The low-end of the black hole mass function at cosmic dawn

Alessandro Trinca,[1,2,3]⋆ Raffaella Schneider [1,2,3,4] Rosa Valiante [2,3] Luca Graziani [1,3,5] Luca Zappacosta[2] and Francesco Shankar [6]

[1]*Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, I-00185 Roma, Italy*
[2]*INAF/Osservatorio Astronomico di Roma, Via di Frascati 33, I-00040 Monte Porzio Catone, Italy*
[3]*INFN, Sezione Roma1, Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, I-00185 Roma, Italy*
[4]*Sapienza School for Advanced Studies, Viale Regina Elena 291, I-00161 Roma, Italy*
[5]*INAF/Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy*
[6]*Department of Physics and Astronomy, University of Southampton, Highfield SO17 1BJ, UK*

ABSTRACT

Understanding the formation and growth of supermassive black holes (SMBHs) at high redshift represents a major challenge for theoretical models. In this work, we investigate the early evolution of the first SMBHs by constraining their distribution in mass and luminosity at $z > 4$. In particular, we focus on the poorly explored low-mass end of the nuclear black hole (BH) distribution down to $z \simeq 4$, and explore its connection with the nature of the first BH seeds and the processes governing their mass growth. To this aim, we have developed CAT (Cosmic Archaeology Tool), a new semi-analytic model that describes the formation of the first stars and BHs in a self-consistent way and follows the co-evolution of nuclear BHs and their host galaxies for a representative population at $z > 4$. We find that current observational constraints favour models where the growth of BH seeds is Eddington limited and occurs at the Bondi–Hoyle–Lyttleton rate or where super-Eddington accretion occurs via a slim disc during gas-rich galaxy mergers. The main difference between these two model variants lies at the low end of the predicted mass and luminosity functions at $4 \le z \le 6$, where a clear gap appears in the first model, reflecting the stunted growth of *light* BH seeds formed as remnants of the first stars. Detecting this signature will be extremely challenging even for the future generation of space observatories, such as *JWST*, *Athena*, and *Lynx*.

Key words: black hole physics – galaxies: active – galaxies: evolution – galaxies: formation – galaxies: high-redshift – quasars: supermassive black holes.

1 INTRODUCTION

In the last decade hyper-luminous quasars ($L \ge 10^{47}\,\mathrm{erg\,s^{-1}}$) have been observed up to $z > 7$, suggesting that supermassive black holes (SMBHs) with masses as high as 10 billion solar have already formed when the Universe was less than 1 Gyr old. The most massive detected system, SDSS J0100+2802, is a quasar observed at $z \sim 6.3$ (Wu et al. 2015) with an estimated mass of $M_{\mathrm{BH}} = (1.2 \pm 0.19) \times 10^{10}\,\mathrm{M_\odot}$, while the most distant quasars are ULAS J1342+0928 ($z = 7.54$, Bañados et al. 2018) powered by a central BH with $M_{\mathrm{BH}} \sim 8 \times 10^8\,\mathrm{M_\odot}$, *Pōniuā'ena* (J1007+2115) at $z = 7.52$ ($M_{\mathrm{BH}} = 1.5 \times 10^9\,\mathrm{M_\odot}$, Yang et al. 2020), and the recently reported J0313−1806 at $z = 7.64$ ($M_{\mathrm{BH}} = 1.6 \times 10^9\,\mathrm{M_\odot}$, Wang et al. 2021).

Explaining how these extreme objects form and grow during the first billion years of cosmic history still represents a major challenge for theoretical models (for thorough reviews see Volonteri 2010; Valiante et al. 2017; Inayoshi, Visbal & Haiman 2020). Therefore, it is crucial to understand what is the nature of SMBH first progenitors (generally referred to as BH 'seeds') and how these efficiently grow through gas accretion and mergers. Many possible scenarios for the formation of BH seeds have been proposed in the literature so far.

Light seeds, with a BH mass $\sim 100\,\mathrm{M_\odot}$, are supposed to be the remnants of the first generation of metal-free (Population III/Pop III) stars and are expected to form with masses ranging between a few 10s and a few 100s $\mathrm{M_\odot}$, inside dark matter (DM) *minihaloes* ($M_{\mathrm{halo}} \sim 10^5$–$10^6\,\mathrm{M_\odot}$) at very high redshifts ($z \ge 20$, see Bromm 2013 for a review, and Hirano et al. 2014, 2015; Hosokawa et al. 2016; Sugimura et al. 2020 for more recent works).

