim 0.1\%$ for $L_{\mathrm{x,\,AGN,\,2-12\,keV}} \geqslant 10^{42.4}$ erg/s. For more massive galaxies with $M_\star = 10^9 - 10^{9.5}\,\mathrm{M}_\odot$, Birchall, Watson & Aird (2020) find a fraction of OF $\sim 0.17\%$ and OF $\sim 0.06\%$, for AGN with $L_{\mathrm{x,\,AGN,\,2-12\,keV}} = 10^{41.5} - 10^{42.4}$ erg/s and $L_{\mathrm{x,\,AGN,\,2-12\,keV}} \geqslant 10^{42.4}$ erg/s respectively.

Mezcua & Domínguez Sánchez (2020) identify 37 galaxies over a sample of 1609 dwarf galaxies with $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$ with signatures consistent with AGN ionisation. This results in a fraction of OF = 2.3% for AGN in the range $L_{\mathrm{bol,\,AGN}} = 10^{38.9} - 10^{41.4}$ erg/s. We convert these values into hard X-ray luminosities using the bolometric correction of Hopkins, Richards & Hernquist (2007b).

Latimer et al. (2021) recently studied the AGN fraction in galaxies with $M_\star \leqslant 3\times10^9\,\mathrm{M}_\odot$ using X-ray observations from the eROSITA eFEDS survey, for $z \leqslant 0.15$. The authors derived an upper limit to the AGN OF of the sample of $\leqslant 1.8\,\%$. The sources have X-ray luminosities in the range $L_{\mathrm{x,\,gal,\,2-10\,keV}} = 10^{38.5} - 10^{40}$ erg/s.

Finally, Birchall et al. (2022) recently reported the AGN fraction in stellar mass bins for luminous AGN with $L_{\mathrm{x,\,AGN,\,2-10\,keV}} = 10^{42} - 10^{44}$ erg/s derived from the MPA-JHU catalogue (based on SDSS DR8) and 3XMM. The AGN fraction increases towards more massive galaxies, from $3-9\times10^{-4}$ in galaxies with $\leqslant 10^{9.5}\,\mathrm{M}_\odot$ to $2-6\times10^{-3}$ in $\sim 10^{10.5}\,\mathrm{M}_\odot$ galaxies. At fixed stellar mass, the fraction also slightly increases towards higher redshifts in the range $z = 0-0.3$.

To summarize, the measured AGN OF depends on the different studies, selection techniques, and sample definitions. Most of the constraints listed in this section span over an order of magnitude or more, but in general favor an OF of $\leqslant 3\times10^{-2}$ in galaxies of $M_\star = 10^{8.5} - 10^{10}\,\mathrm{M}_\odot$. We emphasize here that only a few of these analyses include a correction for galaxy completeness (Aird et al. 2012; Miller et al. 2015; Pardo et al. 2016; Aird, Coil & Georgakakis 2018; Mezcua et al. 2018; Birchall, Watson & Aird 2020), that can mostly be done with deep X-ray surveys. In theory, deriving an OF requires to weigh each AGN detection by the completeness of the corresponding bin (defined by a redshift, galaxy stellar mass, and X-ray luminosity). The weighting function thus depends on the total number of galaxies which should be in the bin, the X-ray sensitivity of the survey (i.e., what is the probability of detecting an AGN of a given luminosity in the survey, the probability varies spatially). We refer the reader to the Section of Aird et al. (2012) for an in-depth description of the methodology.

