Bringing faint active galactic nuclei (AGNs) to light: a view from large-scale cosmological simulations

Adrian P. Schirra,[1][★] Mélanie Habouzit,[1,2] Ralf S. Klessen [ORCID],[1,3] Francesca Fornasini [ORCID],[4,5] Dylan Nelson [ORCID],[1,6] Annalisa Pillepich [ORCID],[2] Daniel Anglés-Alcázar,[7,8] Romeel Davé [ORCID][9] and Francesca Civano[4]

[1]*Zentrum für Astronomie, Institut für Theoretische Astrophysik, Universität Heidelberg, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany*
[2]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[3]*Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany*
[4]*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[5]*Stonehill College, 320 Washington Street, North Easton, MA 02357, USA*
[6]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, D-85740 Garching bei München, Germany*
[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]*Institute for Astronomy, Royal Observatory, University of Edinburgh, Edinburgh EH9 3HJ, UK*

ABSTRACT
The sensitivity of X-ray facilities and our ability to detect fainter active galactic nuclei (AGNs) will increase with the upcoming Athena mission and the AXIS and Lynx concept missions, thus improving our understanding of supermassive black holes (BHs) in a luminosity regime that can be dominated by X-ray binaries. We analyse the population of faint AGNs ($L_{x,2-10\,\mathrm{keV}} \leqslant 10^{42}\,\mathrm{erg\,s^{-1}}$) in the Illustris, TNG100, EAGLE, and SIMBA cosmological simulations, and find that the properties of their host galaxies vary from one simulation to another. In Illustris and EAGLE, faint AGNs are powered by low-mass BHs located in low-mass star-forming galaxies. In TNG100 and SIMBA, they are mostly associated with more massive BHs in quenched massive galaxies. We model the X-ray binary (XRB) populations of the simulated galaxies, and find that AGNs often dominate the galaxy AGN + XRB hard X-ray luminosity at $z > 2$, while XRBs dominate in some simulations at $z < 2$. Whether the AGN or XRB emission dominates in star-forming and quenched galaxies depends on the simulations. These differences in simulations can be used to discriminate between galaxy formation models with future high-resolution X-ray observations. We compare the luminosity of simulated faint AGN host galaxies to observations of stacked galaxies from *Chandra*. Our comparison indicates that the simulations post-processed with our X-ray modelling tend to overestimate the AGN + XRB X-ray luminosity; luminosity that can be strongly affected by AGN obscuration. Some simulations reveal clear AGN trends as a function of stellar mass (e.g. galaxy luminosity drop in massive galaxies), which are not apparent in the observations.

Key words: methods: numerical – black hole physics – galaxies: evolution – galaxies: formation – galaxies: active.

1 INTRODUCTION

Some galaxies show a much broader energy distribution than some others. They are powered by active galactic nuclei (AGNs), i.e. by supermassive black holes (BHs) accreting efficiently at their centres. Evidence shows that BHs could be present in the centre of most massive galaxies around us (Gültekin et al. 2009; Kormendy & Ho 2013; Greene, Strader & Ho 2020, and references therein). The population of AGNs is constrained by the luminosity function of the AGN, e.g. in X-ray (Aird et al. 2015; Georgakakis et al. 2015). However, these constraints from observations do not cover the faint regime defined by hard X-ray (2–10 keV) AGN luminosity of $L_{\mathrm{AGN}} \leqslant 10^{42}\,\mathrm{erg\,s^{-1}}$. The properties of the faint AGN population are especially hard to determine because their luminosities are in

a regime potentially dominated by X-ray binaries (XRBs). In this paper, we aim at constraining the population of faint AGNs, and the relative contribution of the X-ray emission from the AGN and from the XRB population in galaxies of different masses in large-scale cosmological simulations.

XRBs are formed by a compact object, stellar mass BH or neutron star, accreting from a normal star. In a binary system, the mass is transferred from the star to the compact object via Roche lobe overflow or stellar wind mass transfer. The number and luminosity of XRBs are thought to primarily depend on the host galaxy properties, such as mass and specific star formation rate (sSFR; Lehmer et al. 2016, 2019; Fornasini et al. 2018). The galaxy-wide emission from the binary population comes from both high-mass XRBs (HMXBs) and low-mass XRBs (LMXBs). The stellar companion in LMXBs is less massive than the compact object, and LMXB lifetime is longer than 1 Gyr (Lehmer et al. 2010). HMXBs are composed of a neutron star or stellar mass BH and a more massive stellar companion than

[★] E-mail: a.schirra@stud.uni-heidelberg.de

in LMXBs. As such, their lifetime is shorter than 100 Myr. The galaxy-wide number of LMXBs increases with the galaxy stellar mass.

A scaling of the X-ray emission with the stellar mass of the galaxies and their star formation rate (SFR) was found in several analyses of nearby galaxies ($D \leqslant 50$ Mpc; e.g. Grimm, Gilfanov & Sunyaev 2003; Colbert et al. 2004; Gilfanov 2004; Lehmer et al. 2010; Boroson, Kim & Fabbiano 2011; Mineo, Gilfanov & Sunyaev 2012a, b; Zhang, Gilfanov & Bogdán 2012). The most recent scaling relation was derived in Lehmer et al. (2019). In this work, galaxies with X-ray measurements from the *Chandra* data are decomposed into spatially resolved maps of SFR and stellar mass, and the XRB population statistics is extracted for sSFR bins. The parameters of the scaling relation functional form are fitted in these sSFR bins (i.e. fits over the pixels of all the galaxies together, and galaxy by galaxy). The result scaling relation is then tested by fitting each galaxy of the sample. These studies revealed that the total XRB emission of quiescent galaxies is dominated by LMXBs, for which the emission is proportional to the galaxy mass since LMXBs are long lived. However, the total XRB emission of star-forming galaxies is dominated by young short-lived luminous HMXBs (Fabbiano 2006), and the higher the galaxy SFR the higher the total XRB emission. The X-ray emission from LMXBs and HMXBs is comparable for sSFR $\approx 10^{-10}$ yr^{-1} (Lehmer et al. 2019).

The X-ray emission of the XRB population could also depend on the metallicity and stellar ages of the galaxies (Fragos et al. 2013; Lehmer et al. 2016; Madau & Fragos 2017). Studies focusing on more distant and diverse galaxies have found deviations from the empirical relations established in the local Universe. This is the case for samples of low-metallicity galaxies or for samples with a wider range of stellar masses (e.g. Kim & Fabbiano 2010; Lehmer et al. 2014; Basu-Zych et al. 2013a, 2016; Douna et al. 2015; Brorby 2016; Tzanavaris et al. 2016). Finally, an evolution with redshift was found in X-ray stacking analyses (e.g. Lehmer et al. 2007, 2016; Basu-Zych et al. 2013b; Kaaret 2014; Aird, Coil & Georgakakis 2017). The emission from LMXBs per unit stellar mass is thought to be higher at higher redshifts. The donor stars of the LMXBs have higher masses at higher redshifts and the LMXBs are more luminous (Lehmer et al. 2016). Moreover, elliptical galaxies with many globular clusters show an excess of LMXBs. Globular clusters can produce LMXBs efficiently through stellar dynamical interactions (Cheng et al. 2018; Lehmer et al. 2019), i.e. by either capture of a stellar object by the second object, or hardening of soft binaries (hard binaries are binaries with bound energy larger than the kinetic energy of the intruding star and stellar encounters involving hard binaries make them harder, see Cheng et al. 2018). The HMXB emission per unit SFR is higher at higher redshifts. This is connected to the lower metallicities of the stellar populations, as observationally shown in Fornasini et al. (2019) and Fornasini, Civano & Suh (2020). The star formation at low metallicities is expected to result in a larger number of massive compact objects and Roche lobe overflow binaries because the mass-loss of massive stars through stellar winds is less effective (Dray 2006; Linden et al. 2010; Fragos et al. 2013; Lehmer et al. 2016).

There is evidence for the presence of X-ray AGN in all types of galaxies, including dwarf galaxies ($M_\star \leqslant 10^{9.5}$ M$_\odot$; Mezcua et al. 2016; Chilingarian et al. 2018), where the galaxy-wide X-ray emission is higher than expected from the XRB population. In Mezcua & Domínguez Sánchez (2020), the authors identify AGN signatures in 37 local dwarf galaxies in the range $\log_{10} M_\star/$M$_\odot = 8.8 - 9.5$ using integral field unit (IFU) spectroscopy. These AGNs are faint, sometimes with off-centre X-ray emission, and have bolometric luminosity

in the range $\log_{10} L_{\rm bol}/({\rm erg/s}) = 38.9 - 41.4$. BH mass estimates for these AGNs could lie in the range $\log_{10} M_{\rm BH}/$M$_\odot = 5.5 - 8$. The presence of AGNs in dwarf galaxies could also shed new light on the role of AGN feedback in these galaxies, as studied in simulations (Koudmani et al. 2019; Koudmani, Henden & Sijacki 2021), and observed with the detections of fast AGN-driven outflows (Manzano-King, Canalizo & Sales 2019; Liu et al. 2020).

Signatures of faint AGNs have been identified in more massive galaxies than dwarf galaxies, particularly with stacking analysis of *Chandra* COSMOS Legacy galaxies (Georgantopoulos et al. 2017; Fornasini et al. 2018). The BHs powering these faint AGNs can have low accretion rates for different reasons: for example, BH accretion rates regulated by BH feedback, BH accretion rates regulated by the feedback in the environment of the BHs, BHs in quenched galaxies depleted in gas, off-centre BHs that do not benefit from the galaxy potential well gas reservoir, and the accretion rate likely depends on BH mass and galaxy properties. Constraining the regime of faint AGNs will help to understand the co-evolution and interplay between BHs and their host galaxies. In this study, we will investigate in which galaxies faint AGNs reside in several cosmological simulations, and whether the low accretion rates on to the simulated faint AGNs can have different causes (e.g. impact of AGN feedback, absence of cold gas), such as outlined above.

The galaxy-wide emission of the XRB population could depend on the SFR of the host galaxies, and in a similar way correlations between AGN activity and the SFR of their galaxies could exist. Even for the regime of the bright AGN (e.g. $L_{\rm AGN} \geqslant 10^{42}$ erg s^{-1}), the correlations between AGN activity and the properties of galaxies are still not clear, especially in terms of stellar masses, SFR, and sSFR (e.g. Aird, Coil & Georgakakis 2019). For bright AGN of $L_{\rm AGN} \geqslant 10^{44}$ erg s^{-1} in massive galaxies, there is evidence for a correlation between AGN luminosity and SFR (Lutz et al. 2008; Bonfield et al. 2011; Mor et al. 2012; Rosario et al. 2012). There are very few observational constraints in the faint AGN regime, and therefore, in order to theoretically investigate possible dependences/correlations of our results with galaxy properties, we will split the simulated galaxies in starburst, star-forming, or quiescent galaxies.

We are also particularly interested in how we can compare the predictions from cosmological simulations to current and future observations to disentangle simulation subgrid models and improve them. In this work, we compare the total X-ray luminosity of the simulated galaxies hosting faint AGNs to observational constraints of stacked galaxies (Fornasini et al. 2018). The high sensitivity of the upcoming X-ray mission Athena (Nandra et al. 2013) and the NASA concept missions Lynx (Gaskin et al. 2018) and AXIS (Mushotzky et al. 2019) will allow us to investigate the properties of AGNs up to high redshift, and will provide us with new constraints on the faint AGN population. Our work paves the way for future investigations on faint AGNs, a goal aligned with the upcoming X-ray missions. As in all studies on the AGN population, obscuration is critical and a major open question. The attenuation of radiation from intervening gas and dust can affect the contribution of both the AGN and the XRBs to the observed galaxy X-ray luminosity. We will build several models for the obscuration of the faint AGN to account for this.

In this paper, we use the four Illustris, TNG100, EAGLE, and SIMBA cosmological hydrodynamical simulations of $\geqslant 100$ comoving Mpc (cMpc) box side length. For all these simulations, we model both the AGN luminosity and the XRB population of the simulated galaxies using several empirical scaling relations. We describe the physical models of the simulations, as well as our post-processing modelling of AGN and XRB luminosity in Section 3. In Section 4, we analyse the AGN population of the simulations and in which

galaxies they live. In order to derive the total X-ray luminosity of the simulated galaxies, we first analyse in Section 5 the properties of their AGN as a function of their host galaxy's mass and SFR. In Sections 6 and 7, we derive the relative contribution of the AGN and the XRB population to the galaxy total hard X-ray luminosity and we compare these results to recent findings in observations of stacked galaxies from the *Chandra* COSMOS Legacy survey (Fornasini et al. 2018).

2 COSMOLOGICAL SIMULATIONS AND METHODS

In this section, we describe the Illustris, TNG100, EAGLE, and SIMBA cosmological simulations, and our method to compute AGN luminosity and the X-ray luminosity of the XRB population of the simulated galaxies. We also describe our method to divide the simulated galaxy in three samples with different sSFR.

2.1 Cosmological simulations

In the following, we briefly summarize the BH subgrid models, i.e. seeding, accretion, and feedback of the Illustris, TNG100, EAGLE, and SIMBA large-scale cosmological simulations. A more detailed comparison of the modelling of these simulations is presented in Habouzit et al. (2021). Several other aspects differentiate these simulations, for example the presence of magnetic fields in the TNG100 simulation (Pillepich et al. 2018b), the single mode AGN feedback of EAGLE (Schaye et al. 2015), the two mode accretion model of SIMBA (Davé et al. 2019). The detailed models can be found in the references given below.

2.1.1 Illustris

The Illustris hydrodynamical simulation (Genel et al. 2014; Vogelsberger et al. 2014a, b; Sijacki et al. 2015) simultaneously follows the evolution of dark matter (DM) and baryonic matter, in a volume of $(106.5 \, \mathrm{cMpc})^3$. The Illustris simulation data are publicly available (Nelson et al. 2015). The simulation produces a mix of galaxy morphologies in broad agreement with observations (Vogelsberger et al. 2014a). The equations of gravity and hydrodynamics are evolved with the moving-mesh code AREPO (Springel 2010). The galaxy formation model includes gas cooling, star formation, supernova (SN) feedback, and the physics of BHs. BHs are seeded in haloes with halo mass $\geqslant 7.1 \times 10^{10} \, \mathrm{M_\odot}$ with seed mass $M_\mathrm{seed} = 1.4 \times 10^5 \, \mathrm{M_\odot}$. BHs can grow in mass with gas accretion and BH mergers. The BH accretion is modelled with the Bondi–Hoyle–Lyttleton formalism as

$$\dot{M}_\mathrm{BH} = \min[\alpha \dot{M}_\mathrm{Bondi}, \dot{M}_\mathrm{Edd}], \tag{1}$$

with

$$\dot{M}_\mathrm{Bondi} = \frac{4\pi G^2 M_\mathrm{BH}^2 \rho}{(c_s^2 + v_\mathrm{BH}^2)^{3/2}}, \tag{2}$$

where ρ and c_s are the density and the sound speed of the surrounding gas, respectively, v_BH is the velocity of the BH relative to the gas, and G is the gravitational constant. The boost factor $\alpha = 100$ accounts for the unresolved gas around the AGN, which tend to underestimate the accretion rates. The boost factor is set to produce a BH population that agrees with the M_BH–$M_\star$ diagram at $z = 0$. A repositioning scheme for BH sink particles is used which ties them to the local minimum gravitational potential. $\dot{M}_\mathrm{Edd}$ is the Eddington accretion

rate of the BH:

$$\dot{M}_\mathrm{Edd} = \frac{4\pi G m_\mathrm{p}}{\sigma_T \, c \, \epsilon_\mathrm{r}} M_\mathrm{BH}, \tag{3}$$

where m_p is the proton mass, σ_T is the Thomson cross-section, c is the speed of light, and ϵ_r is the radiative efficiency which is set to 0.2 in the simulation. The Eddington ratio is defined as $f_\mathrm{Edd} = \dot{M}_\mathrm{BH}/\dot{M}_\mathrm{Edd}$. The feedback from the AGN depends on the BH accretion rate and can operate in *high-accretion mode* ($f_\mathrm{Edd} > 0.05$) or *low-accretion mode* ($f_\mathrm{Edd} < 0.05$). The low-accretion mode model injects highly bursty thermal energy into large 'bubbles' ($\sim$50 kpc) which are displaced away from the central galaxy. In the high-accretion mode, thermal energy is continuously injected into the surrounding gas. Moreover, radiative AGN feedback that modifies the ionization state of the surrounding gas is included as well (Vogelsberger et al. 2013).

2.1.2 TNG100

The TNG100 simulation builds directly on the Illustris framework and has a volume of $(110.7 \, \mathrm{cMpc})^3$ (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019b; Springel et al. 2018; Pillepich et al. 2018a). The simulation includes magnetic fields. The BH seed mass is $M_\mathrm{seed} = 1.1 \times 10^6 \, \mathrm{M_\odot}$. The BH seed mass was increased in comparison with the Illustris seed mass by nearly one order of magnitude (Pillepich et al. 2018b). The BH accretion rate is described by the Bondi–Hoyle–Lyttleton accretion rate limited by the Eddington accretion rate (Weinberger et al. 2017). However, in the TNG model, the additional boost factor α from equation (1) is removed. The effective sound speed $c_s^2 = c_{s,\mathrm{therm}}^2 + (B^2/4\pi\rho)$ includes the thermal and the magnetic signal propagation. The transition between the AGN feedback modes in the TNG model depends on the BH accretion rate and BH mass. A BH is assumed to be in the high accretion state as long as (Weinberger et al. 2017)

$$f_\mathrm{Edd} \geq \min\left[2 \times 10^{-3}\left(\frac{M_\mathrm{BH}}{10^8 \mathrm{M_\odot}}\right)^2, 0.1\right]. \tag{4}$$

The low accretion feedback is modelled as kinetic outflows from the BHs (Weinberger et al. 2017), while the high accretion mode is modelled as injection of thermal energy in the BH surroundings (Weinberger et al. 2017).

2.1.3 EAGLE

The EAGLE simulation (Crain et al. 2015; Schaye et al. 2015) has a volume of $(100 \, \mathrm{cMpc})^3$, and was run with the code ANARCHY. This modified version of GADGET3 (Springel 2005) is based on the smoothed particle hydrodynamics (SPH) method. The simulation includes subgrid models for radiative cooling, star formation, stellar mass-loss, metal enrichment, and energy feedback from star formation (Schaye et al. 2015). The BH seeding mass is $M_\mathrm{seed} = 1.48 \times 10^5 \, \mathrm{M_\odot}$ and BHs are seeded in haloes with a mass of at least $M_\mathrm{h} = 1.48 \times 10^{10} \, \mathrm{M_\odot}$. The accretion rate is determined by the Bondi–Hoyle–Lyttleton model (Rosas-Guevara et al. 2015, 2016):

$$\dot{M}_\mathrm{acc} = \min(\dot{M}_\mathrm{Bondi} \times \min\left((c_s/V_\Phi)^3/C_\mathrm{visc}, 1\right), \dot{M}_\mathrm{Edd}), \tag{5}$$

with $\dot{M}_\mathrm{Bondi}$ and $\dot{M}_\mathrm{Edd}$ defined as in equations (2) and (3) and $\dot{M}_\mathrm{BH} = (1 - \epsilon_r)\dot{M}_\mathrm{acc}$. The radiative efficiency is $\epsilon_r = 0.1$. Here, c_s represents the sound speed of the surrounding gas, V_Φ is the rotation speed of the gas around the BH, and C_visc is a free parameter. It is related to the viscosity of the accretion disc. The correction factor $\min\left((c_s/V_\Phi)^3/C_\mathrm{visc}, 1\right)$ multiplied with the Bondi rate accounts for a lower accretion rate for gas with angular momentum. In that case,

the accretion is not spherically symmetric and proceeds through an accretion disc (Schaye et al. 2015). The simulation uses a single-mode AGN feedback model. A fraction of the accreted gas on to the BHs is stochastically injected in the surroundings as thermal energy (Schaye et al. 2015).

2.1.4 SIMBA

SIMBA (Davé et al. 2019) has a volume of $(147\,\mathrm{cMpc})^3$ and builds on its predecessor Mufasa (Davé, Thompson & Hopkins 2016), using the code GIZMO (Hopkins 2015, 2017) in its 'meshless finite mass' hydro-dynamics mode. The BH seeding mass is $M_{\mathrm{seed}} = 1.4 \times 10^4\,\mathrm{M_\odot}$ and BHs are seeded in galaxies with $M_\star \sim 10^{9.5}\,\mathrm{M_\odot}$. This is motivated by FIRE simulations showing that stellar feedback strongly suppresses BH growth in lower mass galaxies (Anglés-Alcázar et al. 2017b; Çatmabacak et al. 2020), in agreement with other models (e.g. Habouzit, Volonteri & Dubois 2017). Because of the seeding, we will not draw any conclusion on BH properties in galaxies with $M_\star < 10^{9.5}\,\mathrm{M_\odot}$ for the SIMBA simulation. The BH accretion rate is given by

$$\dot{M}_{\mathrm{BH}} = (1 - \epsilon_{\mathrm{r}})\left[\min(\dot{M}_{\mathrm{Bondi}}, \dot{M}_{\mathrm{Edd}}) + \min(\dot{M}_{\mathrm{Torque}}, 3\,\dot{M}_{\mathrm{Edd}})\right],\tag{6}$$

with $\dot{M}_{\mathrm{Bondi}}$ and $\dot{M}_{\mathrm{Edd}}$ defined as in equations (2) and (3). The radiative efficiency is $\epsilon_{\mathrm{r}} = 0.1$. $\dot{M}_{\mathrm{Torque}}$ describes the gas inflow rate driven by gravitational instabilities from the scale of the galaxy to the accretion disc of the BH (Hopkins & Quataert 2011; Anglés-Alcázar et al. 2017a). This accretion is only evaluated for the cold gas ($T < 10^5$ K). For the hot gas ($T > 10^5$ K), the Bondi–Hoyle–Lyttleton model is applied (Davé et al. 2019). AGN feedback in SIMBA is modelled as an injection of kinetic energy following a two-mode approach, with high-mass loading outflows in the radiative 'quasar' mode and lower mass loading but faster outflows at low Eddington ratios in the jet mode (Davé et al. 2019). BHs begin to transition into jet mode for $f_{\mathrm{Edd}} < 0.2$, reaching full speed jets at $f_{\mathrm{Edd}} \lesssim 0.02$. Also, X-ray feedback is included for galaxies with $M_{\mathrm{gas}} < 0.2\,\mathrm{M_\star}$ and full speed jets, following the implementation of Choi et al. (2012). See Thomas et al. (2019, 2021) and Habouzit et al. (2021) for previous studies of BHs in SIMBA.