Collisions between stars inside dense star clusters are instead supposed to give birth to *intermediate-mass* or *medium-weight* BH seeds, with a mass $\sim 1000\,\mathrm{M_\odot}$ (Omukai, Schneider & Haiman 2008; Devecchi & Volonteri 2009; Katz, Sijacki & Haehnelt 2015; Sakurai et al. 2017; Reinoso et al. 2018, 2019; Sassano et al. 2021), although even more massive BH seeds could originate from merger episodes between stellar mass BHs (aided by strong gas inflows) inside similarly dense environments (Lupi et al. 2014; Boco, Lapi & Danese 2020).

Recent studies have also suggested that very massive BHs ($\sim 10^5$–$10^6\,\mathrm{M_\odot}$), usually referred to as *heavy seeds*, can form via the direct collapse of a supermassive star. The above mechanism is supposed to take place inside DM haloes with virial temperatures $T_{\mathrm{vir}} > 10^4\,\mathrm{K}$, where gas cooling and fragmentation is suppressed by their metal-poor composition and H_2 photodissociation caused by Lyman–Werner (LW) radiation (Omukai 2001; Bromm & Loeb 2003; Wise, Turk & Abel 2008; Regan & Haehnelt 2009; Hosokawa, Omukai & Yorke 2012; Latif et al. 2013; Inayoshi, Omukai & Tasker

⋆ E-mail: alessandro.trinca@inaf.it

2014; Chon et al. 2016; Latif, Schleicher & Hartwig 2016; Becerra et al. 2018) or by dynamical heating (Wise et al. 2019; Lupi, Haiman & Volonteri 2021) associated with strong gas inflows (Lodato & Natarajan 2006, 2007; Mayer et al. 2010, 2015; Haemmerlé et al. 2019; Mayer & Bonoli 2019).

Although the formation rate of heavy BH seeds is yet to be established (see Valiante et al. 2017; Inayoshi et al. 2020 and references therein) and their growth efficiency might depend on the formation site inside the host galaxy (e.g. Chon, Hosokawa & Omukai 2021), they represent one of the most promising formation scenarios to explain the existence of SMBHs at $z \geq 6$ without invoking super-Eddington accretion rates. It has been shown, in fact, that starting from a seed of $M_{BH} \sim 10^5 \, M_\odot$, a BH can reach the observed mass of high-redshift SMBHs through classical Eddington-limited growth (Valiante et al. 2017; Sassano et al. 2021). Many studies also pointed out that less massive seeds would require persistent accretion of gas at the Eddington rate during all their existence to reach a billion solar masses in less than ~ 1 Gyr (Volonteri 2010; Madau, Haardt & Dotti 2014; Bañados et al. 2018). Alternatively, intermittent gas accretion at super- or hyper-Eddington rates may efficiently grow lower mass seeds (Inayoshi, Haiman & Ostriker 2016; Pezzulli, Valiante & Schneider 2016; Pacucci et al. 2017; Pezzulli et al. 2017b; Takeo et al. 2018), provided that suitable conditions are met in their circumnuclear regions at high z (see the discussion in Mayer 2019 and references therein).

Despite the significant progresses made in their theoretical description, discriminating the nature of BH seeds by electromagnetic observations may be very challenging, for a number of reasons: BH seeds keep memory of their birth conditions and genetic origin as long as they live in isolation, accreting gas from their host galaxy. For seed progenitors of $z \sim 6$–7 SMBHs, these conditions are only met at $z > 10$ for <100–150 Myr since their formation (Valiante et al. 2018a). In addition, *light* BH seeds may be too faint to be detectable, even with upcoming facilities, at least in Eddington-limited growth scenarios (Valiante et al. 2018b). Hence, the only chance to constrain their nature would be to detect the gravitational waves emitted during their binary coalescence through third-generation ground-based detectors, such as the Einstein Telescope[1] (Valiante et al. 2020). Detecting the electromagnetic emission from rapidly growing *heavy* BH seeds appears more promising, and photometric selection techniques have been envisaged to help identify the more promising candidates (Pacucci et al. 2015, 2016; Natarajan et al. 2017; Valiante et al. 2018b).

An alternative way to constrain the early evolution of black holes (BHs) is by shedding light on to the low-mass end of the BH mass function (MF) at high z (Volonteri, Lodato & Natarajan 2008; Volonteri & Natarajan 2009; Ricarte & Natarajan 2018b; Piana et al. 2021), which is supposed to be very sensitive to the nature of BH seeds and their growth mode.

Several optical/Near-Infrared surveys such as the Sloan Digital Sky Survey (SDSS[2]), the UKIRT Infrared Deep Sky Survey (UKIDSS[3]), the DECam Legacy Survey (DECaLS[4]), and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS[5]) gave a fundamental contribution to the detection of luminous quasars at redshift $z \geq 6$, characterizing the bright end

of their luminosity function (LF, Willott et al. 2010a; Bañados et al. 2016). At very high redshift, however, these objects are rare and presumably trace extremely biased regions of the Universe.