APPENDIX B: SIMULATED AGN IN THE LOW-MASS GALAXY REGIME

Besides having different AGN occupation fractions at $z = 0$, the simulations have significant differences in their AGN populations. We show the histograms of the AGN luminosities in Fig. B1 for AGN in galaxies with total stellar mass of $\sim 10^9$, $\sim 10^{9.5}$, and $10^{10}\,\mathrm{M}_\odot$. The AGN luminosity distribution in galaxies with $M_\star \sim 10^{10}\,\mathrm{M}_\odot$ peaks at $\log_{10} L_{\mathrm{x}}/(\mathrm{erg/s}) \sim 41$ in SIMBA (median $\log_{10} L_{\mathrm{x}}/(\mathrm{erg/s}) \sim 39$), and at $\log_{10} L_{\mathrm{x}}/(\mathrm{erg/s}) \sim 42$ for Horizon-AGN (median $\log_{10} L_{\mathrm{x}}/(\mathrm{erg/s}) \sim 42$), TNG100

(median $\log_{10} L_{\rm x}/({\rm erg/s}) \sim 40$), and TNG300. However, we find that the EAGLE and Illustris simulations have fainter AGN populations in these intermediate-mass galaxies. In the dwarf regime, i.e. for galaxies with $M_\star \sim 10^9\,{\rm M}_\odot$, AGN are on average fainter for most of the simulations. In Horizon-AGN, the AGN luminosity distribution has a median of $\log_{10} L_{\rm x}/({\rm erg/s}) \sim 37$, $\log_{10} L_{\rm x}/({\rm erg/s}) \sim 31$ in EAGLE, and $\log_{10} L_{\rm x}/({\rm erg/s}) \sim 37$ in Illustris. However, we find that the AGN in the dwarf galaxies produced by TNG100 still have similar luminosities (median $\log_{10} L_{\rm x}/({\rm erg/s}) \sim 40$) than in more massive galaxies with $M_\star \sim 10^{10}\,{\rm M}_\odot$.

Assessing the agreement between the AGN luminosity functions produced by the simulations and those constrained by observations is crucial. A comparison is presented in Habouzit et al. (2022b) and shows that at $z = 0$ the faint end of the hard X-ray luminosity function ($L_{\rm x} \leqslant 10^{43}\,{\rm erg/s}$) of Horizon-AGN, Illustris, and SIMBA are within the constraints of Buchner et al. (2015) and in agreement with those of Hopkins, Richards & Hernquist (2007a); Aird et al. (2015). The function produced by TNG100 is above the constraints, and the one from EAGLE slightly below. A high number of faint AGN at $z = 0$ are located in low-mass galaxies in TNG100 and could explain the higher luminosity function in case of too massive BHs, and/or too efficiently accreting BHs, and/or not strong enough SN feedback compared to observations. In the opposite, the lower luminosity function produced in EAGLE could be explained by lower-mass BHs than in observations (a priori not the case, see Fig. 3) and/or too weak accretion onto the BHs and/or too efficient SN feedback in low-mass galaxies Habouzit et al. (2022b).

APPENDIX C: CONTRIBUTORS TO GALAXY X-RAY LUMINOSITY

In this paper, we use the hard X-ray luminosity of the simulated AGN to assess whether they could be detected, and whether the XRB population could outshine them. However, observations often rely on a galaxy-wide X-ray emission, which is the sum of the AGN emission, the XRB population, and the ISM host gas. Using the X-ray stacking technique, Mullaney et al. (2012) and Aird, Coil & Georgakakis (2018) compute the average AGN luminosity in star-forming galaxies at $z \sim 1$ and $z \sim 2$ and find that the $< L_{AGN} > -M_\star$ relation behave linearly (also referred to as "AGN main sequence"). This is not always true for low-mass galaxies, where Aird, Coil & Georgakakis (2018) find that AGN activity is significantly suppressed. In the following, we show that, in the absence of obscuration, the emission from the AGN in the low-mass end $M_\star \leqslant 10^{10.5}\,{\rm M}_\odot$ is on average larger than the XRBs and ISM for most of the simulations. The average contributions of the AGN, XRB population, and ISM to the galaxy-wide X-ray luminosity are shown in Fig. C1.