2.2 Computation of the AGN luminosity and obscuration of the simulated BHs

2.2.1 AGN luminosity

We compute in post-processing the luminosity of the BHs, following the model of Churazov et al. (2005), i.e. explicitly distinguishing radiatively efficient and radiatively inefficient AGNs. The bolometric luminosity of radiatively efficient BHs ($f_{\mathrm{Edd}} > 0.1$) is given by

$$L_{\mathrm{bol}} = \frac{\epsilon_{\mathrm{r}}}{1 - \epsilon_{\mathrm{r}}}\dot{M}_{\mathrm{BH}}c^2.\tag{7}$$

BHs with small Eddington ratios ($f_{\mathrm{Edd}} \leq 0.1$) are considered radiatively inefficient and their bolometric luminosities are given by (Habouzit et al. 2021)

$$L_{\mathrm{bol}} = 0.1 L_{\mathrm{Edd}}(10 f_{\mathrm{Edd}})^2 = (10 f_{\mathrm{Edd}})\epsilon_{\mathrm{r}}\dot{M}_{\mathrm{BH}}c^2.\tag{8}$$

We compute the hard (2–10 keV) X-ray luminosities of the AGN with the bolometric correction of Hopkins, Richards & Hernquist (2007):

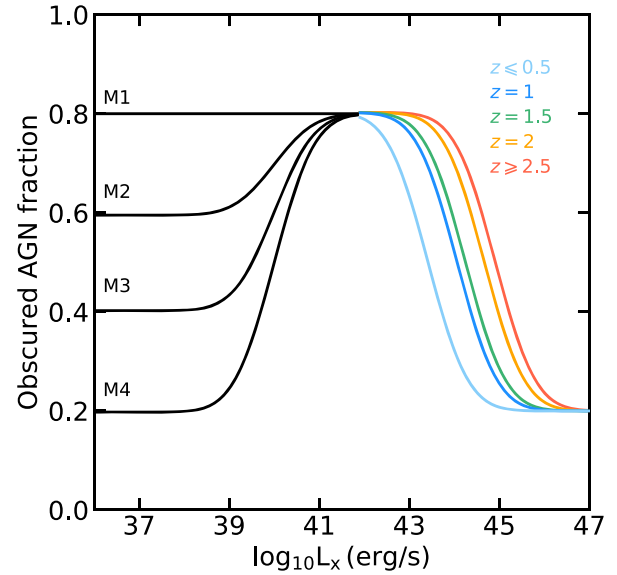$$\log_{10} L_{2-10\,\mathrm{keV},\odot} = \log_{10} L_{\mathrm{bol},\odot} - \log_{10}\mathrm{BC},\tag{9}$$

Figure 1. Our theoretical models for the fraction of obscured AGNs. For relatively bright AGNs with $L_{\mathrm{AGN}} \geqslant 10^{41}\,\mathrm{erg\,s^{-1}}$, the obscured AGN fractions depend on L_{AGN} and redshift (Habouzit et al. 2019); the models are based on observational constraints (Merloni et al. 2014; Ueda et al. 2014). Lacking observational constraints for fainter AGNs ($L_{\mathrm{AGN}} \leqslant 10^{41}\,\mathrm{erg\,s^{-1}}$), we build four different redshift-indepedent models to cover a broad range of possibilities (M1, M2, M3, M4).

with

$$\mathrm{BC} = 10.83\left(\frac{L_{\mathrm{bol},\odot}}{10^{10}\,\mathrm{L_\odot}}\right)^{0.28} + 6.08\left(\frac{L_{\mathrm{bol},\odot}}{10^{10}\,\mathrm{L_\odot}}\right)^{-0.020}.\tag{10}$$

In the following, we present the luminosity of the AGN in $\mathrm{erg\,s^{-1}}$.

2.2.2 AGN obscuration

AGNs can be obscured by gas and/or dust, either physically close to the AGN and at larger scales within the host galaxies. We generally refer to two types of obscured AGNs, either Compton-thin AGN with column densities of $N_{\mathrm{H}} \geqslant 10^{22}\,\mathrm{cm^{-2}}$, or more heavily obscured Compton-thick AGN with $N_{\mathrm{H}} \geqslant 10^{24}\,\mathrm{cm^{-2}}$. While the hard X-ray band suffers less from obscuration than softer bands, the observed luminosity of the AGN could still be attenuated. The obscuration might depend on the intrinsic AGN luminosity and the redshift: there is more gas present at higher redshifts and therefore, the obscuration could be stronger (Gilli et al. 2014; Merloni et al. 2014; Vito et al. 2014, 2016).

We build several models for the fraction of obscured AGNs and apply them to the simulations. Our models depend on redshift and hard X-ray (2–10 keV) AGN luminosity for AGNs with $L_{\mathrm{AGN}} \geqslant 10^{41}\,\mathrm{erg\,s^{-1}}$ (Habouzit et al. 2019, for more details), and only on AGN luminosity for fainter AGNs. Obscuration could in principle also depend on galaxy SFR, a parameter that we do not consider here.

For the AGN with $L_{\mathrm{x}} \geqslant 10^{41}\,\mathrm{erg\,s^{-1}}$, we build the models based on the observational constraints of Merloni et al. (2014) and Ueda et al. (2014): the coloured lines of Fig. 1 are a fit to the observations. These constraints assume that the obscured AGNs mostly include Compton-thin AGNs, but that the presence of Compton-thick AGNs cannot be ruled out. Here, we assume that our models based on these observations include all types of obscured AGNs, and we apply

Table 1. Parametrizations from the literature of the hard X-ray luminosity of the XRB population in galaxies, as a function of their total stellar mass $M_\star$ ($M_\odot$), SFR, and redshift z. In this paper, we mostly use the scaling relation from Lehmer et al. (2019). The $M_\star$, SFR, and z ranges describe the sample of galaxies used to derive the empirical relations.

References	Scaling relations	$M_\star$ ranges	SFR ranges ($M_\odot$ yr^{-1})	Redshift ranges
L10 Lehmer et al. (2010)	$L_{XRB} = 10^{28.96} M_\star + 10^{39.21} \text{SFR}$	$M_\star = 10^{10-11} M_\odot$	$\text{SFR} = 10^{1-2}$	$z \sim 0$
L16 Lehmer et al. (2016)	$L_{XRB} = 10^{29.37}(1+z)^{2.0} M_\star + 10^{39.28}(1+z)^{1.3} \text{SFR}$	$M_\star = 10^{9-12} M_\odot$	$\text{SFR} = 10^{-2-3}$	$z = 0-4$
L19 Lehmer et al. (2019)	$L_{XRB} = 10^{29.15}(1+z)^{2.0} M_\star + 10^{39.73}(1+z)^{1.3} \text{SFR}$	$M_\star = 10^{9-11.5} M_\odot$	$\text{SFR} = 10^{-2-1}$	$z \sim 0$
A17 Aird et al. (2017)	$L_{XRB} = 10^{28.81}(1+z)^{3.9} M_\star + 10^{39.50}(1+z)^{0.7} \text{SFR}^{0.86}$	$M_\star = 10^{8.5-11.5} M_\odot$	$\text{SFR} = 10^{-1-3}$	$z = 0.1-4$
F18 Fornasini et al. (2018)	$L_{XRB} = 10^{29.98}(1+z)^{0.62} M_\star + 10^{39.78}(1+z)^{0.2} \text{SFR}^{0.84}$	$M_\star = 10^{9.5-11.5} M_\odot$	$\text{SFR} = 10^{-1-3}$	$z = 0.1-5$

them on the full AGN simulated samples. Since we are lacking observational constraints on the population of faint AGNs, and even more so on the fraction of those that are obscured, we build four models that cover a broad range of possible fractions of obscured AGNs (Fig. 1). In our model M1, a large fraction (80 per cent) of the faint AGN with $L_{AGN} \leqslant 10^{41}$ erg s^{-1} are obscured, and this fraction does not depend on AGN luminosity. For M2, M3, and M4, the fraction does depend on L_{AGN} and can progressively go down to 60 per cent (M2), 40 per cent (M3), or 20 per cent (M4) of AGNs being obscured. Our choice of low fractions of obscured AGNs among the faint AGN with $L_{AGN} \leqslant 10^{40}$ erg s^{-1} is motivated by the fact that in simulations if these BHs do not have high accretion rates, their close surrounding is depleted of gas. Consequently, if the gas reservoir is small there may be little room for the gas/dust to obscure the emission from the AGN. Since this is highly uncertain, we also have our model M1 with a high fraction of obscured AGNs to bracket the range of possibilities. In the following, we apply our obscuration models only when specified in the text and figures.

2.3 Computation of the XRB population luminosity

We parametrize the X-ray emission ($2-10$ keV band) from the XRB population of the simulations (which is not modelled in the simulations), including both LMXBs and HMXBs. To do so, we employ XRB emission scaling relations that have been derived from observational samples. These relations provide the luminosity of the XRB population for a given galaxy as a function of its SFR, its stellar mass, and for some of them its redshift:

$$L_{XRB} = \alpha_{LMXB} (1+z)^\gamma M_\star + \beta_{HMXB} (1+z)^\delta \text{SFR}. \quad (11)$$

For example, the parameters $\log_{10} \alpha_{LMXB} = 29.37 \pm 0.15$ erg/s/$M_\odot$, $\log_{10} \beta_{HMXB} = 39.28 \pm 0.05$ erg/s/($M_\odot$/yr), $\gamma = 2.03 \pm 0.60$, $\delta = 1.31 \pm 0.13$ have been described, e.g. in Lehmer et al. (2016, 2019). The relations described in Aird et al. (2017) and Fornasini et al. (2018) use a power law of SFR$^{0.86}$ and SFR$^{0.84}$, respectively. While it has been shown that the metallicity of the galaxies can play a role in the amplitude of the XRB emission, especially in some specific regimes of stellar mass and redshift, we do not include any metallicity dependence in our analysis below.

To study the impact of the modelling of the XRB luminosity, we use five different empirical relations (Lehmer et al. 2010, 2016, 2019; Aird et al. 2017; Fornasini et al. 2018). We report these relations in Table 1. We compare the scaling relations in Fig. 2, as a function of galaxy stellar mass (top panels), SFR (middle panels), and sSFR (bottom panels). The left-hand panels show redshift $z = 0$, and the right-hand panels $z = 2$. The difference between the different models can be up to one order of magnitude in luminosity at $z = 0$, and more than an order of magnitude at higher redshift. At fixed SFR, the luminosity of the XRB population increases with the stellar mass of galaxies. We note that for galaxies with high SFR of $\log_{10} \text{SFR}/(M_\odot/\text{yr}) \sim 2$ the luminosity is almost constant
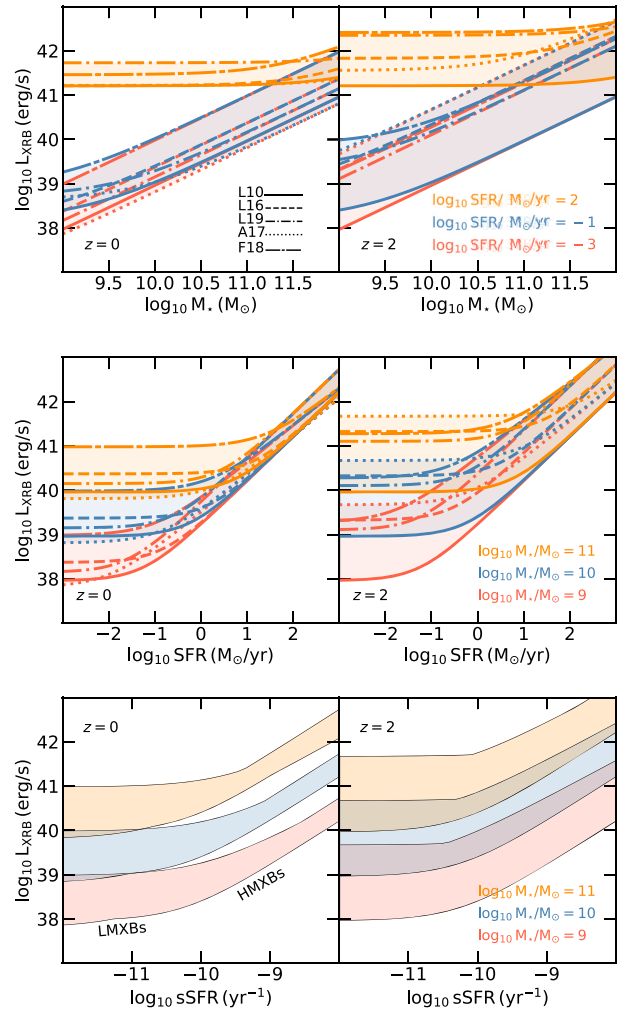

Figure 2. Comparison of the XRB scaling relations from Table 1 (Lehmer et al. 2010, 2016, 2019; Aird et al. 2017; Fornasini et al. 2018), at $z = 0$ (left-hand panels) and $z = 2$ (right-hand panels). The difference between the models can be up to one order of magnitude in hard X-ray (2–10 keV) luminosity at $z = 0$, and more at higher redshift. The XRB luminosity increases with $M_\star$ (at fixed SFR), and with SFR (at fixed $M_\star$), and with redshift for most of the models. LMXBs dominate the hard X-ray (2–10 keV) luminosity of the galaxies with $\log_{10} \text{sSFR}/\text{yr} \leqslant -10.5$. For higher sSFR, the X-ray emission is dominated by HMXBs.

with stellar mass, for all the scaling relations studied here. The XRB population luminosity also increases with SFR, at fixed galaxy stellar mass. The model of Lehmer et al. (2010) does not evolve with redshift. For the other models, the normalization of the scaling relations increases with increasing redshift. In general, the models of
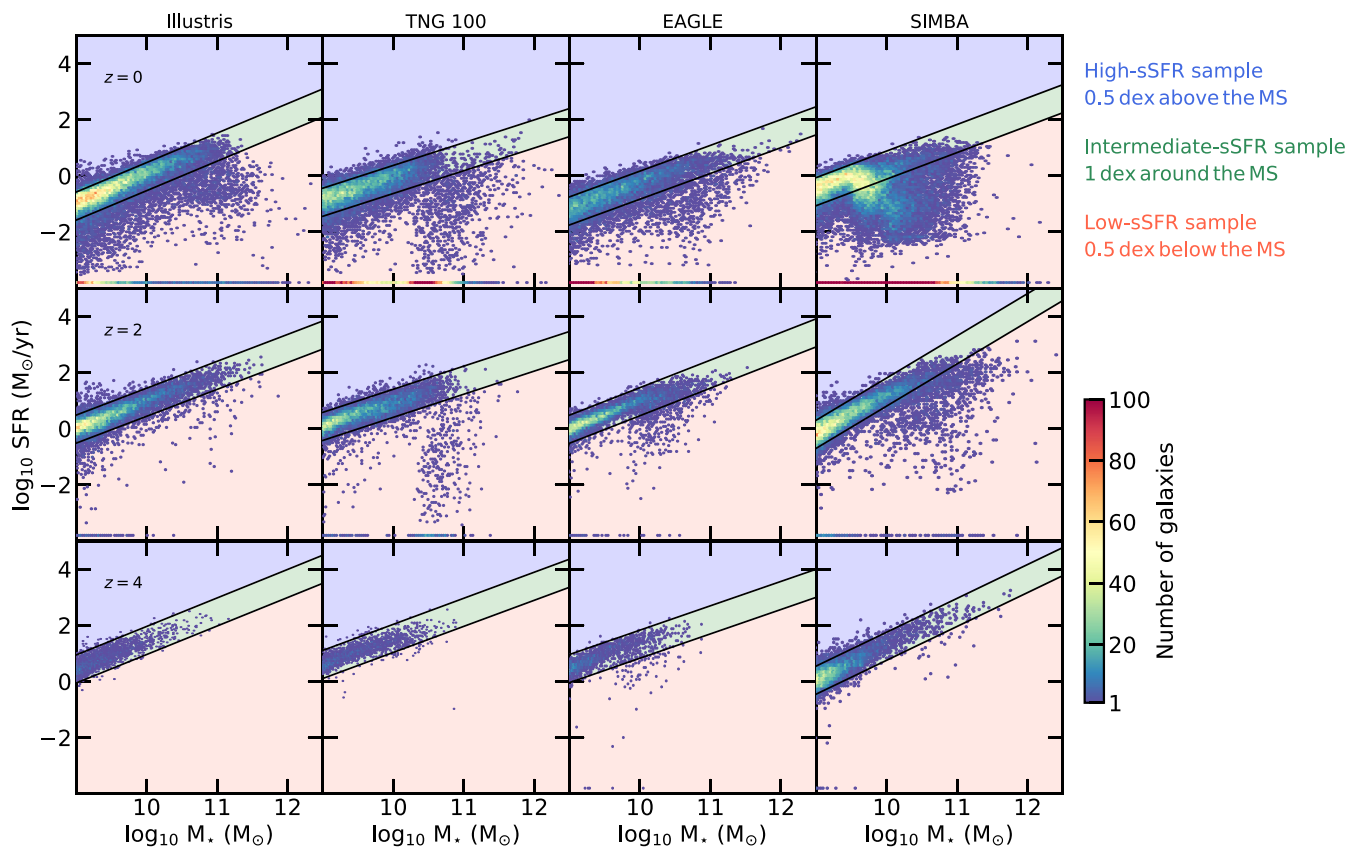
Figure 3. SFR − M$_\star$ plane with hexabins colour coded by the number of galaxies in the bins. Galaxies with SFR $\leqslant 10^{-4}$ M$_\odot$ yr^{-1} are shown with this value. We define three samples of galaxies: the *high-sSFR* sample with galaxies 0.5 dex above the star-forming main sequence, the *intermediate-sSFR* sample with galaxies on the star-forming sequence, and the *low-sSFR* sample with galaxies below 0.5 dex of the main sequence. The *high-sSFR*, *intermediate-sSFR*, and *low-sSFR* samples are, respectively, composed of starburst, main-sequence, and quiescent galaxies.

Lehmer et al. (2019) and Fornasini et al. (2018) provide the highest normalizations of the $L_{\rm XRB}$ relation at $z = 0$. At higher redshift, the model of Aird et al. (2017) also provides high luminosities. For low sSFR galaxies with $\log_{10}$ sSFR/yr < -10.5, the XRB population luminosity is dominated by LMXBs whose luminosity depends on the stellar mass. But for galaxies with higher sSFR of $\log_{10}$ sSFR/yr > -10.5, the XRB luminosity is dominated by HMXBs whose luminosity scales with the SFR (Lehmer et al. 2019).

2.4 Galaxy star-forming main sequence and sSFR galaxy samples

In this paper, we will present our results as a function of galaxy properties, i.e. their stellar mass and SFR. In the following, we define the simulation star-forming main sequence and build three different galaxy samples from the simulations.

2.4.1 Galaxy star-forming main sequence

We show the SFR − M$_\star$ relation for galaxies from the Illustris, TNG100, EAGLE, and SIMBA simulations in Fig. 3. We define the star-forming main sequence of the simulations as the mean SFR in fixed width bins (same width for all the simulations) of galaxy

stellar mass for $10^9 \leqslant M_\star \leqslant 10^{10}$ M$_\odot$[1] (see also Genel et al. 2014; Vogelsberger et al. 2014b; Furlong et al. 2015; Sparre et al. 2015; Bluck et al. 2016; Terrazas et al. 2017; Davé et al. 2019; Donnari et al. 2019; Matthee & Schaye 2019). We purposely exclude more massive galaxies to compute the main sequence as many of these galaxies are quenched. The main sequence can be defined as a power law:

$$\log_{10}[\text{SFR}_{\rm MS}/(\text{M}_\odot\ \text{yr}^{-1})] = \alpha + \beta \log_{10}(\text{M}_\star/\text{M}_\odot), \quad (12)$$

with SFR$_{\rm MS}$ the SFR of the galaxies on the main sequence, in M$_\odot$ yr^{-1}. The parameters α and β are computed for each simulation and redshift and are given in Table A1. The simulations present some differences in the galaxy population in Fig. 3, such as the exact normalization and slope of their star-forming main sequence, or the formation of the population of quenched galaxies, i.e. galaxies with low or null SFR. We note that the SIMBA simulation has the steepest main sequence, but overall we find a good agreement between the simulations for the normalization of the star-forming main sequence. In observations there is, at the moment, no real consensus on the exact slope and normalization of the star-forming main sequence, and the parameters depend on the observed samples. The slope of

---

[1]We use $M_\star \leqslant 10^{9.5}$ M$_\odot$ for SIMBA at $z = 0$ because of the presence of galaxies with reduced SFR at lower stellar mass than in the other simulations.

the star-forming main sequence of observed galaxies can be slightly shallower (Habouzit et al. 2019; Hahn et al. 2019).

When looking at the population of quenched galaxies, we can identify some features of the AGN feedback modelling such as the sharp decrease of the SFR in TNG100 galaxies of $M_\star \geqslant$ a few 10^{10} $M_\odot$ (but still in broad agreement with observational data, see Donnari et al. 2021b). This shows the transition at $M_{BH} \sim$ a few 10^8 $M_\odot$ between the high accretion state and the low accretion state of the AGN feedback modelling (see equation 4). The low accretion mode feedback is efficient, and both regulates the BHs and quenches their host galaxies (Weinberger et al. 2018; Habouzit et al. 2019). Towards $z = 0$, we also note a strong decrease of the SFR of a large fraction of the galaxies with $M_\star \leqslant 10^{9.5}$ $M_\odot$ in all the simulations, which reflects the quenching of satellite galaxies (Donnari et al. 2021a).

2.4.2 Defining three galaxy samples

For the purpose of our analysis, we divide the galaxies into three subsets, which are simulation- and redshift-dependent. The *high-sSFR* subset (blue background in Fig. 3) consists of starburst galaxies with SFR higher than half a dex above the main sequence. The *intermediate-sSFR* subset (green background in Fig. 3) includes star-forming galaxies on the main sequence, i.e. within 1 dex. Finally, the *low-sSFR* subset (red background in Fig. 3) includes all galaxies with SFR below half a dex of the main sequence, i.e. quiescent galaxies or galaxies on their way to quiescence.

We find that a low fraction of the galaxies of $M_\star \geqslant 10^9$ $M_\odot$ are in the *high-sSFR* sample, typically less than 5 per cent in all the simulations and at all redshifts. The *intermediate-sSFR* represents the largest fraction of galaxies ($\leqslant 90$ per cent, $z = 4$) at high redshift in all the simulations. With time, the number of galaxies in the *intermediate-sSFR* samples decreases, for all the simulations. The percentage of galaxies in the *low-sSFR* samples increases with time for all the simulations, as more and more galaxies quench. Most of the galaxies in the *low-sSFR* samples are massive with $M_\star \gtrsim 10^{10}$ $M_\odot$ and have reduced SFR because of AGN feedback. For example, in TNG there is a sharp decrease of the SFR in galaxies of $M_\star \gtrsim 10^{10.5}$ $M_\odot$ due to the kinetic low-accretion mode of the AGN feedback model. Still, some of the galaxies present in the *low-sSFR* samples have lower masses ($M_\star < 10^{10}$ $M_\odot$), and can have lower SFR due to gas starvation, SN feedback, and environmental quenching. The final percentage of galaxies in each sample at $z = 0$ varies from simulation to simulation. We present these numbers in Table A2. We use these three samples in Section 5.

From now on, we only consider galaxies of total stellar mass $M_\star \geqslant 10^9$ $M_\odot$ which host a BH. The BH can be an AGN, i.e. accreting mass, or a non-accreting BH, and in that case $L_{AGN} = 0$. The mass of the BHs is the mass of individual BHs in all the simulations, except in TNG and Illustris for which the BH mass is the sum of the mass of all the BHs within a galaxy. In practice, only a couple of galaxies per output host several BHs at the same time. The total stellar mass that we use here for all the simulations is not exactly computed in the same way in all the simulations, but we prefer to use each simulation definition. The total stellar mass of the galaxies is computed, e.g. as twice the stellar mass in the half mass radius for the Illustris and TNG100 simulation, within an aperture of 30 kpc in EAGLE. In SIMBA, galaxies are identified using a friends-of-friends galaxy finder, assuming a spatial linking length of 0.0056 times the mean interparticle spacing (equivalent to twice the minimum softening length). Discussion on the impact of these different definitions can

be found in Pillepich et al. (2018a) for TNG. Finally, we do not distinguish between central and satellite galaxies, and all our results group these two types of galaxies.

3 IN WHICH GALAXIES LIVE THE FAINT AGN?

In this section, we first investigate the properties of galaxies hosting faint AGNs. We divide the galaxies in subsamples depending on their AGN luminosity: $L_{AGN} = 10^{37.5-38.5}$ erg s^{-1}, $L_{AGN} = 10^{39.5-40.5}$ erg s^{-1}, and $L_{AGN} = 10^{41.5-42.5}$ erg s^{-1}. For each subsample as well as for the whole galaxy sample, we show the distributions of stellar mass, BH mass, and sSFR in Fig. 4. Galaxies with $\log_{10}$ sSFR/yr$^{-1} < -13$ are shown with this value.

To quantify the findings described in the following, we perform several Kolmogorov–Smirnov (KS) tests for the distributions of the faint AGN with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ and $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 42$, for all the simulations and for $M_\star$, M_{BH}, and SFR (Table B2). We cannot reject the null hypothesis that two distributions were drawn from the same distribution for high values of p, i.e. $p \geq 0.01$. However, if $p < 0.01$ the two distributions can be considered as statistically different.

3.1 Illustris

In the Illustris simulation, faint AGNs with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ are mostly located in low-mass galaxies of $M_\star \leqslant 10^{10}$ $M_\odot$ at any redshift. These AGNs also tend to be low-mass BHs of $M_{BH} \leqslant 10^8$ $M_\odot$. More luminous AGNs with $L_{AGN} \sim 10^{42}$ erg s^{-1} are also located in low-mass galaxies at high redshift, but with time they start deviating from the $L_{AGN} \sim 10^{38}$ erg s^{-1} AGNs and are largely located in massive galaxies of $M_\star \sim 10^{11}$ $M_\odot$ at $z = 0$. These AGNs are also mostly massive BHs of $M_{BH} \sim 10^9$ $M_\odot$ at $z = 0$.

For Illustris, we find with our KS test that the distributions of $M_\star$, M_{BH}, and sSFR for $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ and $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 42$ are always statistically different, except the sSFR distribution at $z = 3$.

3.2 TNG100

In the TNG100 simulation, the faint AGNs are located in galaxies with generally different properties. From high redshift to low redshift, the AGNs with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ are mostly located in massive galaxies of $M_\star \geqslant 10^{10.5}$ $M_\odot$. There is also a population of AGN with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ in lower mass galaxies of $M_\star \sim 10^9$ $M_\odot$, similar to the population found in Illustris. Brighter AGNs with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 42$ are mostly hosted in lower mass galaxies with $M_\star \leqslant 10^{10}$ $M_\odot$. Since the $M_{BH}-M_\star$ relationship is tight in TNG100 (Habouzit et al. 2021, and Fig. 5 below), the $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ AGNs are also mostly massive BHs of $M_{BH} \sim 10^9$ $M_\odot$. More luminous AGNs are BHs of $M_{BH} \leqslant 10^7$ $M_\odot$ at high redshift ($z \geq 2$) and BHs of $M_{BH} \leqslant 10^8$ $M_\odot$ at $z \leq 2$. The distributions of the $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 38$ and the $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 42$ AGN are statistically different (KS test). At $z = 3$, there are only few AGNs with $\log_{10} L_{AGN}/(\text{erg s}^{-1}) \sim 42$ and the KS test cannot be performed.

The fact that faint AGNs ($\log_{10} L_{AGN}/(\text{erg s}^{-1}) < 40$) are found in massive galaxies of a few 10^{10} $M_\odot$ and that they are powered by BHs of a few 10^8 $M_\odot$ is due to the efficient low-accretion mode of the TNG feedback model. These AGNs are found in quenched galaxies with reduced SFR. The population of quenched galaxies can also be observed in the sSFR distribution.
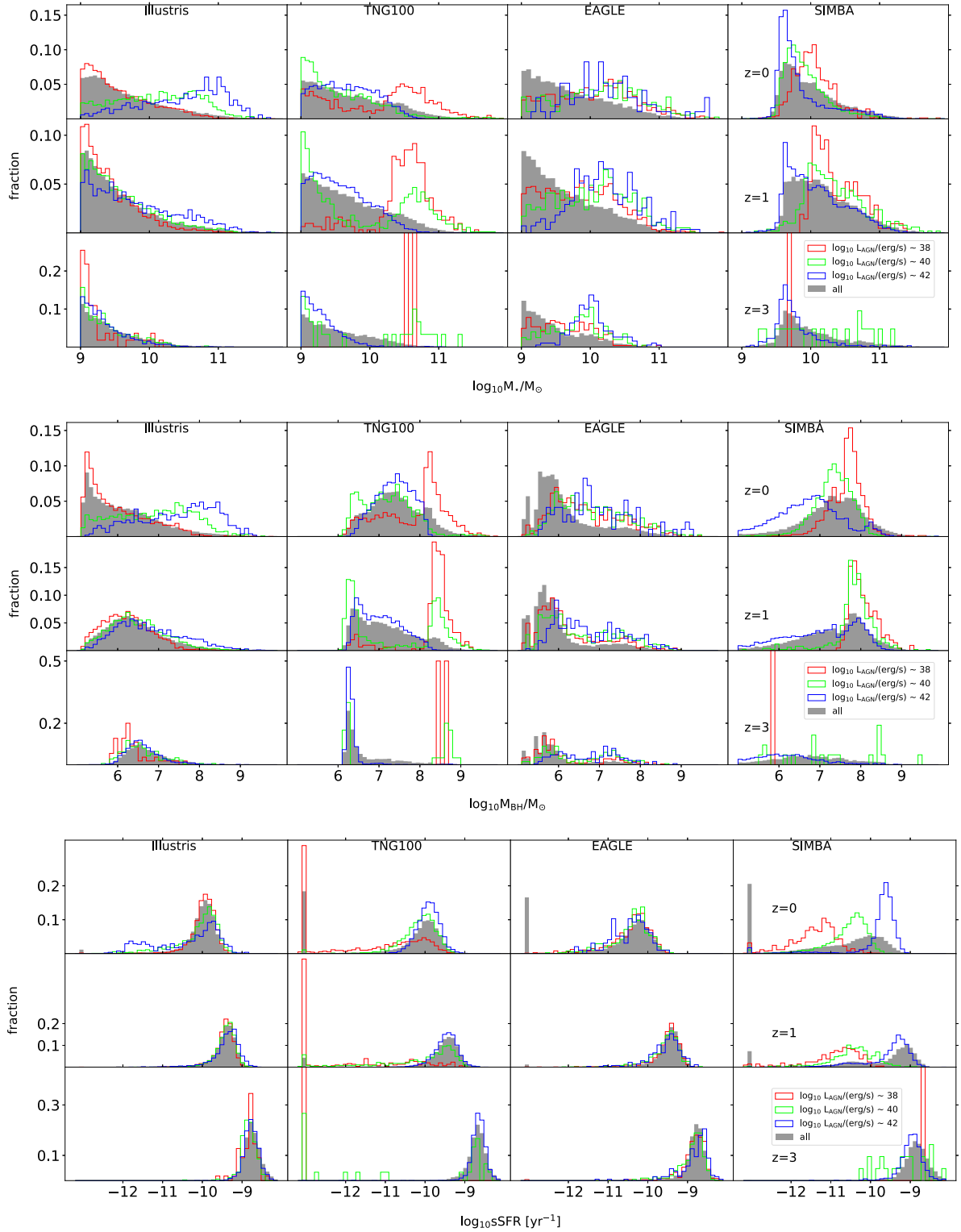
Figure 4. Normalized distributions of galaxy stellar mass, BH mass, and sSFR, for $z = 0$, 1, and 3 (top, middle, and bottom panels). Red distributions only include galaxies with $L_{AGN} \sim 10^{38}$ erg s^{-1} (bin $10^{37.5}$–$10^{38.5}$ erg s^{-1}), $L_{AGN} \sim 10^{40}$ erg s^{-1} (bin $10^{39.5}$–$10^{40.5}$ erg s^{-1}) for the green distributions, and $L_{AGN} = \sim 10^{42}$ erg s^{-1} (bin $10^{41.5}$–$10^{42.5}$ erg s^{-1}) for the blue ones. Distributions are normalized by the total number of galaxies in the specific luminosity bin. The grey shaded histograms show the distributions for the full galaxy population. Faint AGNs of different luminosities are powered by different BHs and reside in different stellar mass galaxies in the different simulations. For example, the AGNs with $L_{AGN} \sim 10^{38}$ erg s^{-1} reside in low-mass galaxies in Illustris and in galaxies with much higher masses in TNG100 and SIMBA.
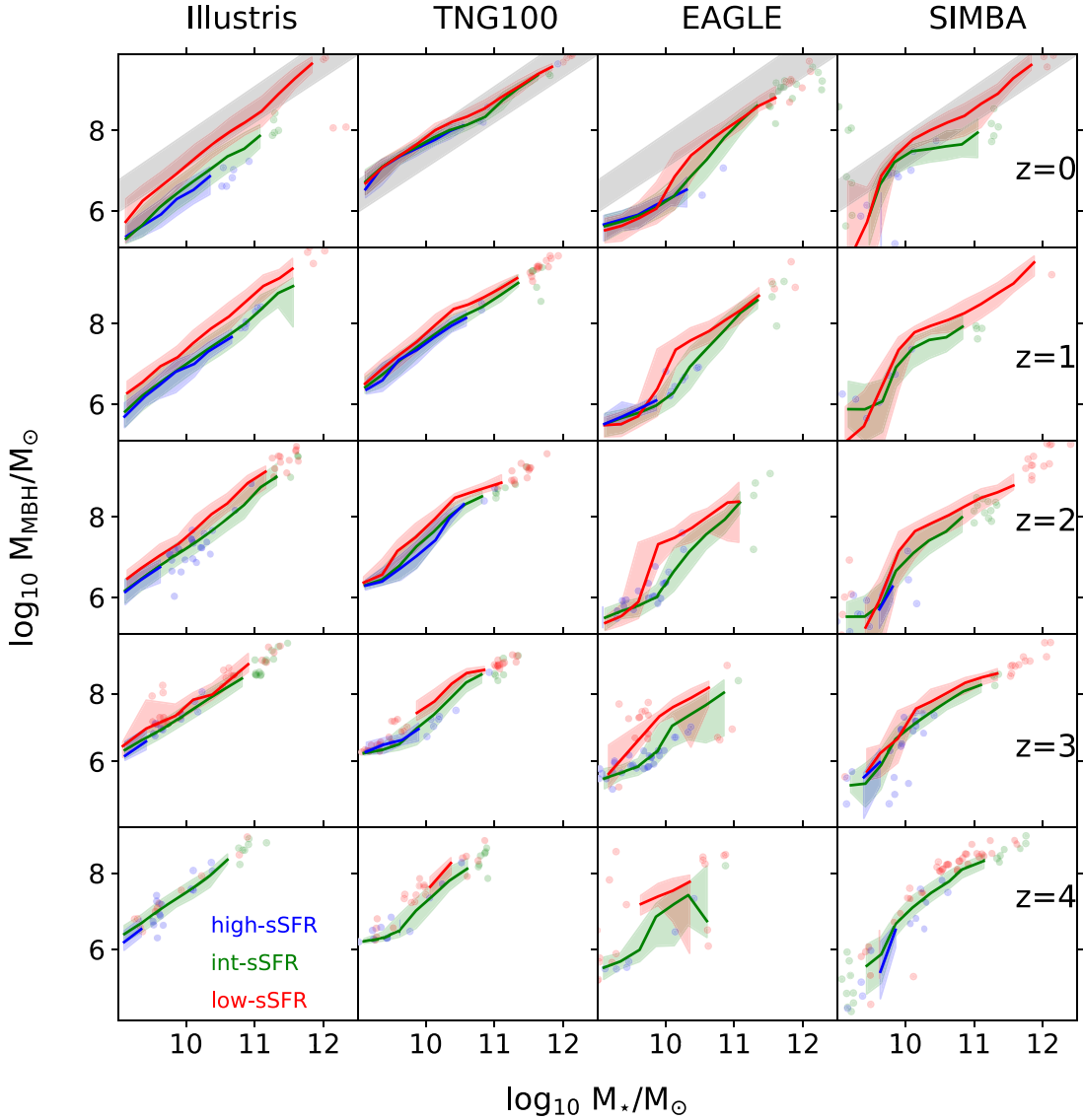
Figure 5. Median of the M_{BH}–$M_{\star}$ relation of the simulations for $z = 0$, 1, 2, 3, and 4. We divide the galaxy population into the three galaxy samples (*high-sSFR* in blue, *intermediate-sSFR* in green, *low-sSFR* in red). Shaded regions show the 15th–85th percentiles of the distribution at fixed total stellar mass. The width of the stellar mass bins is 0.25 dex. Individual galaxies are shown with dots in bins with less than 10 galaxies. The grey region at $z = 0$ encloses the empirical scaling relations of Kormendy & Ho (2013) (higher normalization of the scaling relation), McConnell & Ma (2013), and Häring & Rix (2004) (lower normalization of the relation). In these simulations, the median BH mass of low-sSFR galaxies is higher than for higher sSFR galaxies, at fixed stellar mass.

3.3 EAGLE

In the EAGLE simulation, the AGNs with $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 42$ are mostly located in galaxies of $M_{\star} \sim 10^{10}\,\mathrm{M_{\odot}}$ at $z = 3$. The distribution of the stellar mass of their host galaxies becomes broader with time, with AGNs located in the range $M_{\star} = 10^{9.5} - 10^{11}\,\mathrm{M_{\odot}}$. At high redshift $z \geq 2$, the distribution of the stellar mass of the galaxies hosting AGNs with $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 38$ is different from the $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 42$ AGNs. The $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 38$ AGNs are located in less massive galaxies. However, at lower redshift ($z < 1$) the $M_{\star}$ distribution of $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 38$ extends to more massive galaxies and become very similar to the distribution of more luminous AGNs with $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 42$. For the mass of the BHs powering these AGNs, we note some differences at high redshift ($z \sim 3$) between AGNs of $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 38$ and $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 42$, with fainter AGNs being lower mass

BHs of $M_{BH} \sim 10^6\,\mathrm{M_{\odot}}$. The M_{BH} distributions have broad humps for the two types of AGNs at high redshift. At $z = 0$ when the distributions are broader, the peak of the distributions is still different but the distributions are less distinguishable.

In EAGLE, the faintest AGNs are found in two BH populations: predominantly in low-mass BHs of $M_{BH} \leqslant 10^6\,\mathrm{M_{\odot}}$, but also in more massive BHs of $M_{BH} \sim 10^7\,\mathrm{M_{\odot}}$. The first BH population does not power luminous AGNs because they reside in low-mass galaxies that are regulated by SN feedback. The second BH population consists of more massive BHs located in more massive galaxies, and are regulated by AGN feedback.

For the EAGLE simulation, by applying a KS test we find that all the $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 38$ and $\log_{10} L_{AGN}/(\mathrm{erg\,s^{-1}}) \sim 42$ distributions are significantly different, except the sSFR distributions at $z = 0$ and $z = 3$ and the $M_{\star}$ distribution at $z = 0$.

3.4 SIMBA

The distributions of the faint AGN with $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 38$ and $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 42$ of the SIMBA simulation are qualitatively similar to the distribution of TNG100. However, the $M_\star$ distributions of TNG100 extend to lower values than the $M_\star$ distributions of SIMBA. Our KS test indicates that the distributions of $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 38$ and $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 42$ in SIMBA are statistically different. At $z = 3$, there are only few AGNs with $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 42$ and we cannot perform the KS test.

AGNs with $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 38$ are located in more massive galaxies than the AGNs with $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 42$. The peak of the two $M_\star$ corresponding distributions are separated by less than an order of magnitude. The $M_\star$ distributions are less separated at higher redshift. For the BHs, we note that at $z = 1$, the peak of the BH mass is very similar, but separate at later time. At $z = 0$, simulated AGNs with $\log_{10} L_{\rm AGN}/({\rm erg/s}) \sim 42$ are found mostly as BHs of $M_{\rm BH} \sim 10^7 \, {\rm M_\odot}$. In contrast, the distribution of $\log_{10} L_{\rm AGN}/({\rm erg \, s^{-1}}) \sim 38$ AGN peaks at $M_{\rm BH} \sim 10^8 \, {\rm M_\odot}$.

The sSFR distributions are very different for the different AGN luminosities in SIMBA. The AGNs with $\log_{10} L_{\rm AGN}/({\rm erg \, s^{-1}}) \sim 38$ have a peak in their sSFR distribution which is nearly two orders of magnitude smaller than the AGNs with $\log_{10} L_{\rm AGN}/({\rm erg \, s^{-1}}) \sim 42$. In SIMBA, the faintest AGNs are in galaxies with a low sSFR, they are quenched, and the AGNs with $\log_{10} L_{\rm AGN}/({\rm erg \, s^{-1}}) \sim 42$ are in star-forming galaxies. There is a large population of galaxies that quench with time, and with them a large population of fainter AGNs at $z = 0$.

In this first section, we demonstrated that the BHs powering the faint AGN in the Illustris, TNG100, EAGLE, and SIMBA simulations, as well as the properties of their host galaxies, are different among the different simulations. Indeed, faint AGNs of $L_{\rm AGN} \sim 10^{38} \, {\rm erg \, s^{-1}}$ can be powered by relatively massive BHs and be located in massive galaxies ($M_\star \gtrsim 10^{10} \, {\rm M_\odot}$) with reduced SFR (TNG100, SIMBA), or be powered by lower mass BHs in less massive galaxies ($M_\star \lesssim 10^{10} \, {\rm M_\odot}$) still forming stars (Illustris, EAGLE). Both galaxy and BH mass appear to be fundamental quantities to understand the faint AGN populations in simulations. Moreover, in some simulations we already see that the sSFR of the host galaxies can play an important role. In the next section, we go further and divide the full simulated galaxy samples into subsamples as a function of the galaxies distance from the star-forming main sequence.

4 AGN HARD X-RAY LUMINOSITY

4.1 $M_{\rm BH}$–$M_\star$ properties in the three galaxy samples

We show the median of the mass of the BHs located in the three galaxy samples in Fig. 5, as a function of the total stellar mass of the galaxies and for several redshifts from $z = 0$ (top row) to $z = 4$ (bottom row).

On average we find that galaxies with lower SFR (*low-sSFR* samples, red lines) host more massive BHs at fixed stellar mass; the effect is mild and in most cases within the 15th–85th percentiles of the sample distributions. We find this for the Illustris and TNG100 simulations at all stellar masses, and for galaxies of $M_\star \geqslant 10^{10} \, {\rm M_\odot}$ in the EAGLE and SIMBA simulations. At low redshift, the EAGLE *high-sSFR* galaxies with $M_\star < 10^{10} \, {\rm M_\odot}$ (blue lines) host BHs slightly more massive than galaxies with lower SFR. In SIMBA, we cannot really assess the behaviour of BH mass of galaxies with $M_\star \leqslant 10^{10} \, {\rm M_\odot}$ as the seeding takes place in galaxies

of $M_\star \sim 10^{9.5} \, {\rm M_\odot}$. Most simulations have a tight $M_{\rm BH}$–$M_\star$ relation (not shown here, but see Habouzit et al. 2021), TNG100 having the tightest relation, which explains the very mild difference in the median of BH mass for the three samples at all redshifts. We do not show it here, but the $M_{\rm BH}$–$M_\star$ diagram of the simulations at $z = 0$ is in broad agreement with the observational samples of e.g. Reines & Volonteri (2015). To guide the eye, we show in grey shaded area in the $z = 0$ panels a region enclosing several $M_{\rm BH}$–$M_{\rm bulge}$ empirical scaling relations (e.g. Häring & Rix 2004; Kormendy & Ho 2013; McConnell & Ma 2013) that have been used to calibrate the simulations. It should be noted that the calibrations are done on the entire galaxy population of the simulations and not on one given sSFR subsample, whereas in this paper we are testing the simulation outcomes to deeper and more constraining details. It may hence be not surprising that some discrepancies as the ones revealed above are in place between e.g. the high-sSFR $M_{\rm BH}$–$M_\star$ relation and the empirical scaling relations.