The contribution of the AGN differs from a simulation to another. AGN in EAGLE are on average the faintest followed by Illustris, both which have a median hard X-ray luminosity $L_{\rm AGN} \leqslant 10^{38}\,{\rm erg/s}$. The relative contribution of the XRB population is higher for these simulations at all stellar masses, and is higher in the other simulations in regimes where the AGN luminosity is lower: typically dwarf galaxies with $\sim 10^9\,{\rm M}_\odot$ (e.g., Horizon, TNG300), and more massive galaxies with $M_\star \geqslant 10^{10.2}\,{\rm M}_\odot$ in TNG when the efficient kinetic AGN feedback mode kicks out. In these regimes, the XRB population is able to outshine the AGN. We find that the contribution from ISM to the galaxy-wide X-ray luminosity is very small for all the simulations.

APPENDIX D: AGN AND GALAXY-WIDE XRB LUMINOSITIES

In Table D1, we provide three different percentages for three different galaxy types (starburst, main sequence and quenched), and for three stellar mass bins: $\sim 10^9, 10^{9.5}$ and $10^{10}\,{\rm M}_\odot$:

- The percentage of galaxy types in each simulation, referred as "starburst", "Main Seq" and "Quenched" in Table D1. This is defined as the number of galaxies of a given galaxy type in a given stellar mass bin, divided by the total number of galaxies in that bin.
- The percentage of galaxies, within a given galaxy type, hosting an AGN with $L_{\rm AGN} \geqslant 10^{38}\,{\rm erg/s}$.
- The percentage of galaxies, within a given galaxy type, that host an AGN with $L_{\rm AGN} \geqslant 10^{38}\,{\rm erg/s}$ and satisfy $L_{\rm AGN} > L_{\rm XRB}$ (i.e., an AGN that outshines the galaxy-wide XRB population).

Our definitions of the galaxy types are simulation dependent. We also provide the percentages for AGN with $L_{\rm AGN} \geqslant 10^{40}\,{\rm erg/s}$.

APPENDIX E: OBSERVATIONAL SAMPLES PRESENTED IN FIG. 7

We describe in this section the observational samples presented in Fig. 7. We use the nearby host galaxies ($z < 0.5$) of Mezcua et al. (2018, see Appendix A for details) where the AGN luminosity ranges from $L_{\rm x,\,AGN,\,0.5\text{–}10keV} \sim 10^{39} - 10^{43}\,{\rm erg/s}$ and the SFR of their hosts from $\log_{10}{\rm SFR}/({\rm M}_\odot/{\rm yr}) = [-2.63, 1.11]$. We also show 37 AGN candidates of Mezcua & Domínguez Sánchez (2020), identified in galaxies with SFR $\log_{10}{\rm SFR}/({\rm M}_\odot/{\rm yr}) = -2.5, 1.5$. For a fair comparison, we have converted their AGN luminosities from bolometric to hard (2-10 keV) X-ray (using the correction of Hopkins, Richards & Hernquist 2007a).

We also include the sample of Birchall, Watson & Aird (2020) presenting 61 dwarf galaxies as potential AGN hosts with SFR in the range $\log_{10}{\rm SFR}/({\rm M}_\odot/{\rm yr}) = -3.8, 1.2$. The X-ray luminosity of this work are in the $2 - 12$ keV band and for the galaxy as a whole. To derive the luminosities of the AGN (as shown in Fig. 7), we subtract the predicted contribution of the XRB population (using the relation of Lehmer et al. 2016) and hot gas from the ISM (using the relation of Mineo, Gilfanov & Sunyaev 2012) from the X-ray galaxy luminosity given in their sample.