In this section, we showed that, on average, lower sSFR galaxies (quenched or on their way to quiescence) tend to host more massive BHs in cosmological simulations. Since the $M_{\rm BH}$–$M_\star$ relation is tight in simulations, the effect is small. This is in agreement with the results of Thomas et al. (2019) for SIMBA. Moreover, this is consistent with what is found in observations (Terrazas et al. 2017), but such observational samples with estimates of both dynamical BH mass and SFR are still restrictively small. Our previous paper (Habouzit et al. 2021) provides more information and comparisons between the simulated population of BHs in these simulations and observations.

4.2 AGN X-ray (2–10 keV) luminosity in the three galaxy samples

In this paper, we aim at deriving the galaxy total X-ray luminosity of galaxies hosting faint AGNs and comparing it to observational constraints, for different types of galaxies. As a first step we derive the X-ray emission of the AGN, a crucial component of the total luminosity of galaxies.

We present in Fig. 6 the median of the AGN luminosity as a function of the stellar mass of galaxies. In order to focus on analysis on the faint AGN, we only consider AGNs with hard X-ray luminosity below $L_{\rm AGN} \sim 10^{42} \, {\rm erg \, s^{-1}}$. In practise, we use the luminosity limit of the COSMOS survey to detect AGNs as individual sources. This allows us to evolve the luminosity limit with redshift in a way that we will be able to compare our results with observations in Section 7. In Section 7, we will also apply this limit and describe in detail how we compute it. For our analysis we remove all AGNs that are brighter than this limit. In this section, we also do not consider AGN obscuration. The shaded regions indicate the 15th–85th percentiles of the distribution. A median value is only calculated for stellar mass bins with more than 10 galaxies, otherwise individual galaxies are shown as points.

In general, the median AGN luminosity increases with redshift. Furthermore, the galaxies with a higher sSFR (e.g. the blue lines in Fig. 6) have higher luminosities in the simulations than lower sSFR galaxies (red lines), at fixed stellar masses. There is an exception in Illustris at redshifts $z = 0, 1$. Here, the AGN luminosity of the *low-sSFR* sample (in red) is higher for stellar masses between $M_\star = 10^{10} \, {\rm M_\odot}$ and a few $10^{10} \, {\rm M_\odot}$. The median AGN luminosity drops for more massive galaxies in Illustris.

In the panels of the TNG100 simulation, the low-accretion mode of the AGN feedback (kinetic mode) leads to a sharp decrease of the AGN luminosity (Weinberger et al. 2018) for galaxies with stellar mass $\log_{10} M_\star/{\rm M_\odot} = 10.25$ at redshifts $z = 0, 1, 2, 3$ in the *low-*
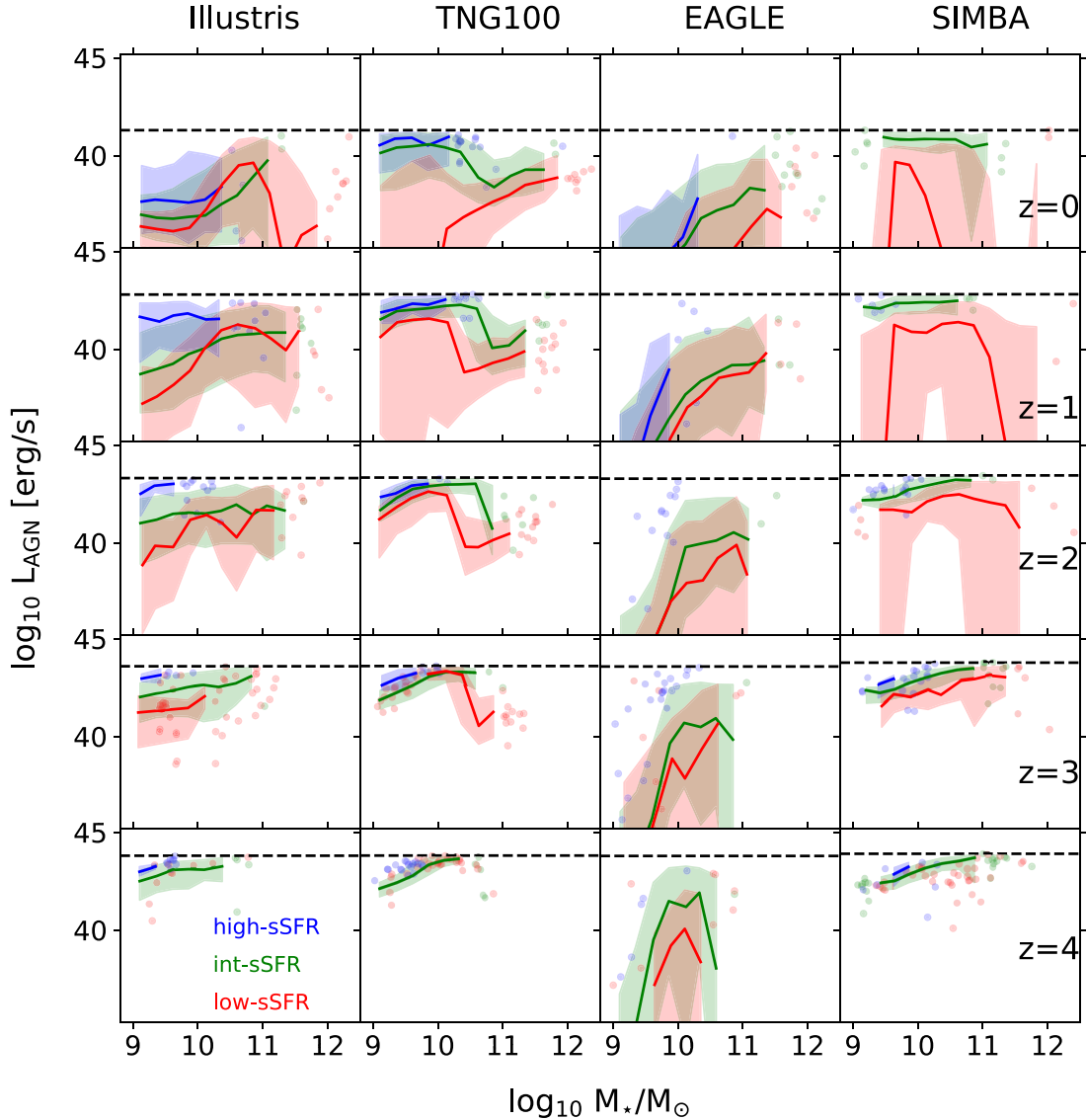
Figure 6. Median AGN hard X-ray (2–10 keV) luminosity for different galaxy stellar mass bins and redshifts from $z = 0$ (top row) to $z = 4$ (bottom row). Individual galaxies are shown as points in bins with less than 10 galaxies. The shaded regions indicate the 15–85th percentiles. We do not consider AGN obscuration here. All AGNs that could be detected individually by the COSMOS survey are excluded (black dashed line). AGNs are on average brighter in galaxies with high sSFR. AGN feedback is responsible for the decrease of luminosity in massive galaxies, in some simulations.

sSFR sample. We note here that the sharp decrease is present at $z = 0$, but masked by the low L_{AGN} in lower mass galaxies due to the rarefaction of cold gas (lower accretion rates on to the BHs) at low redshift.

A large shaded region indicates a broad distribution of luminosities in a given stellar mass bin. The size of the shaded regions decreases with increasing redshift. Moreover, the shaded regions are large for *low-sSFR* galaxies, smaller for *intermediate-sSFR* galaxies, and even smaller for *high-sSFR* galaxies. Many different luminosities are possible for *low-sSFR* galaxies depending on how much gas there is still available for the central BH. The amount of gas and its properties (e.g. temperature) can be affected by both the SN feedback in low-mass galaxies and by AGN feedback primarily in more massive galaxies (e.g. Habouzit et al. 2017; Nelson et al. 2019a; Zinger et al. 2020). In general, this leads to broad distributions.

5 HARD X-RAY LUMINOSITY FROM X-RAY BINARIES: CAN THEY OUTSHINE THE AGN?

In the previous section, we derived the AGN luminosity of the host galaxies of faint AGNs. In this section, we derive the luminosity of the XRB population of the simulated galaxies, and evaluate whether the XRB luminosity can be higher than the AGN luminosity, and for which galaxies.

5.1 XRB hard X-ray (2–10 keV) luminosity as a function of galaxy properties and redshift

In Fig. 7, we show as an example the X-ray luminosity of the XRB population in the simulated galaxies of Illustris, for different redshifts. We use the relation of Lehmer et al. (2019), which is the most recent among the models discussed in this paper. The error bars
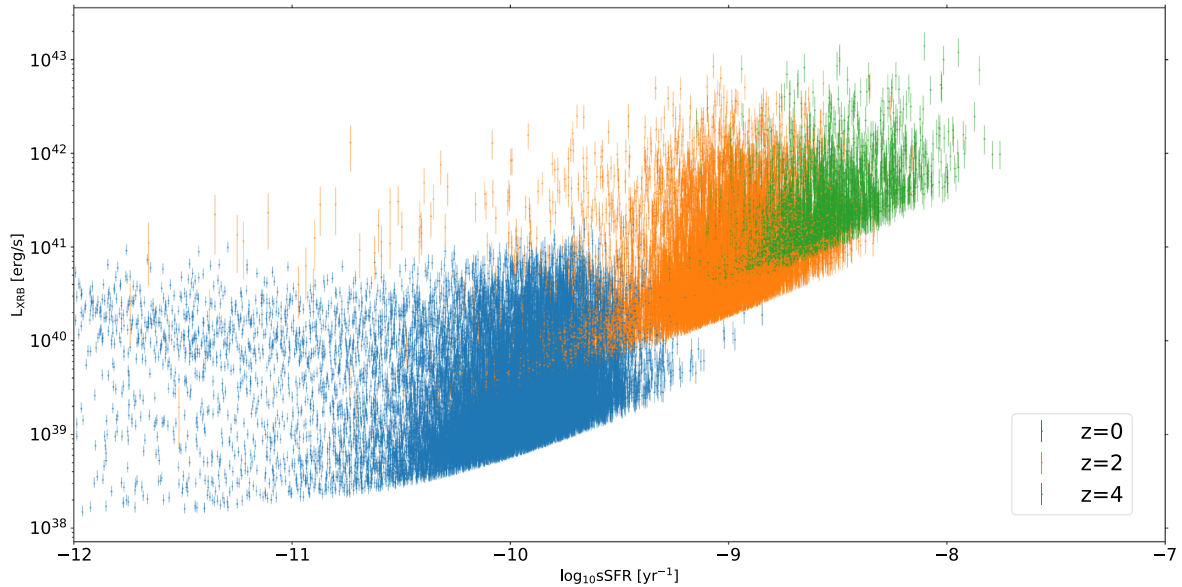
Figure 7. Hard X-ray (2–10 keV) luminosity of the XRB population for the simulated galaxies of Illustris. The XRB population luminosity is estimated from the galaxy SFR, stellar mass, and redshift. Here, we use the model from Lehmer et al. (2019). The error bars are calculated from the uncertainties of the parameters in the formula for the XRB luminosity. In the modelling, massive XRBs dominate the XRB population luminosity in galaxies with high sSFR ($\log_{10}$ sSFR/yr $\geqslant -10.5$). Low-mass binaries dominate at lower sSFR ($\log_{10}$ sSFR/yr < -10.5).

indicate the uncertainties on the parameters in equation (11). At high redshift ($z \sim 4$), most of the simulated galaxies form stars efficiently and have high sSFR. For the Illustris simulation and this given empirical XRB relation, this corresponds to hard X-ray luminosity of the XRB populations in the range $L_{\mathrm{XRB}} = 10^{41} - 10^{43}$ erg/s. With time, there are more and more galaxies with lower SFR. As a consequence the diagram extends towards low sSFR values, and also towards lower XRB population luminosity. At $z = 0$, the X-ray luminosity of the XRB population in the Illustris galaxies is $L_{\mathrm{XRB}} \leqslant 10^{41}$ erg/s.

5.2 Relative hard X-ray luminosity contributions of the AGN and the X-ray binary populations

Fig. 8 compares the luminosity contributions from the AGN to the luminosity contribution from XRBs. Here, we use the model of Lehmer et al. (2019). We show the contributions for different redshift with different colours. There are 50 bins (per dimension) in the 2D histogram and a contour line is drawn if at least five galaxies are present. We draw a grey solid line to show equal contributions from the XRB population and the AGN. Since we only consider galaxies with $M_\star \geq 10^9$ M$_\odot$, the minimum XRB luminosity is $L_{\mathrm{XRB}} = 10^{38.15}$ erg s^{-1} (minimum possible value from the empirical XRB scaling relations).

At high redshift ($z = 4$), we find that in the Illustris, TNG100, and SIMBA simulations the X-ray luminosity of the AGN is higher than the luminosity of the XRB population for most of the galaxy population. At $z = 4$, almost all the galaxies are on the main sequence, independently of their stellar mass. The emission of the XRB population is more important for galaxies with higher sSFR, at fixed stellar mass (see Fig. 2). The XRB populations have high luminosities of $L_{\mathrm{XRB}} = 10^{41} - 10^{43}$ erg/s. In these high-redshift star-forming galaxies, there is also gas available to feed the central BHs, and therefore, we also find bright AGNs in these galaxies with $L_{\mathrm{AGN}} = 10^{41} - 10^{45}$ erg/s, on average. At $z = 4$, > 90 per cent of the galaxies in Illustris, TNG100, and SIMBA

have a higher relative AGN luminosity $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$. We note that the picture is different in the EAGLE simulation. The XRB population in the EAGLE simulation covers the same luminosity range. However, there is a larger diversity of luminosity for the AGN, and they have (much) lower luminosity than the other simulations, on average. As a result, at $z = 4$ a large fraction of the galaxies have $L_{\mathrm{XRB}} > L_{\mathrm{AGN}}$ (below the grey line). Only 13 per cent of the galaxies have high AGN luminosities compared to their XRB populations.

With time, the luminosity of the XRB population generally decreases, as well as the luminosity of the AGN. In all the simulations, more and more galaxies reach a regime with $L_{\mathrm{XRB}} > L_{\mathrm{AGN}}$. At $z = 0$, most of the galaxies have $L_{\mathrm{XRB}} > L_{\mathrm{AGN}}$ in EAGLE (only 5 per cent of galaxies have $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$). This is also the case for a large fraction of the Illustris simulation with 10 per cent of galaxies with $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$. Finally, we note that in TNG100 and SIMBA a large number of galaxies still host bright AGNs (~ 63 per cent for both simulations), and therefore are still in the $L_{\mathrm{AGN}} > L_{\mathrm{XRB}}$ regime (above the grey line).

Now if we only consider faint AGNs in Fig. 8, i.e. by only considering the galaxies below an horizontal line at $L_{\mathrm{AGN}} \leqslant 10^{40}$ erg/s or $L_{\mathrm{AGN}} \leqslant 10^{42}$ erg/s (shown as grey dashed lines in Fig. 8), we find that a significant fraction of the galaxies are dominated by the XRB populations. We quantify this in Table 2. The percentage of XRB dominated galaxies with AGN luminosities $L_{\mathrm{AGN}} \leq 10^{42}$ erg s^{-1} is 92 per cent in Illustris at redshift $z = 0$ (compared to 90 per cent in the full sample). This increases further if we only include galaxies with AGN luminosities $L_{\mathrm{AGN}} \leq 10^{40}$ erg/s (98 per cent). We find similar trends for TNG100, EAGLE, and SIMBA.

5.3 Comparison between the XRB hard X-ray luminosity and the AGN luminosity as a function of galaxy properties

In the previous section, we have considered all galaxies independently of their properties ($M_\star$, SFR). In the following, we
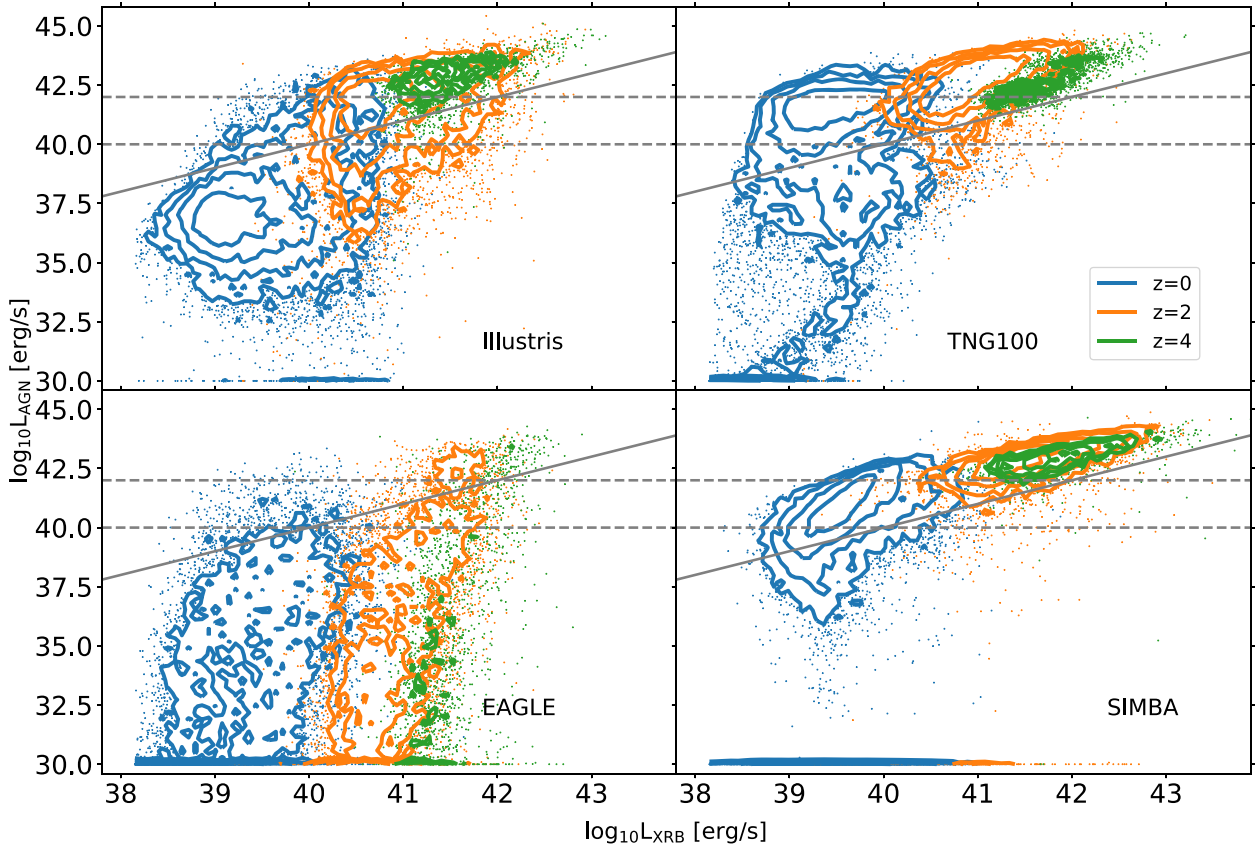
Figure 8. Hard X-ray (2–10 keV) luminosity of the AGN and XRB population of the simulated galaxies. Galaxies with AGN luminosity smaller than 10^{30} erg s^{-1} are plotted at that luminosity. Different colours indicate different redshifts. We only consider galaxies with $M_\star \geqslant 10^9$ M$_\odot$. We draw a grey solid line to show equal luminosities from the XRB population and the AGN. The fraction of galaxies with $L_{XRB} > L_{AGN}$ increases at lower redshifts. In the faint AGN regime (illustrated by the regions below the horizontal dashed lines) more and more galaxies have higher XRB luminosity than AGN luminosity with time.

Table 2. Percentage (per cent) of all the simulated galaxies with $M_\star \geqslant 10^9$ M$_\odot$ with $L_{XRB} > L_{AGN}$, at redshift $z = 4, 2, 0$. In the second and third columns, we only consider galaxies with a hard X-ray (2–10 keV) AGN luminosity $\leq 10^{42}$ erg s^{-1} or $\leq 10^{40}$ erg s^{-1} and show the percentages of galaxies with $L_{XRB} > L_{AGN}$ again. We neglect AGN obscuration here.

| | | Percentage of galaxies with $L_{XRB} > L_{AGN}$ (per cent) | | |
		All galaxies	$L_{AGN} \leqslant 10^{42}$ erg s^{-1}	$L_{AGN} \leqslant 10^{40}$ erg s^{-1}
Illustris	$z = 0$	90	92	98
Illustris	$z = 2$	39	56	>99
Illustris	$z = 4$	8	44	100
TNG100	$z = 0$	37	42	85
TNG100	$z = 2$	6	17	99
TNG100	$z = 4$	2	11	–
EAGLE	$z = 0$	95	95	99
EAGLE	$z = 2$	93	98	100
EAGLE	$z = 4$	87	98	100
SIMBA	$z = 0$	37	39	78
SIMBA	$z = 2$	9	42	100
SIMBA	$z = 4$	3	37	100

investigate the relative luminosities of the AGN compared to the XRB populations as a function of galaxy properties.

For illustration we show in Fig. 9 the L_{AGN}/L_{XRB} ratio between AGN and XRB luminosity as function of stellar mass and SFR for

all simulations at $z = 0$. The median L_{AGN}/L_{XRB} vary from one simulation to another. In Illustris and EAGLE, there is a significant population of galaxies for which the XRB population outshine the AGN, and even more so for lower mass galaxies and for galaxies with lower SFR. In TNG and SIMBA, instead, AGNs dominate the galaxy X-ray luminosity for $M_\star \leqslant 10^{10.5}$ M$_\odot$ and also in galaxies with $\log_{10}$ SFR/(M$_\odot$/yr) $\geqslant -1$.