For our analysis, we search for X-ray counterparts to the AGN of Martín-Navarro, Shankar & Mezcua (2021) using the Chandra Source Catalog (Evans et al. 2010) version 2 (CSC 2.0) Master Source Catalog. We use a search radius of 5 arcsec, which returns 50 X-ray detections. For these, we take from CSC 2.0 the aperture-corrected net energy

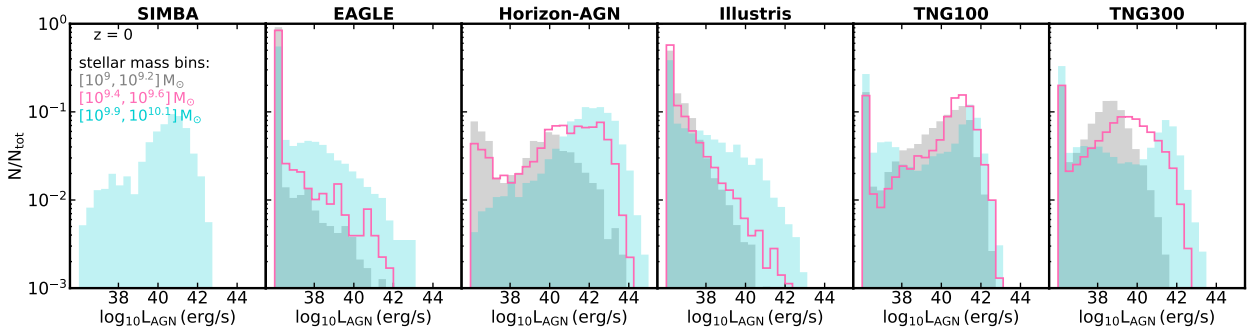

Figure B1. The normalised distribution of X-ray AGN luminosity at $z = 0$ galaxies with $M_\star \sim 10^9$, $\sim 10^{9.5}$, $\sim 10^{10}\,\mathrm{M}_\odot$ for all the six simulations. For SIMBA, we do not display the BH population for $M_\star < 10^{9.5}\mathrm{M}_\odot$ since the simulation starts seeding BHs at the aforementioned value.

Table D1. Percentages of galaxy types (starburst, main sequence and quenched) at $z = 0$ in three different stellar bins: $\log_{10} M_\star/\mathrm{M}_\odot \sim$ 9, 9.5, 10. For each galaxy type, we derived the percentage of galaxies that host an AGN with a hard X-ray luminosity $L_{\mathrm{AGN}} \geqslant 10^{38}$ erg/s and $L_{\mathrm{AGN}} \geqslant 10^{40}$ erg/s. We also add the percentages of galaxies with an AGN outshining the galaxy-wide XRB population.

$\log_{10} M_\star/\mathrm{M}_\odot$	% of	EAGLE	SIMBA	Horizon-AGN	Illustris	TNG100	TNG300
9-9.2	Starburst	6	-	4	2	4	4
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	6	-	2	30	89	80
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	3	-	1	10	80	42
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	2	-	1	7	78	27
	$L_{\mathrm{AGN}} > 10^{40}$ and AGN $> L_{\mathrm{XRB}}$	2	-	1	7	78	27
	Main Seq.	61	-	75	86	74	70
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	4	-	3	6	82	64
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	2	-	2	0	70	22
	AGN host($L_{\mathrm{AGN}} > 10^{40}$)	1	-	2	1	50	5
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	1	-	2	1	50	5
	Quenched	33	-	21	12	22	26
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	0	-	3	0	1	0
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	0	-	2	0	0	0
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	0	-	1	0	0	0
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	0	-	1	0	0	0
9.4-9.6	Starburst	5	-	4	1	2	5
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	12	-	7	38	93	87
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	6	-	7	20	80	60
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	5	-	7	21	85	57
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	5	-	7	21	85	56
	Main Seq.	73	-	82	91	78	71
	AGN host($L_{\mathrm{AGN}} > 10^{38}$)	9	-	18	10	88	74
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	4	-	16	2	70	46
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	3	-	15	2	69	29
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	3	-	15	2	69	29
	Quenched	23	-	13	8	20	24
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	0	-	12	0	3	1
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	0	-	10	0	0	0
	AGN host($L_{\mathrm{AGN}} > 10^{40}$)	0	-	6	0	0	0
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	0	-	6	0	0	0
9.9-10.1	Starburst	2	4	4	1	3	5
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	27	99	28	75	100	92
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	18	98	7	50	90	76
	AGN host($L_{\mathrm{AGN}} > 10^{40}$)	18	98	27	58	93	81
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} \geq L_{\mathrm{XRB}}$	5	98	27	42	90	76
	Main Seq.	76	59	87	92	79	73
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	19	95	50	22	93	85
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	7	64	17	6	71	69
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	6	64	45	6	74	65
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} \geq L_{\mathrm{XRB}}$	2	64	45	6	74	64
	Quenched	22	38	9	8	18	22
	AGN host ($L_{\mathrm{AGN}} > 10^{38}$)	3	30	27	0	4	7
	$L_{\mathrm{AGN}} > 10^{38}$ and $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$	2	14	10	0	0	1
	AGN host ($L_{\mathrm{AGN}} > 10^{40}$)	2	6	17	0	2	1
	$L_{\mathrm{AGN}} > 10^{40}$ and $L_{\mathrm{AGN}} \geq L_{\mathrm{XRB}}$	0	5	17	0	2	0