We quantify our findings and the redshift evolution below, and in Table 3. In Illustris, almost all low-mass galaxies have a higher XRB luminosity than AGN luminosity, for all the sSFR galaxy groups. Indeed, we find that 81 per cent of the low-mass *high-sSFR* galaxies with $M_\star = 10^9$–$10^{9.5}$ M$_\odot$ have a higher XRB luminosity than AGN luminosity. We find 96 per cent in the *intermediate-sSFR* sample and > 99 per cent in the *low-sSFR* sample. For more massive galaxies in the range $M_\star = 10^{9.5}$–$10^{10.5}$ M$_\odot$, the XRB contribution dominates in 75 per cent in the *high-sSFR* sample, 91 per cent in the *intermediate-sSFR* sample, and 93 per cent in the *low-sSFR* sample. The corresponding values for the other simulations are given in Table 3. At high redshift, the galaxy total luminosity is dominated by the AGN emission in general for all the simulations, except EAGLE which produces fainter AGNs. More galaxies are dominated by the XRB emission with time, for all the Illustris, TNG100, EAGLE, and SIMBA simulations. In more detail, the number of Illustris and EAGLE galaxies dominated by XRB at low redshift decreases at higher galaxy stellar mass, for all the sSFR galaxy subsets. However, in TNG100 the number of galaxies with $L_{XRB} > L_{AGN}$ increases for
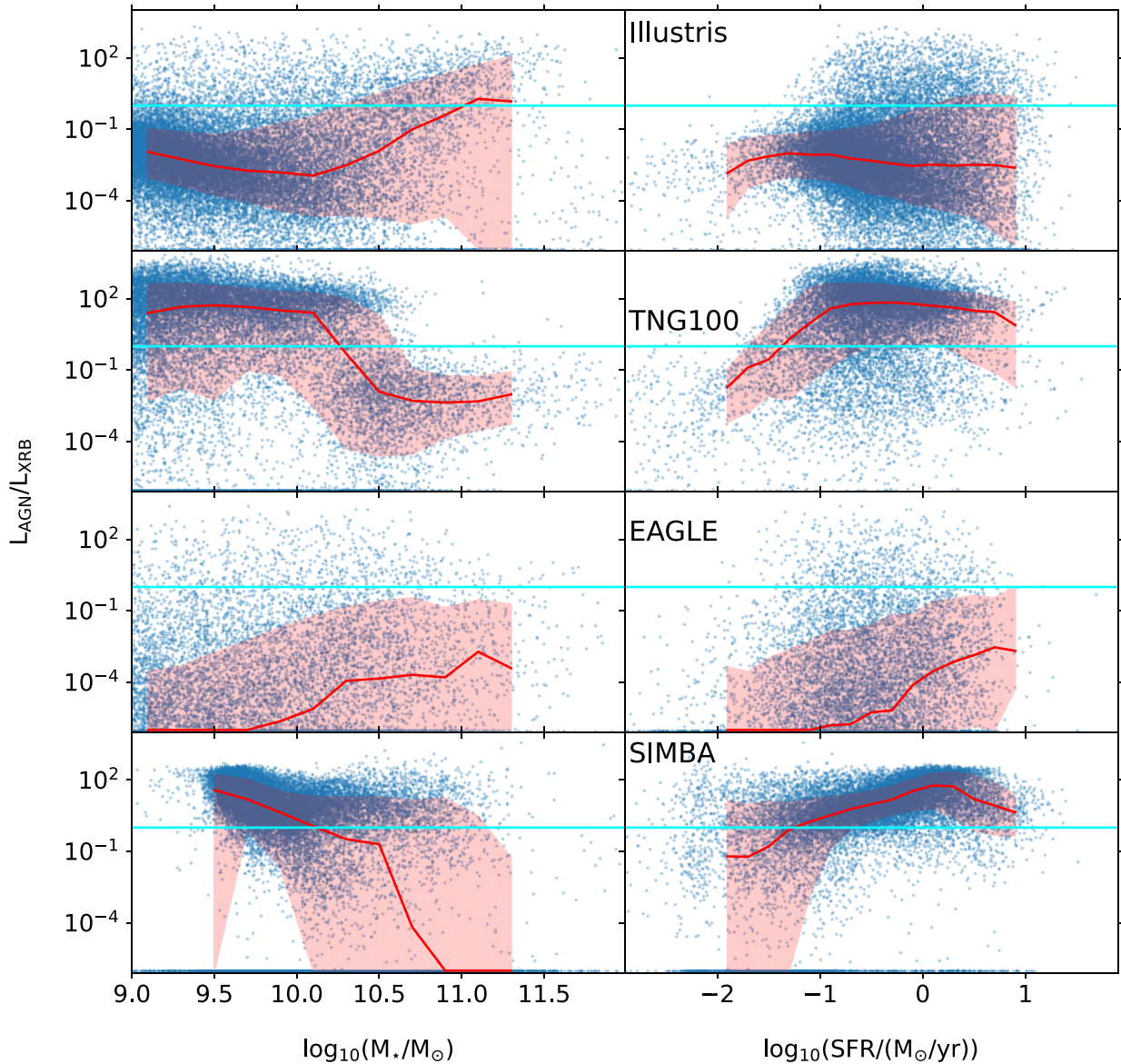
Figure 9. Ratio between AGN and XRB luminosities at $z = 0$. The blue lines indicate equal contributions and red lines show the running median in bins of width 0.2 dex. The shaded regions are the 15–85th percentiles of the distributions. Ratios below 10^{-6} are displayed at this value. The median L_{AGN}/L_{XRB} varies from one simulation to another. While there is a significant population of galaxies with $L_{AGN} < L_{XRB}$ in Illustris and EAGLE for all $M_\star$, we find that AGNs dominate the luminosity for $M_\star \leqslant 10^{10.5}\,\mathrm{M_\odot}$, and also $\log_{10} \mathrm{SFR/(M_\odot/yr)} \geqslant -1$ in TNG and SIMBA.

more massive galaxies. This is because at low redshift a significant fraction of massive galaxies experiences quenching. In general for all the simulations, we find that the XRB emission is more likely to dominate the total galaxy emission in low-sSFR galaxies than in higher sSFR galaxies.

In this section, we demonstrated that while the simulations can have similar trends (more *low-sSFR* galaxies with $L_{XRB} > L_{AGN}$), the XRB emission can dominate already at high redshift (EAGLE), or not (Illustris, TNG100, SIMBA), can dominate more in low-redshift low-mass galaxies (Illustris, EAGLE) or more in massive galaxies (TNG100).

Our findings here have a large implication for the detection of AGNs in dwarf galaxies. As indicated in Table 3, we find that all the simulations presented here have a very large fraction

of low-mass galaxies ($M_\star \leqslant 10^{9.5}\,\mathrm{M_\odot}$) with $L_{XRB} > L_{AGN}$ at $z = 0$. This makes the confirmation of the presence of an AGN in these galaxies challenging, and is in agreement with results from the current search for these AGNs (Reines, Greene & Geha 2013; Baldassare et al. 2015; Mezcua et al. 2016, 2018; Mezcua, Suh & Civano 2019; Greene et al. 2020; Reines et al. 2020). Interestingly, some simulations predict a large fraction of $L_{XRB} > L_{AGN}$ galaxies independently of the SFR of these galaxies (Illustris, EAGLE), some other simulations predict that AGN emission dominates in star-forming main-sequence and starburst galaxies (TNG100, SIMBA). Faint AGNs in low-mass galaxies should be more detectable in hard X-rays in galaxies forming stars more efficiently. If the trend identified in TNG100 is correct, an enhancement of X-ray emission due to the AGN would be observed more

Table 3. The percentages of all galaxies dominated by XRB emission ($L_{XRB} > L_{AGN}$) are given for different sSFR and stellar mass samples. Percentages of galaxies hosting faint AGNs with $L_{AGN} \leq 10^{42}$ erg s^{-1} with $L_{XRB} > L_{AGN}$ are also given. Percentages for subsamples with less than 10 galaxies are written in parentheses. We do not consider AGN obscuration here.

| | Redshift | $M_\star$ (M$_\odot$) | Percentage of all galaxies and faint AGN hosts with $L_{XRB} > L_{AGN}$ (per cent) | | | | | |
| | | | *Low-sSFR* | | *Intermediate-sSFR* | | *High-sSFR* | |
			All galaxies	Faint AGN hosts	All gal	Faint hosts	All gal	Faint hosts
Illustris	$z=0$	10^9–$10^{9.5}$	>99	>99	96	96	81	83
		$10^{9.5}$–$10^{10.5}$	93	93	91	92	75	83
		$10^{10.5}$–$10^{11.5}$	50	57	74	78	(67)	(100)
	$z=3$	10^9–$10^{9.5}$	28	37	17	38	3	(100)
		$10^{9.5}$–$10^{10.5}$	32	52	22	70	0	–
		$10^{10.5}$–$10^{11.5}$	16	(100)	12	100	–	–
TNG100	$z=0$	10^9–$10^{9.5}$	82	82	15	17	7	10
		$10^{9.5}$–$10^{10.5}$	85	86	12	16	6	11
		$10^{10.5}$–$10^{11.5}$	99	99	64	78	17	36
	$z=3$	10^9–$10^{9.5}$	0	(0)	3	6	4	(50)
		$10^{9.5}$–$10^{10.5}$	6	70	3	33	0	–
		$10^{10.5}$–$10^{11.5}$	56	83	1	(100)	(0)	–
EAGLE	$z=0$	10^9–$10^{9.5}$	>99	>99	96	96	95	95
		$10^{9.5}$–$10^{10.5}$	96	96	93	93	94	97
		$10^{10.5}$–$10^{11.5}$	91	92	89	90	(50)	(50)
	$z=3$	10^9–$10^{9.5}$	92	100	99	99	91	100
		$10^{9.5}$–$10^{10.5}$	82	96	81	97	24	(80)
		$10^{10.5}$–$10^{11.5}$	57	86	72	100	–	–
SIMBA	$z=0$	$10^{9.5}$–$10^{10.5}$	50	50	2	2	5	(67)
		$10^{10.5}$–$10^{11.5}$	73	76	25	28	–	–
	$z=3$	$10^{9.5}$–$10^{10.5}$	8	21	4	27	10	(100)
		$10^{10.5}$–$10^{11.5}$	22	97	6	(100)	–	–

often in low-redshift star-forming galaxies than in more quiescent galaxies.

6 GALAXY TOTAL HARD X-RAY LUMINOSITY OF FAINT AGN HOSTS IN SIMULATIONS AND COMPARISON TO OBSERVATIONS

In the previous sections, we have analysed the luminosity of the faint AGN, but also the luminosity of the XRB population in their host galaxies, as a function of galaxy properties. In this section, we predict the average total hard X-ray (2–10 keV) luminosity of the faint AGN host galaxies, from the Illustris, TNG100, EAGLE, and SIMBA simulations. We compare our predictions to recent observations of stacked galaxies (Fornasini et al. 2018).

6.1 Total hard X-ray luminosity of faint AGN host galaxies in simulations

We compare the total X-ray luminosity of the galaxies, defined as $L_{total} = L_{AGN} + L_{XRB}$, in the different simulations. We do not include the emission from the hot gas. Here, we purposely work in the hard X-ray band (2–10 keV), in which the emission from the hot gas is thought to be the smallest contribution and much lower than the XRB contribution (Lehmer et al. 2016). The hot gas contributes to the X-ray emission with a diffuse, soft thermal component (Fornasini et al. 2018). We address the contribution of the hot gas further in Section 7.

In Fig. 10, we show the median of the total X-ray luminosity of the simulated galaxies for different stellar mass bins for the four simulations in coloured lines. We show the median AGN luminosity (without obscuration) as black solid lines. Similarly as in Fig. 6, we

only consider the galaxies hosting faint AGNs, below the individual detection limit of the X-ray COSMOS survey. In the regime of faint AGN host galaxies, we have demonstrated in the previous section that the luminosity of the XRB populations could overshine those of the AGN. This, of course, also depends on the X-ray scaling relations that we use to compute the XRB population luminosity. To understand the impact of these different relations, we show the median values of the galaxy total luminosity after applying the different scaling relations (indicated by different linestyles for the coloured lines). The shaded regions and individual galaxies (plotted as points) are only given for the XRB scaling relation of Lehmer et al. (2019).

In general, the median galaxy total hard X-ray luminosity increases with stellar mass. The shape of the total luminosity as a function of the galaxy stellar mass follows the shape of the median AGN luminosity, when the latter is brighter on average than the XRB population. In Illustris and EAGLE, the median total luminosity increases with stellar mass, up to the most massive galaxies. In TNG100, we find a strong evidence for the efficient low-accretion mode AGN feedback effect in massive galaxies, especially in the low-SFR sample. We find the same feature in the SIMBA populations. Depending on the scaling relations for the XRB that we employ, the decrease due to AGN feedback of the galaxy total hard X-ray luminosity in massive galaxies can be washed out by high luminosities of the XRB or can also still be there in case of lower XRB luminosities.

Different XRB scaling relations have a strong influence on the total luminosity, changing it by more than one order of magnitude depending on the relations employed. In Illustris, TNG100, and SIMBA, the choice of the XRB scaling relation particularly affects the low redshifts $z \leq 1$. We note that it affects TNG100 and SIMBA even more in the low-sSFR galaxies (red lines): particularly in massive galaxies of $M_\star \geq 10^{10.2}$ M$_\odot$ in TNG100 at $z \geq 1$ and all
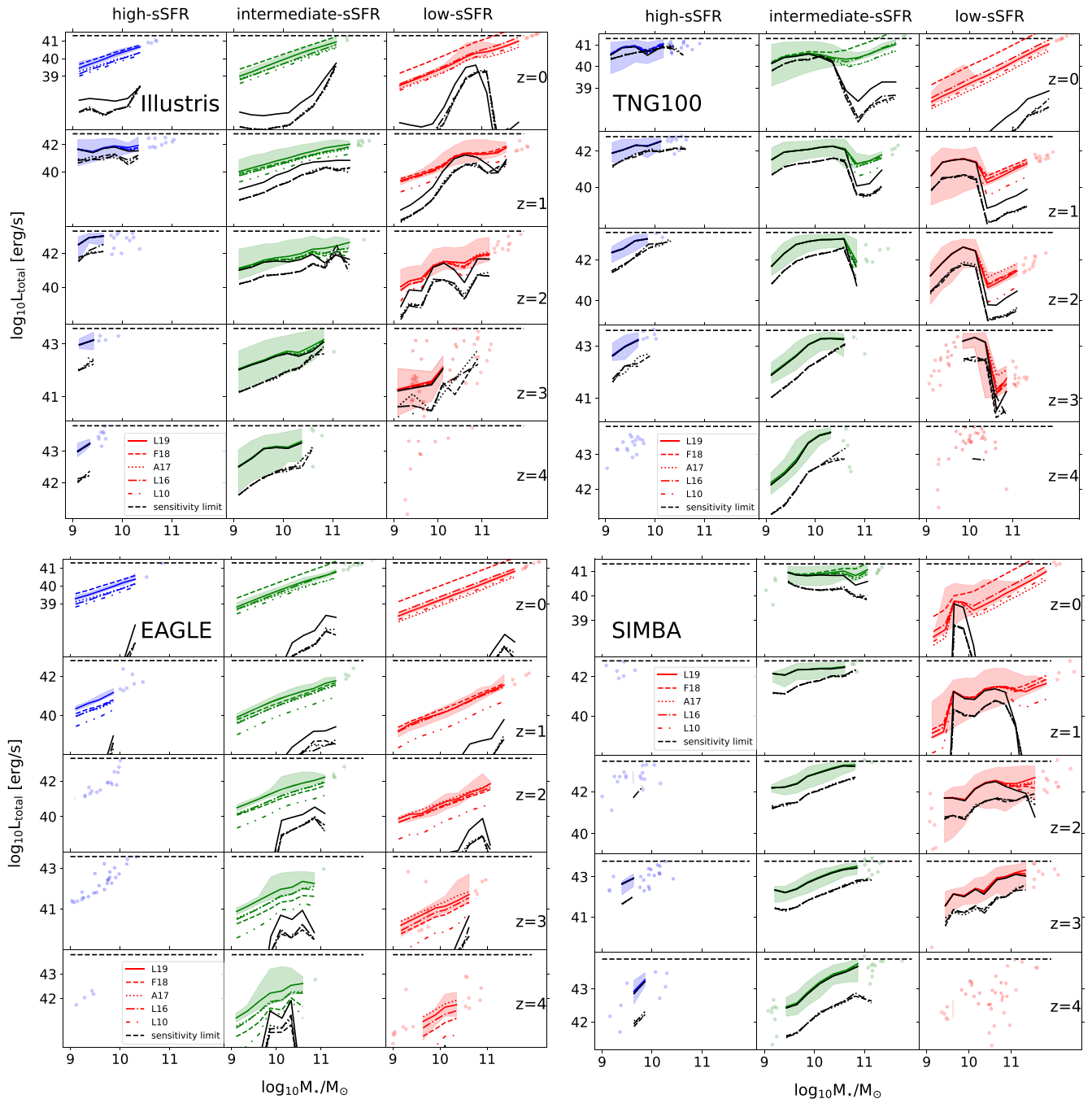
Figure 10. Total galaxy hard X-ray (2–10 keV) luminosity ($L_{total} = L_{AGN} + L_{XRB}$) as a function of galaxy stellar mass for different sSFR samples. We only include galaxies hosting too faint AGNs to be detected by current X-ray facilities as individual sources, i.e. below the COSMOS Legacy detection limit. Shaded regions indicate the 15–85th percentiles. Percentiles and individual galaxies are only shown for the XRB correction from Lehmer et al. (2019). Different line styles of the coloured lines indicate different XRB empirical scaling relations, as explained in Table 1. Black solid lines indicate the median AGN luminosity (without obscuration) as shown in Fig. 6, and entering in the computation of L_{total}. We also show the effect of AGN obscuration with the other black lines: (dashed for model M1, dotted for M2, dash–dotted for M3, and loosely dotted for M4). A larger fraction of obscured faint AGN leads to a more linear shape of the L_{total}–$M_\star$ relation, because the total luminosity is dominated by the XRB contribution.

galaxies at $z = 0$, and both in low-mass galaxies of $M_\star \leqslant 10^{9.5}\,M_\odot$ and massive galaxies of $M_\star \geqslant 10^{10.5}\,M_\odot$ in SIMBA at all redshifts. In EAGLE, the choice of the XRB scaling relation impacts all galaxy types (*high-sSFR*, *intermediate-sSFR*, and *low-sSFR* samples), at all redshifts. The differences that we find here in the impact of the scaling relation are inversely related to the median of the AGN luminosity

(black lines in Fig. 10). If the AGNs are luminous and brighter than the XRB population the choice of the scaling relation does not make a huge difference, because the XRBs do not contribute significantly to the total luminosity of the galaxies. This is the case in most of the simulations at high redshift. Instead, if the median AGN luminosity is small, as it is the case in EAGLE at all redshifts, the contribution

of the XRBs becomes more important, and the choice of the XRB scaling relation as well.

The obscuration of a given fraction of the faint AGN could reinforce the contribution of the XRB in the total X-ray luminosity of the galaxies. In Fig. 10, we show the median hard X-ray luminosity of the AGN populations with different models for their obscuration in dashed (model M1), dotted (M2), dash–dotted (M3), and loosely dotted (M4) lines. We only plot the models for which we decrease the luminosity of the obscured faint AGN by one order of magnitude. For the Illustris, TNG100, and SIMBA simulations, the largest impact of our obscuration models is found at the highest redshift ($z = 4$). At this redshift the obscuration can lead to almost one order of magnitude decrease in the AGN median luminosity. In general, we find very little difference between our models M1, M2, M3, and M4 in the median of the AGN luminosity. At lower redshifts, the effect of obscuration slightly diminishes but is still significant. In some simulations, the effect of obscuration decreases with the host galaxy stellar mass. In practice, if there is a large fraction of obscured AGNs among the faint AGNs, the relative contribution of the XRB population to the observable galaxy X-ray luminosity would be larger. Consequently the shape of the L_{total}–$M_\star$ relation would be almost completely driven by the XRB in these simulations, for all galaxy sSFR groups, and at any stellar mass $M_\star$ (except e.g. the low-mass end of TNG100 and SIMBA).

6.2 Comparison of the galaxy total hard X-ray luminosity with recent observations

Here, we compare the population of simulated faint AGNs to observations of stacked galaxies presented in Fornasini et al. (2018). The observational study presents the total galaxy X-ray luminosity extracted from ∼75 000 stacked star-forming galaxies in the redshift range $0.1 < z < 5$, using the *Chandra* COSMOS Legacy survey.

6.2.1 Mimicking the observation analysis: defining new sSFR galaxy samples

We follow Fornasini et al. (2018), and we divide the simulated galaxy populations into three new subsets. The *high-sSFR* subset now consists of galaxies with an sSFR higher than $10^{-8.5}$ yr^{-1}. The *intermediate-sSFR* subset includes galaxies whose sSFR range from $10^{-9.5}$ to $10^{-8.5}$ yr^{-1}. Finally, the *low-sSFR* subset includes all galaxies with a sSFR lower than $10^{-9.5}$ yr^{-1}. To mimic the observation selection, we also now exclude galaxies with sSFR $\leqslant 10^{-11}$ M$_\odot$/yr from our low-sSFR sample. For the simulations, this means that we remove galaxies with very low sSFR, which in some simulations corresponds to a non-negligible number of quenched galaxies. The new samples are shown in Fig. A1 for all the simulations. Compared to our previous samples, the most important difference is that the SFR range covered by galaxies in a given sample will not change with redshift. For example, galaxies in the *intermediate-sSFR* sample are mostly on the star-forming main sequence of the simulations at $z = 4$ or $z = 2$, while they are clearly among the most star-forming galaxies at $z = 0$. We provide a summary of the sample definitions employed in this section and in the previous sections in Table 4.

6.2.2 Mimicking the observation analysis: the COSMOS sensitivity curve and upper luminosity limits

As before, we only include galaxies whose AGN could not be detected as a point source by the X-ray instrument *Chandra* (COS-

MOS survey), i.e. galaxies with AGN luminosities smaller than the sensitivity limit of the COSMOS survey.

The COSMOS survey luminosity sensitivity to detect objects depends on a set of parameters, including their distances to the instrument (i.e. their redshift), and the sensitivity of the instrument. Here, we use the flux-area sensitivity curve from Civano et al. (2016) (their fig. 16, for the $0.5 - 2$ keV band). The sensitivity limit of the COSMOS survey depends on the area of sky covered by the instrument, with a lower sensitivity at the edges of the pointings that are combined to build the survey. We compute the $0.5 - 2$ keV flux sensitivity of the COSMOS survey for the Illustris, TNG100, EAGLE, and SIMBA simulations at redshifts $z = 0, 1, 2, 3, 4$ (using the simulations cosmology). We convert these sensitivity limits from the $0.5 - 2$ keV band to the $2 - 10$ keV band with the following k-correction:

$$F_{2-10\,\text{keV}} = F_{0.5-2\,\text{keV}} \frac{10^{2-\gamma} - 2^{2-\gamma}}{2^{2-\gamma} - 0.5^{2-\gamma}} (1+z)^{\gamma-2} \qquad (13)$$

with $\gamma = 1.4$ (De Luca & Molendi 2004). For redshift $z = 0$, we directly use the same flux sensitivity limit as Fornasini et al. (2018). The flux received from a given object with an intrinsic luminosity L depends on the redshift of this object: $F = L / 4\pi d_{\text{L}}^2$, with d_{L} the luminosity distance. The luminosity distance d_{L} is computed from the angular diameter distance with $d_{\text{L}} = (1 + z)^2 d_{\text{A}}$. We invert the flux/luminosity equation and convolve it with the instrument flux sensitivity limit to compute the luminosity limit for individual AGN detection in the COSMOS survey. We use this luminosity upper limit in the following: we only keep galaxies with AGNs fainter than this limit. At $z = 0$, the hard X-ray COSMOS limit is $L_{\text{AGN}} \sim 2 \times 10^{41}$ erg/s for all the simulations. The limit is $L_{\text{AGN}} \sim 6.2 - 6.5 \times 10^{42}$, $1.9 - 2.8 \times 10^{43}$, $3.8 - 5.9 \times 10^{43}$, $6.2 - 8.0 \times 10^{43}$ erg/s for $z = 1, 2, 3, 4$, respectively. The exact value for each simulation depends on the cosmology and volume of the given simulation, but the variations do not affect our findings.

The effect of the different cuts that we use is shown in Fig. 11, at $z = 2$ for illustration. Solid lines indicate the median AGN luminosity in the three galaxy samples *high-sSFR*, *intermediate-sSFR*, and *low-sSFR*, without any cuts. We first apply the upper limit for individual detection in the COSMOS survey (dashed lines). In Illustris, this cut leads to a decrease of the median luminosity for the *intermediate-sSFR* galaxies (green lines) with $\log_{10} M_\star/\text{M}_\odot \gtrsim 10$ up to one order of magnitude at $\log_{10} M_\star/\text{M}_\odot \approx 10.8$. In TNG100, the median luminosity decreases for $\log_{10} M_\star/\text{M}_\odot \gtrsim 9.5$. The effect of this cut increases with the galaxy stellar mass. This cut does not affect significantly the *low-sSFR* galaxies. For the *high-sSFR* and *intermediate-sSFR* galaxies, the cut has a similar effect in the SIMBA simulation. We note a stronger effect for the *low-sSFR* galaxies than in the TNG100. The effect of the upper luminosity cut is the smallest in the EAGLE simulation.

The lower sSFR $\geqslant 10^{-11}$ yr^{-1} cut only affects the median values for low-sSFR galaxies (red lines) and its effect can be seen as the difference between the dashed line (when we apply the COSMOS cut) and the dotted line (COSMOS cut + sSFR cut). In Illustris, the sSFR cut does not have any effect on the median luminosity. In TNG100, the sSFR cut leads to an increase of two orders of magnitude of the median AGN luminosity in the stellar mass bin $10.25 < \log_{10} M_\star/\text{M}_\odot < 10.5$ and to an increase of one order of magnitude in the stellar mass bin $10.5 < \log_{10} M_\star/\text{M}_\odot < 10.75$. This corresponds to the stellar mass ranges in which a significant fraction of these massive galaxies have reduced SFR due to the efficient quenching of the kinetic AGN feedback mode in the TNG model (see Fig. 3). Many *low-sSFR* galaxies with small luminosities

Table 4. Different galaxy samples and different cuts applied to the simulations.

|	
Analysis of the large-scale cosmological simulations Illustris, TNG100, EAGLE, and SIMBA. Used in sections 2, 4, 5, 6.1 and Figs 3, 5, 6, 10	
Galaxy samples defined as redshift- and simulation-dependent.	
High-sSFR sample	0.5 dex above the star-forming main sequence.
Intermediate-sSFR sample	1 dex around the star-forming main sequence.
Low-sSFR sample	0.5 dex below the star-forming main sequence.
Comparison to Fornasini et al. (2018) Used in Section 6.2 and Fig. 11, 12, 13, B1, C1, C2	
Galaxy samples identical to Fornasini et al. (2018).	
High-sSFR sample	$\mathrm{sSFR/yr} > 10^{-8.5}$
Intermediate-sSFR sample	$10^{-9.5} < \mathrm{sSFR/yr} < 10^{-8.5}$
Low-sSFR sample	$10^{-11} < \mathrm{sSFR/yr} < 10^{-9.5}$
Cuts to mimic the study of Fornasini et al. (2018)	
sSFR cut	$\mathrm{sSFR} \geqslant 10^{-11}\ \mathrm{yr}^{-1}$
Upper luminosity cut	Detection of individual AGN with the COSMOS survey in the 2–10 keV band. Depends on redshift and volume of the simulation. AGN brighter than this upper limit are removed from the samples.
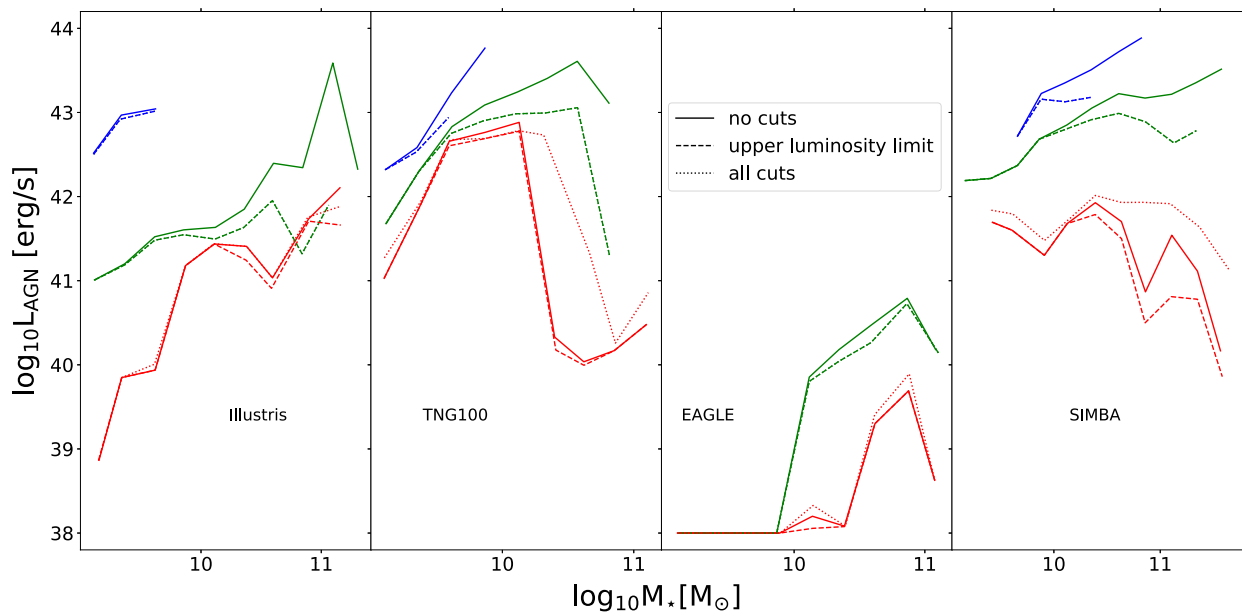
Figure 11. Impact of the COSMOS sensitivity cut for individual AGN detection (dashed lines) and the sSFR cut (dotted lines) on the median AGN hard X-ray luminosities (2–10 keV, solid lines) for the three galaxy samples of all the simulations (blue for *high-sSFR* galaxies, green for *intermediate-sSFR* ones, and red for *low-sSFR* galaxies). Only galaxies with $\mathrm{sSFR} \geqslant 10^{-11}\ \mathrm{yr}^{-1}$, and hosted AGN fainter than individual detections in the COSMOS surveys are considered here. Median luminosity values lower than $10^{38}\ \mathrm{erg\ s}^{-1}$ are shown at this value. We only show the results for $z = 2$, for which we have statistics in the three galaxy samples, but we find similar results for other redshifts. The COSMOS sensitivity cut leads to a lower median AGN luminosity and the sSFR cut leads to a higher median AGN luminosity in the *low-sSFR* sample.

are excluded which leads to an increase in the median luminosity. We find a similar behaviour in the SIMBA simulation. There is almost no impact of the sSFR cut on the median AGN luminosity in the EAGLE, only a small increase of the luminosity for the *low-sSFR* galaxies.

6.2.3 Comparison of the galaxy total X-ray luminosity to observations

We show the median of the total X-ray luminosity of the galaxies $L_{\mathrm{total}} = L_{\mathrm{AGN}} + L_{\mathrm{XRB}}$ (excluding the hot gas component) in Fig. 12

for the *high-sSFR* sample (blue, left-hand panels), the *intermediate-sSFR* galaxies (green, middle panels), and the *low-sSFR* galaxies (red, right-hand panels). Here, we decided to work with the median of the galaxy luminosity but the mean luminosity provides almost identical values. All the limits/cuts explained above are applied.[2]

[2]Galaxies with $\mathrm{sSFR} < 10^{-11}\ \mathrm{yr}^{-1}$ are excluded as well as galaxies with AGN fainter than the individual COSMOS detection limit which is indicated by the black dashed line in all the panels. The COSMOS upper limits (dashed black lines) are shown for Illustris only in Fig. 12, which is why points for other simulations may be higher.
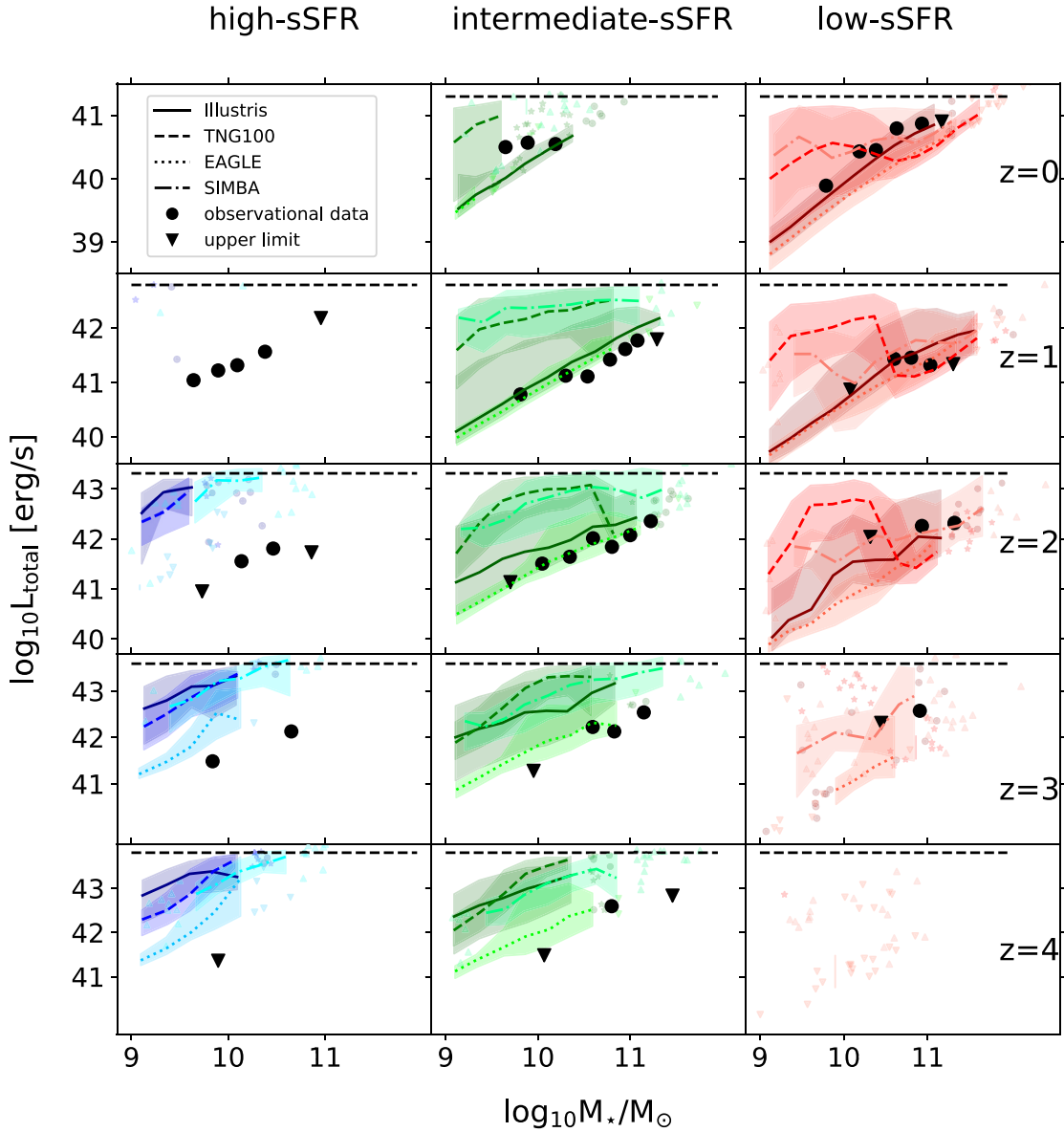
Figure 12. Median galaxy total hard X-ray (2–10 keV) luminosities ($L_{total} = L_{AGN} + L_{XRB}$). Different line styles are used for the Illustris, TNG100, EAGLE, and SIMBA simulations. The galaxies are divided into three different sSFR subsets: red for sSFR $< 10^{-9.5}$ yr^{-1}, green for $10^{-9.5}$ yr^{-1} $<$sSFR $< 10^{-8.5}$ yr^{-1}, and blue for sSFR $> 10^{-8.5}$ yr^{-1}. Furthermore, galaxies with sSFR $< 10^{-11}$ yr^{-1} are excluded as well as galaxies with AGN fainter than the individual COSMOS detection limit which is indicated by the black dashed line (we only show the limit computed for the Illustris simulation for simplicity). The lines indicate the median values for stellar mass bins with at least 10 galaxies, otherwise we show the individual galaxies (coloured circles for Illustris, stars for TNG100, downward triangles for EAGLE, and upward triangles for SIMBA). The shaded regions indicate the 15–85th percentile range of each subset. Here, the XRB correction from Lehmer et al. (2019) is used. The results are compared to the observed total luminosities of stacked galaxies from Fornasini et al. (2018; black dots). The triangle symbols indicate upper limits of the observations. The uncertainty of the observational data is smaller than the width of the symbols used here.

The results are compared to the observed total luminosities of stacked galaxies from Fornasini et al. (2018; black dots). The contribution from the hot gas is included in the total X-ray luminosity in Fornasini et al. (2018), and not in the simulations. The triangle symbols indicate upper limits of the observations. The observed X-ray luminosities are corrected for obscuration based on the measured X-ray hardness ratios which provide a rough measure, and mostly impact the low-redshift data (Fornasini et al. 2018). We discuss this further in the following section.

In the observations, the total X-ray luminosity of the galaxies increases with their stellar mass for both the galaxies in the *high-sSFR* and *intermediate-sSFR* samples. For the *low-sSFR* sample, the correlation is less obvious (except at $z = 0$, for which an increase with $M_\star$ is favoured), and the total X-ray luminosity is more or less constant for massive galaxies with $M_\star \geqslant 10^{10.5}$ M$_\odot$. In the following, we describe our general findings, first for these models ignoring the impact of obscuration, on the agreement between the simulations and the observations for each of the galaxy sSFR subsets.

High-sSFR samples (sSFR $> 10^{-8.5}$ yr^{-1}):

We find that all the simulations have a higher median of the X-ray total luminosity of their galaxy population for the *high-sSFR* samples, compared to the observational constraints of Fornasini et al. (2018). The luminosity for the simulations does not include the hot gas component, and therefore, the difference with the observations could be potentially larger. All the simulations have an increasing galaxy total luminosity with the stellar mass of the galaxies, in agreement with the trend in the observations.

Intermediate-sSFR samples ($10^{-9.5}$ yr^{-1} < sSFR < $10^{-8.5}$ yr^{-1}):

The trend obtained for the *intermediate-sSFR* sample of simulated galaxies is also similar to the observed one. We find a good agreement between the EAGLE simulation and the observations, for all redshift except $z = 0$. The other simulations overpredict the total X-ray luminosity of their *intermediate-sSFR* galaxies, on average. The overprediction is stronger towards less massive galaxies, and can be more than one order of magnitude.

Low-sSFR samples (10^{-11} yr^{-1} < sSFR < $10^{-9.5}$ yr^{-1}):

Finally, for the *low-sSFR* sample galaxies, we find that on average, the TNG100 and SIMBA simulations overpredict the total luminosity for the faint AGN low-mass host galaxies. A better general agreement is found for the EAGLE and SIMBA simulations, which show an increasing total luminosity with increasing stellar mass. In more detail, we find that the best agreement at $z = 3, 2$ is obtained for the SIMBA simulation. At $z = 1$, the Illustris, EAGLE, and SIMBA simulations provide a good agreement with the observations. While the Illustris and EAGLE have increasing galaxy total luminosity with stellar mass, the X-ray median luminosity is decreasing in SIMBA up to $M_\star \sim 10^{10}$ M$_\odot$, and then increasing slightly to $M_\star \sim 10^{11}$ M$_\odot$. For more massive galaxies, the median luminosity of SIMBA stays constant, as in the observations. At $z = 0$, the best agreement is found for the Illustris simulation. This is particularly true for the low-mass galaxies $M_\star \leqslant 10^{10}$ M$_\odot$ where the median luminosity of the Illustris simulation is lower than the SIMBA simulation.

For different redshifts and galaxy types, we find that the highest/lowest galaxy total X-ray luminosity is not always produced by the same simulations. We develop this in the following. For the three galaxy samples and all the redshifts, we find that the EAGLE simulation produces the lowest median total X-ray luminosity (dotted lines). This is because EAGLE is the simulation with the largest population of faint AGN (Rosas-Guevara et al. 2016). The AGN X-ray luminosity function of EAGLE is below all the other simulations (e.g. Sijacki et al. 2015; Habouzit et al. 2019; Thomas et al. 2019). The luminosity function is in agreement with Aird et al. (2015) and Buchner et al. (2015) in the range $\log_{10} L_{\mathrm{x},2-10\,\mathrm{keV}} \leqslant 10^{44}$ erg/s at high redshift for $z \geq 2$, in the range $\log_{10} L_{\mathrm{x},2-10\,\mathrm{keV}} \leqslant 10^{43}$ erg/s at $z = 2$, and is underestimated otherwise. At $z = 0$, the EAGLE simulation is below the constraints of Aird et al. (2015) and Buchner et al. (2015). These constraints do not cover AGN with fainter luminosity than $\log_{10} L_{\mathrm{x},2-10\,\mathrm{keV}} = 10^{42}$ erg/s. We find that for the *high-sSFR* samples the highest median is in Illustris for $M_\star \leqslant 10^{10}$ M$_\odot$, and TNG100 for $M_\star \sim 10^{10}$ M$_\odot$, and in SIMBA for more massive galaxies. For the *intermediate-sSFR* samples, TNG100 has the highest median for galaxies with $10^{9.5}$ M$_\odot \leqslant$ M$_\star \leqslant 10^{10.5}$ M$_\odot$ at $z = 2, 3, 4$, while SIMBA has the highest median at $z = 1$ and for galaxies with $M_\star \geqslant 10^{10.5}$ M$_\odot$ at $z = 1, 2, 3, 4$. TNG100 produces the highest total luminosity median for *low-sSFR* sample galaxies of $M_\star \leqslant$ a few 10^{10} M$_\odot$, and SIMBA and Illustris for more massive galaxies.

Interestingly, we know from Fig. 11 that the fact that observations here exclude galaxies with sSFR < 10^{-11} yr^{-1} can artificially boost the galaxy total X-ray luminosity for the *low-sSFR* galaxies. This would particularly affect the median in the bins containing massive

galaxies. In our analysis we use the same sSFR cut as in the observations of Fornasini et al. (2018), however, this does not ensure that the sSFR distributions of the simulated and observed samples are the same. Depending on the number of galaxies with sSFR close to the limit sSFR < 10^{-11} yr^{-1} in the observations, we may still be counting too many simulated galaxies with very low sSFR. This could explain the lower total luminosity in massive galaxies, particularly for the TNG100 and SIMBA simulations (as they were already affected by the sSFR < 10^{-11} yr^{-1} cut in Fig. 11).

Observing faint systems is challenging. We cannot rule out that the faintest systems that we see in simulations are actually missed by the observations. Observations are always limited when trying to observe fainter and fainter systems. The completeness of the observational samples is particularly difficult to address at the faint end of the galaxy distribution. This can potentially create significant discrepancies with the samples of simulated galaxies, since in the latter we have access to the faintest galaxies and AGNs. Therefore, in a separate test, we applied different lower luminosity cuts to our samples of simulated galaxies. In practice, to mimic what could take place in the observations we assumed that (i) galaxies with total X-ray luminosity below 10^{40} or 10^{41} erg/s (for $z = 0, 1$ and $z = 2, 3, 4$, respectively, following Fornasini et al. 2018) are not detected, or (ii) that galaxies with total luminosity two or three orders of magnitude below the COSMOS Legacy limit are not detected. We found that missing the faintest galaxies when *observing* the simulations would mostly impact the total luminosity of low-mass galaxies with $M_\star \leqslant 10^{10.5}$ M$_\odot$ for the strongest luminosity cut. The median/mean luminosity of the detected galaxies would be increased compared to the median/mean luminosity of the full galaxy population with the same physical characteristics, but it would not strongly impact our main conclusions regarding the agreement between simulations and the observations of Fornasini et al. (2018).

6.2.4 Comparison of the galaxy total X-ray luminosity to observations when accounting for AGN obscuration

The observations of Fornasini et al. (2018) already include some rough estimate for obscuration, assuming that the XRB and the AGN are subject to the same obscuration (which may or may not be a valid assumption, see the discussion in Fornasini et al. 2018). Their correction mostly impacts the total luminosity at low redshift. An upper limit on the impact of the correction is ∼0.4 dex difference for $z < 0.6$, 0.2 dex difference in the range $0.6 < z < 2.3$, and negligible differences at higher redshift. Given the uncertainties on this *average* correction applied to the XRB and AGN, and in order to quantify the role of AGN obscuration on the total luminosity, we apply our correction to the simulations.