flux, inferred from the source region aperture, in the ACIS $2-7.0$ keV energy band. This flux corresponds to the best estimate derived from the longest block of a multi-band, flux-ordered Bayesian Block analysis of the contributing observations. The K-corrected X-ray luminosity in the $2-10$ keV band is derived from this aperture-corrected flux assuming a typical AGN photon index of 1.8.

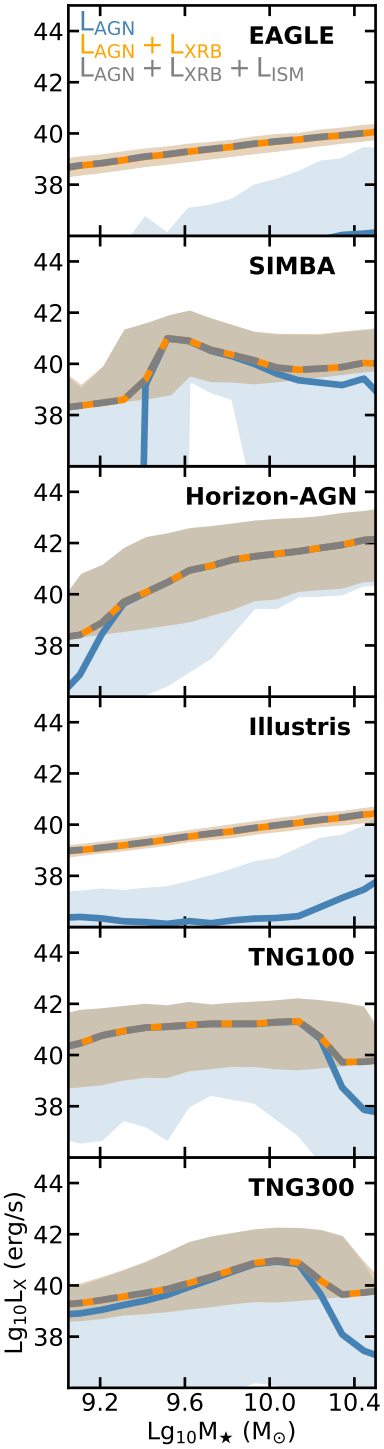

Figure C1. Median of the hard X-ray luminosity as function of stellar mass in the low-mass regime. We display three contributors to the galaxy X-ray luminosity: $L_{\rm AGN}$ (blue), $L_{\rm AGN}+L_{\rm XRB}$ (orange) and $L_{\rm AGN}+L_{\rm XRB}+L_{\rm ISM}$ (grey). The 15-85th percentiles are illustrated by the shaded areas.