In Fig. 13, we show the same figure as Fig. 12, i.e. the median of the galaxy hard X-ray luminosity, but this time we account for obscured AGN. We note that there are no strong differences between the model variation M1, M2, M3, M4 in our results, which means that the fractions of obscured AGN with $L_{\mathrm{AGN}} \leqslant 10^{41}$ erg/s does not impact the total galaxy luminosity. In practice, this is because the median of the galaxy X-ray luminosity is often higher than 10^{41} erg/s. We apply the M1 obscuration model to the simulated AGN samples. More precisely, we show two sets of obscuration models, for which we either decrease by one order of magnitude the AGN luminosity (*fainter AGN* models), or either assume $L_{\mathrm{AGN}} = 0$ (*missed AGN*). We describe below the impact of our two sets of obscuration models on the total galaxy luminosity, and re-examine our previous conclusions.
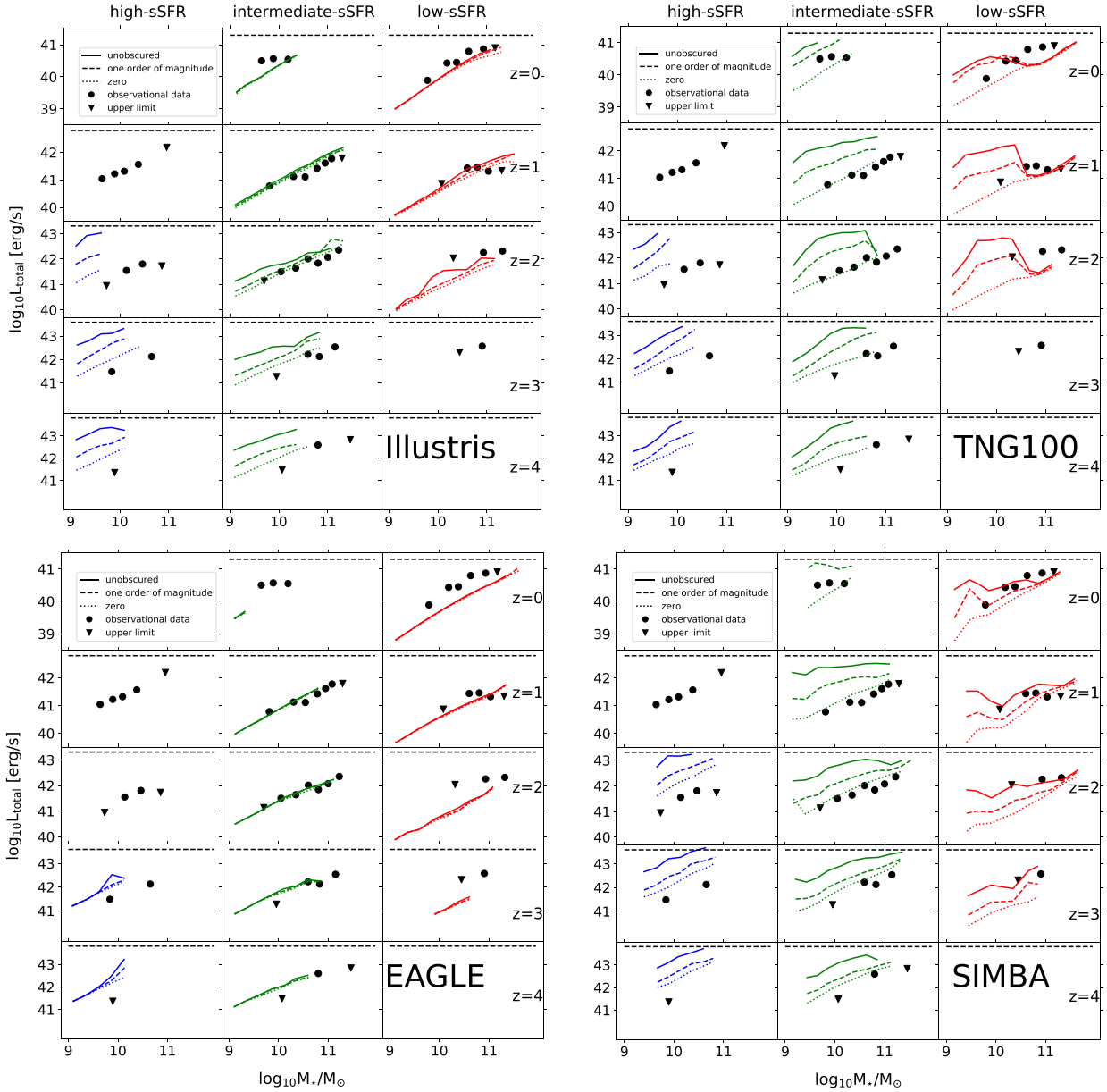
Figure 13. Median galaxy total hard X-ray (2–10 keV) luminosities ($L_{\text{total}} = L_{\text{AGN}} + L_{\text{XRB}}$). Same as Fig. 12 but here we account for given fractions of obscured AGN (see Fig. 1). Solid lines show L_{total} without considering that any AGN is obscured. The dashed lines indicate the model where the luminosity of obscured AGN is reduced by one order of magnitude (*fainter AGN* M1 model) and the dotted lines indicate the model where the luminosity of obscured AGN is set to zero (*missed AGN* M1 model). The results are compared to the observed total luminosities of stacked galaxies from Fornasini et al. (2018) (black symbols). Typically, the median (or mean) L_{total} can be affected by about half an order of magnitude when we account for AGN obscuration (*fainter AGN* model), and even more if we assume that the obscured AGN would be completely missed by the observations (*missed AGN* model). With strong AGN obscuration models, the AGN features that were visible in the galaxy total luminosity (i.e. non-linearity of the L_{tot}–$M_\star$ relation, peak due to large fraction of bright AGN, decrease of L_{tot} because of AGN feedback) can be completely erased. Here, we used the XRB scaling relation of Lehmer et al. (2019) which predicts the highest XRB luminosity among all XRB models. Using other XRB models or having more obscured AGN leads to smaller total galaxy luminosity.

In general, we find that obscuration can decrease the total galaxy luminosity by up to one order of magnitude for the *fainter AGN* models. When we assume that we would completely miss the obscured AGN (*missed AGN* models), the impact can be higher and reach two orders of magnitudes at stellar masses for which simulations produce a large fraction of AGN. This is the case in the TNG100 and SIMBA simulations, which have a large fraction of AGN for galaxies with stellar masses of $M_\star \leqslant 10^{10.5}\,\text{M}_\odot$. When we set $L_{\text{AGN}} = 0$ for the obscured AGN,

the total galaxy luminosity is fully driven by the XRB population.

For the *high-sSFR* galaxies, adding the obscuration models does not impact our previous conclusions: all the simulations overpredict the total galaxy luminosity. For all the simulations except EAGLE, the galaxy luminosity of the *high-sSFR* subsets is dominated by the AGN luminosity, which explains the decrease of the luminosity when we obscure some of the AGN. We note that even when we fully remove the obscured AGN, i.e. when the galaxy luminosity

is dominated by the XRB population, the luminosity is higher than for the observations. This could mean that the empirical scaling relations for the luminosity of the XRB population overpredict the median/mean hard X-ray galaxy luminosity.

For galaxies in the *intermediate-sSFR* group and for most of the simulations, our fiducial model with the XRB scaling relation of Lehmer et al. (2019) and the *fainter-AGN* obscuration models still produce higher galaxy luminosity (see Fig. 13). A better agreement is found if we assume lower XRB scaling relations (Lehmer et al. 2010, 2016; Aird et al. 2017; Fornasini et al. 2018) or that the obscured faint AGNs are completely missed by the observations.

For galaxies in the *low-sSFR* subsets in Illustris and EAGLE, we find that the total galaxy luminosity is in general smaller than in the observations. For these two simulations, the discrepancy would be larger for other XRB scaling relations (as shown in Fig. 10) or obscuration models accounting for more obscured AGN. In SIMBA, the agreement between the total galaxy luminosity from simulations (with the L19 XRB relation and without AGN obscuration) and observations is good for $z > 1$. We note some differences at $z = 0$, with a higher galaxy luminosity median for simulated galaxies with $M_\star \leqslant 10^{10}\,M_\star$, and lower median for $M_\star \geqslant 10^{10.5}\,M_\star$ than in observations. Obscuration models assuming more obscured AGN would in general lower the good agreement of SIMBA with observations. TNG100 has a particular shape of the total galaxy luminosity median, which is not clearly visible in observations, with a higher luminosity median in low-mass galaxies, and lower luminosity median in massive galaxies, compared to observations. Assuming more obscured AGN in the simulations reduces the differences with observations.

From our analysis we find that for some simulations such as Illustris or EAGLE, whose total luminosity at low redshift is dominated by XRB, are not impacted by our modelling of AGN obscuration. For the other simulations with a higher contribution of the AGN luminosity to the total galaxy luminosity, without considering our obscuration correction we often find an excess of luminosity in the simulations compared to the observations. We have demonstrated that this excess can be reduced by obscuration at low redshift for all *intermediate-sSFR* galaxies and *high-sSFR* galaxies with $M_\star \leqslant 10^{10.5}\,M_\odot - 10^{11}\,M_\odot$ (depending on the simulation). Similarly, the excess could be due to an overestimate of the obscuration in the observation data (Fornasini et al. 2018) at low redshift.

7 DISCUSSION

In this section, we discuss different aspects that could impact our results, and our comparison with current observations of galaxy total hard X-ray luminosity.

7.1 Calibration of the cosmological simulations and their AGN X-ray luminosity functions

The BH subgrid models of these simulations have been broadly calibrated against one of the empirical M_{BH}–M_{bulge} relations available in the literature. While the Illustris, TNG100, and EAGLE simulations adjust the efficiency parameter of the AGN feedback model, the simulation SIMBA instead calibrates the accretion efficiency. None of the simulations studied here were calibrated against the AGN luminosity function.

At $z = 0$, these simulations have X-ray luminosity functions in good overall agreement with the constraints of, e.g. Buchner et al. (2015). TNG100 has a higher normalization of the luminosity function (Habouzit et al. 2019), and EAGLE a lower normalization,

for $\log_{10} L_{AGN}/(\mathrm{erg/s}) \leqslant 43$ (Rosas-Guevara et al. 2016). Illustris and SIMBA lie within the constraints of Buchner et al. (2015) for this luminosity regime (see Sijacki et al. 2015; Thomas et al. 2019, for the analysis of AGN properties with observational constraints). At higher redshift, all the simulations except EAGLE overpredict the observational constraints in the range $\log_{10} L_{AGN}/(\mathrm{erg/s}) \leqslant 42.5 - 43$ (Habouzit et al. 2021). The regime that we investigate here ($\log_{10} L_{AGN}/(\mathrm{erg/s}) \leqslant 42$) is below the range which is usually constrained by observations of the X-ray luminosity function. The differences between the simulations and the observational constraints[3] on the X-ray luminosity function could affect the galaxy total X-ray luminosity in a non-trivial way, for the different galaxy sSFR samples.

7.2 Comparison to observations

To make the analysis of the simulations consistent with the observations from Fornasini et al. (2018) (stacked galaxies from the *Chandra* COSMOS Legacy survey), we have adopted the same sSFR cut (excluding simulated galaxies with sSFR $\leqslant 10^{-11}\,\mathrm{yr}^{-1}$), and the same upper luminosity limit to exclude galaxies with AGN detectable by the *Chandra* COSMOS Legacy survey. These same cuts do not ensure that the sSFR distributions of the observation and simulation samples are similar. For example, if a *high-sSFR* simulation sample has a distribution peaking at larger sSFR values than the corresponding observational sample, this will lead to higher total galaxy luminosities in the simulations than in observations. Therefore, we cannot exclude that some discrepancies between the observations and the simulations are due to different distributions of their respectives sSFR samples.

Two additional caveats for our comparison in this paper are the X-ray emission from hot gas and the obscuration.

7.2.1 Hot gas contribution

In our analysis, we have neglected the contribution to the X-ray emission from the hot interstellar medium (ISM). This contribution was, however, studied in the observations considered here. The hot gas is expected to contribute to the X-ray emission with a diffuse, soft thermal component (Fornasini et al. 2018). The hot gas component is thought to be significant and to dominate the X-ray emission over the XRB emission for rest-frame energy of <1.5 keV (Lehmer et al. 2016, their fig. 4). For high energies of >1.5 keV, as it is the case in our analysis, the XRB emission is expected to dominate (Lehmer et al. 2016). There is currently no definitive quantification (from observations or theory) across galaxy masses, galaxy apertures, types, and redshifts. The contribution of the hot gas could not only depend on energy band, but could also decrease with increasing redshift, and could be important mostly for $z < 1$ (Lehmer et al. (2016), results obtained assuming that the spectral energy distribution (SED) of star-forming galaxies do not change strongly with redshift. According to Mineo et al. (2012b), the hot gas X-ray emission could increase with the SFR of a galaxy. In that case, the influence of the hot gas emission would be higher in the high-sSFR galaxy sample at a fixed stellar mass and at higher stellar masses at fixed sSFR. From a numerical and theoretical perspective, definitive assessments of the X-ray emission from the hot gas is still needed, but some preliminary

---

[3]See also Sijacki et al. (2015), Weinberger et al. (2018), Volonteri et al. (2016), Rosas-Guevara et al. (2016), and Thomas et al. (2019) for studies of the BH populations in the different simulations.

estimates of the gas X-ray emission within galaxies in the soft bands have been derived from the Illustris and TNG simulations (Truong et al. 2020). We postpone the task of including the contribution of the X-ray emitting hot gas within galaxies to future work, as this requires a careful assessment of the dependence on X-ray wavelength and on aperture within which the mock or real observations are taken.

7.2.2 Obscuration

Gas and dust obscuration can play a crucial role when comparing observations to simulations. It is not clear yet if the obscuration originates from material close to the galaxies nuclear region. In that case, the luminosity of the XRB population distributed within the galaxies would not be strongly obscured, but the AGN would be. The obscuration could also originate from material in the whole galaxy, and in that case it would lead to the obscuration of both the XRB population and the AGN. Buchner & Bauer (2017) find that the gas on the galaxy scale is only responsible for a part of the Compton-thin AGN, and does not provide Compton-thick lines of sight. The heavily obscured (Compton-thick) AGNs would therefore mostly result from obscuration in the nuclear region. If this is the case, the hard X-ray emission from the galaxy-wide XRB population would be less impacted than the emission from the AGN.

Simulations do not consistently capture the obscuration of the AGN, as obscuration can arise from regions close to the AGN on spatial scales below the simulation resolution. Therefore, we have tested the role of AGN obscuration by applying four different models to the simulated AGN samples. Our models depend on redshift and hard X-ray (2–10 keV) AGN luminosity for AGN with $L_{AGN} \geqslant 10^{41}$ erg/s (Habouzit et al. 2019, for more details), and only on hard X-ray AGN luminosity for fainter AGN. Obscuration could also depend on galaxy SFR, a parameter that we do not consider here. We have not applied any further obscuration model to the XRB populations.

In practice, the galaxy total hard X-ray luminosity can be impacted significantly depending on the fraction of obscured AGN that we assume. We find that the more the faint AGNs are obscured the more the shape of the total galaxy luminosity as a function of galaxy stellar mass is driven by the XRB luminosity (since we do not apply any obscuration model to the XRB emission). If the linearity of the XRB empirical L_{XRB}–$M_{\star}$ scaling relations is a good estimate (as found in e.g. Lehmer et al. 2019) and in the presence of a large population of obscured faint AGN (or just a population of very faint AGN, such as in EAGLE), we should observe a linear L_{total}–$M_{\star}$ relation in the observations, independently of the galaxy sSFR. In that case, we find that any deviation from a linear L_{total}–$M_{\star}$ relation would be due to features of the AGN populations.

7.3 Detection of AGNs in dwarf galaxies

It is now demonstrated that AGNs can exist in dwarf galaxies (Reines et al. 2013, 2020; Baldassare et al. 2015; Mezcua et al. 2016, 2018, 2019; Greene et al. 2020). Quantifying the fraction of galaxies hosting BHs and the BH mass distribution in dwarf galaxies can constrain the theoretical models of BH formation (Greene 2012). Since these galaxies have not evolved much over cosmic times compared to their massive counterparts, they could have retained the initial properties of BH formation: initial BH mass and initial BH formation efficiency. While the BH occupation fraction in low-mass galaxies in such large-scale simulations may not be relevant/accurate because of the simple seeding of BHs in massive

galaxies or haloes, the AGN occupation fraction is fundamental to address the connection between the AGN and their host galaxies (e.g. the correlations between AGN activity and the SFR of their galaxies).

The AGN found in dwarf galaxies in observations can generally be qualified as faint AGN (Mezcua et al. 2016; Chilingarian et al. 2018; Mezcua & Domínguez Sánchez 2020, and references therein). In this work, we have shown that in the Illustris, TNG100, and EAGLE simulations the XRB population in galaxies of $M_{\star} \leqslant 10^{9.5}$ M$_{\odot}$ can outshine the AGN emission in hard X-rays. This is a significant issue when trying to detect an AGN in X-rays; detection in X-ray is one of the most common method to detect low-mass AGN in low-mass galaxies today. What is interesting is that the simulations do not predict the same trend with SFR. In Illustris and EAGLE, the XRB population outshine the AGN in > 90 per cent of the galaxies, whether these galaxies form stars efficiently (starburst) or not (below the star-forming main sequence). However, in TNG100, the XRB population outshine the AGN only in galaxies below the main sequence, but not in main-sequence galaxies or starburst galaxies. There, AGN activity is enhanced when SFR activity is enhanced. Confronting current and future observations of AGN in dwarf galaxies to our results on the AGN population predicted by cosmological simulations will help us to understand the observations and at the same time to constrain our modelling of BH and galaxy physics in simulations.

8 CONCLUSION

We have analysed the properties of the faint AGN (in the hard X-ray band 2–10 keV, $L_{AGN} \leqslant 10^{42}$ erg/s) and their host galaxies of the four large-scale cosmological simulations Illustris, TNG100, EAGLE, and SIMBA. We have modelled the contribution from the XRB population and from the AGN (including their possible obscuration) to the total galaxy hard X-ray luminosity. We summarize below our main findings.

(i) The properties of the faint AGN host galaxies vary from simulation to simulation (Fig. 4). Faint AGN of $L_{AGN} \sim 10^{38}$ erg/s can be powered by relatively massive BHs and be located in massive galaxies ($M_{\star} \gtrsim 10^{10}$ M$_{\odot}$) with reduced SFR (TNG100, SIMBA), or be powered by lower mass BHs in less massive galaxies ($M_{\star} \lesssim 10^{10}$ M$_{\odot}$) still forming stars (Illustris, EAGLE).

(ii) We find that the two possible behaviours described above depend on the effectiveness of AGN feedback in massive galaxies. In TNG100 and SIMBA, the efficient feedback taking place in massive galaxies reduces both their sSFR (Fig. 4), but also the ability of their BHs to accrete efficiently. AGN feedback drives the build up of the faint AGN population in these simulated galaxies.

(iii) In all the simulations, except EAGLE, most of the galaxies have brighter AGN than the XRB population, at high redshift ($z > 2$, Fig. 8). With time, the AGN number density decreases and consequently more and more galaxies have a brighter XRB population than their AGN. The general fainter population of AGN in EAGLE (at all redshifts) compared to the other simulations leads to a significant number of galaxies with a brighter XRB population than AGN.

(iv) The relative contribution of the AGN and the XRB population to the XRB + AGN total galaxy hard X-ray luminosity depends on the stellar mass and the SFR of the galaxies (Fig. 10). Starburst galaxies host brighter AGN in all simulations and across redshift: the AGN luminosity dominates over the XRB luminosity in most of these galaxies. At low redshifts ($z \leqslant 3$), the efficient AGN feedback in TNG100 and SIMBA leads to a strong decrease of the median

AGN luminosity in massive galaxies ($M_\star \geqslant 10^{10}\,\mathrm{M}_\odot$) with reduced sSFR, and more galaxies are dominated by XRB emission.

(v) In low-mass galaxies of $M_\star \leqslant 10^{9.5}\,\mathrm{M}_\odot$ at $z = 0$, we find that the XRB emission always outshines the AGN emission in low-sSFR galaxies in all the simulations (Table 3). The XRB still dominates in main-sequence and startburst galaxies in Illustris and EAGLE, but does not outshine the AGN in TNG100. This has important implications for the search of AGN in dwarf galaxies.

(vi) The total AGN + XRB hard X-ray luminosity of faint AGN host galaxies (i.e. neglecting the hot ISM X-ray emission) increases with increasing $M_\star$, for all redshifts and all the simulations (Fig. 10). We note a turnover for the massive TNG100 and SIMBA galaxies for which the lower AGN median luminosity (due to AGN feedback) propagates to the total galaxy hard X-ray luminosity.

(vii) We find that a non-linear L_{total}–$M_\star$ relation in faint AGN galaxies (Fig. 10 and Fig. 13) is explained by a non-linear L_{XRB}–$M_\star$ scaling relation (in that case XRB luminosity models need to be updated), or by peaks of AGN activity at some stellar masses. We find that the obscuration of faint AGN can completely erase these AGN signatures in the L_{total}–$M_\star$ relation (see Fig. 13). In that case, the shape of the L_{total}–$M_\star$ relation would be fully driven by the XRB emission.

(viii) The simulations, with our modelling of AGN and XRB luminosity, tend to overestimate the total AGN + XRB galaxy X-ray luminosity in the high-sSFR sample and for most of the simulations in the intermediate-sSFR sample (neglecting the hot gas ISM X-ray emission) compared to the observations of the COSMOS Legacy stacked galaxies (Fornasini et al. 2018). Simulated galaxies with sSFR $> 10^{-9.5}\,\mathrm{yr}^{-1}$ are too bright. Galaxies with $10^{-9.5}\,\mathrm{yr}^{-1} <$ sSFR $< 10^{-8.5}\,\mathrm{yr}^{-1}$ are also too bright, except a good agreement for EAGLE. For low-sSFR galaxies of $10^{-11}\,\mathrm{yr}^{-1} <$ sSFR $< 10^{-9.5}\,\mathrm{yr}^{-1}$, we find that some simulations underestimate or overestimate the median galaxy luminosity (Fig. 12).

(ix) In both simulations and observations (Fornasini et al. 2018), *high-sSFR* galaxies have higher total galaxy X-ray luminosity than *low-sSFR* galaxies at fixed stellar mass, in general (Fig. 12).

(x) The empirically driven XRB scaling relations used in this work span 0.5 dex in luminosity (at fixed $M_\star$), which is about the same order of magnitude as some of our obscuration models (Fig. 13). These two aspects are highly degenerate and further observational constraints will be needed to disentangle them.

Our work and predictions pave the way for upcoming and concept space missions such as Athena, AXIS, and Lynx. These missions will increase by several orders of magnitude the sensitivity of the current X-ray instruments, and will allow us to make promising progress on our understanding of faint AGN, a luminosity regime that as we have demonstrated can be dominated by XRBs for specific sSFR and $M_\star$ regimes.

DATA AVAILABILITY

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.

REFERENCES

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892

Aird J., Coil A. L., Georgakakis A., 2017, MNRAS, 465, 3390

Aird J., Coil A. L., Georgakakis A., 2019, MNRAS, 484, 4360

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

Baldassare V. F., Reines A. E., Gallo E., Greene J. E., 2015, ApJ, 809, L14

Basu-Zych A., Lehmer B., Hornschemeier A. E., Ptak A., 2013a, Exploring the Influence of Metallicity on X-ray Binary Formation in Nearby and Distant UV-selected Galaxies, AAS/High Energy Astrophysics Division, American Astronomical Society, HEAD meeting #13, id.301.04

Basu-Zych A. R. et al., 2013b, ApJ, 762, 45

Basu-Zych A., Lehmer B., Hornschemeier A. E., Fragos T., Zezas A., Yukita M., Ptak A., 2016, A Local Perspective on HMXB Populations in the Early Universe, AAS/High Energy Astrophysics Division, American Astronomical Society, HEAD meeting #15, id.402.03

Bluck A. F. L. et al., 2016, MNRAS, 462, 2559

Bonfield D. G. et al., 2011, MNRAS, 416, 13

Boroson B., Kim D.-W., Fabbiano G., 2011, ApJ, 729, 12

Brorby M., 2016, Chandra Proposal, 4835

Buchner J., Bauer F. E., 2017, MNRAS, 465, 4348

Buchner J. et al., 2015, ApJ, 802, 89

Çatmabacak O., Feldmann R., Anglés-Alcázar D., Faucher-Giguère C.-A., Hopkins P. F., Kereš D., 2020, preprint (arXiv:2007.12185)

Cheng Z., Li Z., Xu X., Li X., Zhu Z., Fang T., 2018, ApJ, 869, 52

Chilingarian I. V., Katkov I. Y., Zolotukhin I. Y., Grishin K. A., Beletsky Y., Boutsia K., Osip D. J., 2018, ApJ, 863, 1

Choi E., Ostriker J. P., Naab T., Johansson P. H., 2012, ApJ, 754, 125

Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91

Civano F. et al., 2016, ApJ, 819, 62

Colbert E. J. M., Strickland D. K., Veilleux S., Weaver K. A., 2004, American Astronomical Society Meeting Abstracts, Vol. 205, #155.01

Crain R. A. et al., 2015, MNRAS, 450, 1937

Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

De Luca A., Molendi S., 2004, A&A, 419, 837

Donnari M. et al., 2019, MNRAS, 485, 4817

Donnari M. et al., 2021a, MNRAS, 500, 4004

Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2021b, MNRAS, 506, 4760

Douna V. M., Pellizza L. J., Mirabel I. F., Pedrosa S. E., 2015, A&A, 579, A44

Dray L. M., 2006, MNRAS, 370, 2079

Fabbiano G., 2006, ARA&A, 44, 323

Fornasini F. M., Civano F., Fabbiano G., Elvis M., Marchesi S., Miyaji T., Zezas A., 2018, ApJ, 865, 43

Fornasini F. M. et al., 2019, ApJ, 885, 65

Fornasini F. M., Civano F., Suh H., 2020, MNRAS, 495, 771

Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013, ApJ, 776, L31

Furlong M. et al., 2015, MNRAS, 450, 4486

Gaskin J. A. et al., 2018, in den Herder J.-W. A., Nikzad S., Nakazawa K., eds, Proc. SPIE Conf. Ser. Vol. 10699, Space Telescopes and

Instrumentation 2018: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 106990N

Genel S. et al., 2014, MNRAS, 445, 175

Georgakakis A. et al., 2015, MNRAS, 453, 1946

Georgantopoulos I. et al., 2017, in Ness J.-U., Migliari S., eds, The X-ray Universe 2017, Proceedings of the conference held 6-9 June, 2017 in Rome, Italy, p. 88

Gilfanov M., 2004, MNRAS, 349, 146

Gilli R. et al., 2014, A&A, 562, A67

Greene J. E., 2012, Nat. Commun., 3, 1304

Greene J. E., Strader J., Ho L. C., 2020, ARA&A, 58, 257

Grimm H.-J., Gilfanov M., Sunyaev R., 2003, Chin. J. Astron. Astrophys. Suppl., 3, 257

Gültekin K. et al., 2009, ApJ, 698, 198

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

Habouzit M. et al., 2019, MNRAS, 484, 4413

Habouzit M. et al., 2021, MNRAS, 503, 1940

Hahn C. et al., 2019, ApJ, 872, 160

Häring N., Rix H.-W., 2004, ApJ, 604, L89

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., 2017, preprint (arXiv:1712.01294)

Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027

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

Kaaret P., 2014, MNRAS, 440, L26

Kim D.-W., Fabbiano G., 2010, ApJ, 721, 1523

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Koudmani S., Sijacki D., Bourne M. A., Smith M. C., 2019, MNRAS, 484, 2047

Koudmani S., Henden N. A., Sijacki D., 2021, MNRAS, 503, 3568

Lehmer B. D. et al., 2007, ApJ, 657, 681

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., 2014, ApJ, 789, 52

Lehmer B. D. et al., 2016, ApJ, 825, 7

Lehmer B. D. et al., 2019, ApJS, 243, 3

Linden T., Kalogera V., Sepinsky J. F., Prestwich A., Zezas A., Gallagher J. S., 2010, ApJ, 725, 1984

Liu W., Veilleux S., Canalizo G., Rupke D. S. N., Manzano-King C. M., Bohn T. U V., 2020, ApJ, 905, 166

Lutz D. et al., 2008, ApJ, 684, 853

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

Madau P., Fragos T., 2017, ApJ, 840, 39

Manzano-King C. M., Canalizo G., Sales L. V., 2019, ApJ, 884, 54

Marinacci F. et al., 2018, MNRAS, 480, 5113

Matthee J., Schaye J., 2019, MNRAS, 484, 915

Merloni A. et al., 2014, MNRAS, 437, 3550

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., Suh H., Civano F., 2019, MNRAS, 488, 685

Mezcua M., Domínguez Sánchez H., 2020, ApJ, 898, L30

Mineo S., Gilfanov M., Sunyaev R., 2012a, MNRAS, 419, 2095

Mineo S., Gilfanov M., Sunyaev R., 2012b, MNRAS, 426, 1870

Mor R., Netzer H., Trakhtenbrot B., Shemmer O., Lira P., 2012, ApJ, 749, L25

Mushotzky R. et al., 2019, BAAS, 51, 107

Naiman J. P. et al., 2018, MNRAS, 477, 1206

Nandra K. et al., 2013, preprint (arXiv:1306.2307)

Nelson D. et al., 2015, Astron. Comput., 13, 12

Nelson D. et al., 2018, MNRAS, 475, 624

Nelson D. et al., 2019a, MNRAS, 490, 3234

Nelson D. et al., 2019b, Comput. Astrophys. Cosmol., 6, 2

Pillepich A. et al., 2018a, MNRAS, 475, 648

Pillepich A. et al., 2018b, MNRAS, 473, 4077

Reines A. E., Volonteri M., 2015, ApJ, 813, 82

Reines A. E., Greene J. E., Geha M., 2013, ApJ, 775, 116

Reines A. E., Condon J. J., Darling J., Greene J. E., 2020, ApJ, 888, 36

Rosario D. J. et al., 2012, A&A, 545, A45

Rosas-Guevara Y. M. et al., 2015, MNRAS, 454, 1038

Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, MNRAS, 462, 190

Schaye J. et al., 2015, MNRAS, 446, 521

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Sparre M. et al., 2015, MNRAS, 447, 3548

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V. et al., 2018, MNRAS, 475, 676

Terrazas B. A., Bell E. F., Woo J., Henriques B. M. B., 2017, ApJ, 844, 170

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

Truong N. et al., 2020, MNRAS, 494, 549

Tzanavaris P., Hornschemeier A. E., Gallagher S. C., Lenkić L., Desjardins T. D., Walker L. M., Johnson K. E., Mulchaey J. S., 2016, ApJ, 817, 95

Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 104

Vito F., Gilli R., Vignali C., Comastri A., Brusa M., Cappelluti N., Iwasawa K., 2014, MNRAS, 445, 3557

Vito F. et al., 2016, MNRAS, 463, 348

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M. et al., 2014a, Nature, 509, 177

Vogelsberger M. et al., 2014b, MNRAS, 444, 1518

Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, MNRAS, 460, 2979

Weinberger R. et al., 2017, MNRAS, 465, 3291

Weinberger R. et al., 2018, MNRAS, 479, 4056

Zhang Z., Gilfanov M., Bogdán Á., 2012, A&A, 546, A36

Zinger E. et al., 2020, MNRAS, 499, 768

APPENDIX A: DEFINITIONS OF THE MAIN-SEQUENCE AND GALAXY NUMBER DENSITIES OF THE GALAXY SSFR SAMPLES

A1 Definition of the main sequence of all the simulations

In Table A1, we present the parameters α and β of the main sequence of each simulation for different redshifts as discussed in Section 2.4.1.

Table A1. Parameters α and β of the main sequence of each simulation for different redshifts. The galaxy populations and respective main sequences are shown in Fig. 3.

	z	α	β
Illustris	0	-10.52	1.05
	2	-8.63	0.96
	4	-8.69	1.01
TNG100	0	-8.23	0.81
	2	-7.33	0.82
	4	-7.75	0.93
EAGLE	0	-9.50	0.92
	2	-8.84	0.98
	4	-7.35	0.87
SIMBA	0	-9.07	0.94
	2	-13.73	1.50
	4	-10.80	1.21

Table A2. Percentage of galaxies (per cent) and number density of galaxies n (10^{-5} cMpc^{-3}) in the three *high-sSFR*, *intermediate-sSFR*, and *low-sSFR* samples, for redshift $z = 0, 1, 2, 3, 4$. Only galaxies with $M_\star \geqslant 10^9$ M$_\odot$ are included. This table refers to the division shown in Fig. 3 and Fig. A1.

| | | Figure 3 (simulations) | | | | | | Figure A1 (comparison to observations) | | | | | |
| | | High-sSFR | | Interm-sSFR | | Low-sSFR | | High-sSFR | | Interm-sSFR | | Low-sSFR | |
	z	per cent	n	per cent	n	per cent	n	per cent	n	per cent	n	per cent	n
Illustris	0	3.19	72.27	79.58	1801.4	17.23	390.08	0.0	0.0	2.07	46.77	97.93	2216.98
	1	2.04	34.27	88.65	1492.36	9.31	156.71	0.21	3.56	65.29	1099.14	34.49	580.65
	2	1.79	17.22	91.99	885.47	6.22	59.85	1.78	17.14	91.69	882.57	6.53	62.83
	3	1.32	6.04	96.36	440.08	2.32	10.6	11.18	51.08	88.06	402.17	0.76	3.48
	4	2.57	4.97	96.79	187.18	0.64	1.24	36.47	70.53	63.53	122.85	0.0	0.0
TNG100	0	4.67	63.1	60.69	820.52	34.64	468.31	0.0	0.0	3.5	47.25	96.5	1304.69
	1	3.21	36.78	80.63	924.83	16.16	185.39	0.08	0.96	51.73	593.33	48.19	552.72
	2	2.28	17.32	89.36	679.58	8.36	63.62	1.77	13.49	87.8	667.71	10.43	79.32
	3	1.14	4.42	94.88	369.09	3.98	15.48	16.6	64.57	81.75	318.01	1.65	6.41
	4	0.96	1.62	96.91	163.94	2.14	3.61	63.14	106.81	36.73	62.14	0.13	0.22
EAGLE	0	3.4	45.2	61.95	823.9	34.65	460.9	0.0	0.0	0.41	5.5	99.59	1324.5
	1	1.11	14.0	87.49	1101.5	11.4	143.5	0.01	0.1	59.34	747.1	40.65	511.8
	2	0.47	3.8	93.21	749.3	6.32	50.8	0.34	2.7	92.71	745.3	6.95	55.9
	3	0.73	2.9	94.34	377.0	4.93	19.7	5.11	20.4	91.27	364.7	3.63	14.5
	4	0.39	0.7	94.03	170.0	5.59	10.1	35.56	64.3	61.95	112.0	2.49	4.5
SIMBA	0	1.68	18.52	48.05	529.75	50.28	554.34	0.0	0.0	18.02	198.66	81.98	903.96
	1	1.14	7.51	66.22	434.83	32.64	214.32	0.23	1.54	67.65	444.26	32.11	210.86
	2	2.81	10.82	81.78	315.0	15.41	59.36	3.26	12.55	85.31	328.61	11.43	44.02
	3	3.9	9.97	89.45	228.37	6.65	16.98	5.3	13.52	87.92	224.47	6.79	17.33
	4	3.06	5.41	92.09	162.59	4.84	8.55	10.51	18.55	87.78	154.98	1.71	3.02

A2 Definitions and galaxy number densities of the galaxy sSFR samples

In Table A2, we present the percentages and number densities for the three different subsets that we used in Sections 3, 5, 6, and 7. We only consider galaxies that are well resolved in all the simulations, i.e. $M_\star \geqslant 10^9$ M$_\odot$. The *high-sSFR* sample groups starburst galaxies 0.5 dex above the main sequence. The *intermediate-sSFR* galaxies represent the galaxies on the main sequence (within 1 dex). Finally, *low-sSFR* galaxies are galaxies below 0.5 dex of the main sequence.

To compare the simulations to the results from observations of Fornasini et al. (2018), we have changed the definition of our galaxy sSFR samples. *High-sSFR* galaxies are defined by sSFR/yr $> 10^{-8.5}$, *intermediate-sSFR* galaxies by $10^{-9.5} <$ sSFR/yr $< 10^{-8.5}$, and *low-sSFR* galaxies by $10^{-11} <$ sSFR/yr $< 10^{-9.5}$. Fig. A1 shows the SFR as a function of the stellar mass with these new definitions. The background colours show the three sSFR samples. Contrary to Fig. 3, the division does not depend on redshift or simulation and galaxies on the star-forming main sequence of the simulations are not always included in the *intermediate-sSFR* sample. In Table A2, we add the percentages and number densities of the different subsets for our second set of definitions.
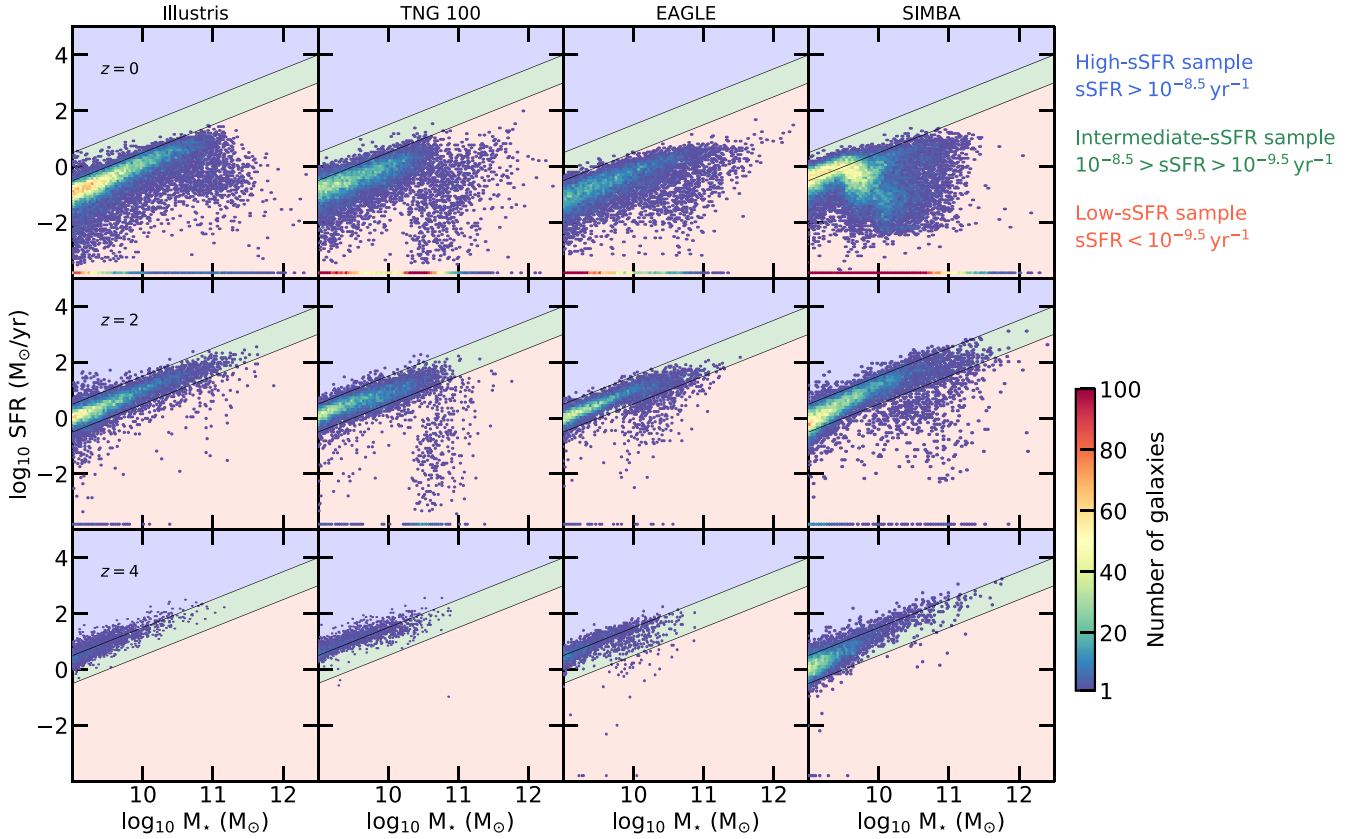
Figure A1. SFR as a function of galaxy stellar mass for the Illustris, TNG100, EAGLE, and SIMBA simulations. Hexabins are colour coded by the number of galaxies in bins. Finally, we show the three samples that we use with background colours: blue for galaxies with sSFR $> 10^{-8.5}$ yr^{-1}, green for galaxies with $10^{-9.5} <$ sSFR $< 10^{-8.5}$ yr^{-1}, and red for $10^{-11} <$ sSFR $< 10^{-9.5}$ yr^{-1}. Galaxies with $sSFR < 10^{-11}$ yr^{-1} are shown at $\log_{10}$SFR/(M$_\odot$/yr) $= -4$, but are not considered in the comparison with observations. These definitions follow Fornasini et al. (2018), and are different from the ones used in the first sections of the paper and showed in Fig. 3, especially for $z < 2$.

APPENDIX B: IN WHICH GALAXIES LIVE THE FAINT AGN: NUMBER DENSITIES OF GALAXIES PRESENT IN THE THREE AGN LUMINOSITY SUBSAMPLES

In Table B1, we report the number densities of the galaxies in the three subsets with different AGN luminosity ranges ($L_{AGN} = 10^{37.5}$–$10^{38.5}$ erg/s, $L_{AGN} = 10^{39.5}$–$10^{40.5}$ erg/s, $L_{AGN} = 10^{41.5}$–$10^{42.5}$ erg/s). From redshift $z = 3$ to $z = 1$, all number densities increase for all simulations. However, from $z = 1$ to $z = 0$, only the number densities for the galaxies with $L_{AGN} \sim 10^{38}$ erg/s increases in Illustris. In TNG100 and SIMBA, the number densities of galaxies with $L_{AGN} \sim 10^{38}$ erg/s and $L_{AGN} \sim 10^{40}$ erg/s both increase.

In Table B2, we present the p values from the Kolmogorov–Smirnov test with distributions of the galaxy samples with $L_{AGN} \sim 10^{38}$ erg/s and $L_{AGN} \sim 10^{42}$ erg/s. These values are used in Section 4.

Table B1. The number densities (in 10^{-5}cMpc^{-3}) of galaxies present in the three AGN luminosity subsamples of Fig. 4, i.e. with hard X-ray luminosity of $L_{AGN} \sim 10^{38}$, 10^{40}, 10^{42} erg/s, for the redshifts $z = 0, 1, 3$. We also report the number densities of galaxies in the full sample of galaxies, showed as grey histograms in Fig. 4.

z	L_{AGN} (erg/s)	Illustris	TNG100	EAGLE	SIMBA
0	$10^{37.5}$–$10^{38.5}$	333.21	102.02	73.3	21.70
	$10^{39.5}$–$10^{40.5}$	100.17	136.96	37.8	149.54
	$10^{41.5}$–$10^{42.5}$	42.47	314.76	9.7	107.47
	All	1975.74	1241.73	1245.5	725.20
1	$10^{37.5}$–$10^{38.5}$	239.83	22.56	107.1	6.60
	$10^{39.5}$–$10^{40.5}$	256.30	64.87	75.8	15.82
	$10^{41.5}$–$10^{42.5}$	181.13	487.63	27.4	122.63
	All	1551.72	1091.28	1229.0	379.27
3	$10^{37.5}$–$10^{38.5}$	4.55	0.15	24.4	0.03
	$10^{39.5}$–$10^{40.5}$	26.91	2.21	20.1	0.66
	$10^{41.5}$–$10^{42.5}$	150.91	167.63	19.0	26.32
	All	438.76	380.81	395.1	78.61

Table B2. p values from the Kolmogorov-Smirnov test with the distributions of the galaxy samples with $L_{AGN} \sim 10^{38}$ erg/s and $L_{AGN} \sim 10^{42}$ erg/s. We only use the test for samples with at least 50 galaxies. The null hypothesis can be rejected if $p < 0.01$, meaning that the distributions of different AGN luminosities are statistically different. We find that this is the case for almost all simulations and redshifts here.

Simulation	z	$M_\star$ p	M_{BH} p	sSFR p
Illustris	0	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
	1	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
	3	0.0005	$<10^{-5}$	0.0391
TNG100	0	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
	1	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
	3	–	–	–
EAGLE	0	0.0196	0.0017	0.1504
	1	$<10^{-5}$	$<10^{-5}$	0.0035
	3	$<10^{-5}$	$<10^{-5}$	0.0137
SIMBA	0	$<10^{-5}$	$<10^{-5}$	0
	1	$<10^{-5}$	$<10^{-5}$	$<10^{-5}$
	3	–	–	–

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