Large-scale cosmological simulations predict that a much larger population of fainter active galactic nuclei (AGNs), powered by less massive BHs, is assembling and growing together with their host galaxies at these cosmic epochs (Di Matteo et al. 2012; Vogelsberger et al. 2014; Schaye et al. 2015; Feng et al. 2016). This population has eluded direct detection until recent surveys, in particular the Subaru High-z Exploration of Low-Luminosity Quasars (SHELLQS), have started to explore fainter magnitudes, unveiling a large sample of low-luminosity quasars. These findings have raised the total number of observed AGNs at $z > 6$ above 200, constraining for the first time the faint-end of their LF (Matsuoka et al. 2018, 2019).

In recent years, a number of semi-analytical models have been applied to investigate observational signatures of BH seeds and of their early co-evolution with their host galaxies (Somerville et al. 2008; Volonteri et al. 2008; Devecchi et al. 2012; Salvaterra et al. 2012; Yue et al. 2013; Bonoli, Mayer & Callegari 2014; Pezzulli et al. 2016, 2017a; Valiante et al. 2016, 2018b; Ricarte & Natarajan 2018a, b; Dayal et al. 2020; Piana et al. 2021). In these models the evolution of the baryonic components of DM haloes is described through physically and/or observationally motivated prescriptions and the growth of DM haloes can be either generated analytically or extracted from numerical simulations (Baugh 2006; De Lucia 2019).

In this work, we investigate the MF of SMBHs, and the corresponding AGNs LF, by following their redshift evolution from the epoch of BH seeds formation down to $z = 4$. In particular, we focus on the large population of low-mass faint systems that will be targeted by upcoming facilities, such as the *James Webb Space Telescope* (*JWST*), the *Advanced Telescope for High Energy Astrophysics* (*ATHENA*), and the *Lynx X-ray Observatory*. We also investigate whether initial BH seed masses and their growth rates leave an imprint on the low-mass (faint) end of the MF (LF) that may potentially be used to discriminate among different scenarios. To this aim, we developed CAT (Cosmic Archaeology Tool), a semi-analytical model that describes structure formation in the first billion years of cosmic evolution following the hierarchical growth of DM haloes, their stellar and gas content, and their nuclear BHs. CAT can account for a wide range of halo masses: from the $10^6 \, M_\odot$ minihaloes hosting the collapse of the first stars at $z = 20$–30, up to the largest galaxies with $M_h \sim 10^{12}$–$10^{14} \, M_\odot$ where the most powerful quasars at $z = 4$–7 are supposed to reside. CAT can explore a statistics which is still prohibitively expensive for cosmological hydrodynamical simulations (Di Matteo, Springel & Hernquist 2005; Springel et al. 2005; Sijacki et al. 2007, 2015; DeGraf et al. 2012, 2015; Hirschmann et al. 2014; Di Matteo et al. 2017; McAlpine et al. 2017; Weinberger et al. 2017; Habouzit et al. 2019), and can be achieved only through zoom-in techniques on a small number of systems pre-selected in very large lower resolution simulations (Lupi et al. 2019; Regan et al. 2019; Zhu et al. 2020).

CAT builds on our semi-analytical model GAMETEQSODUST (GQD), which was successfully applied to study the co-evolution of SMBHs and their host galaxies at $z \geq 6$ (Valiante et al. 2011, 2014, 2016). In the present version, CAT enables to follow the formation of both *light* and *heavy* BH seeds depending on the environmental properties and to explore their contribution to the BH MF (AGNs LF) across a wide range of redshifts, mass (luminosity) scales and physical parameters, quantifying the conditions for seed formation and their mass growth rate. This kind of investigation is certainly beyond the modelling capabilities of current large-scale cosmological hydrodynamic simulations, which usually adopt simplified prescriptions for

---

[1] http://www.et-gw.eu/
[2] https://www.sdss.org/
[3] https://www.nottingham.ac.uk/astronomy/UDS/
[4] https://www.legacysurvey.org/decamls/
[5] https://panstarrs.stsci.edu/

BH seeding and assume a fixed BH seed mass that is planted in DM haloes above a given threshold mass, independently of their internal properties (Di Matteo et al. 2012; Vogelsberger et al. 2014; Khandai et al. 2015; Schaye et al. 2015; Feng et al. 2016). More physically sound BH seeding prescriptions have been adopted in smaller scale or zoom-in simulations (Bellovary et al. 2011; Habouzit, Volonteri & Dubois 2017; Tremmel et al. 2017; Huang et al. 2020), but at the price of being unable to simultaneously explore the low- and high-mass (luminosity) ends of the BH MF (LF).

To explore how different high-z formation scenarios leave their imprints on the low-mass (luminosity) end of the BH MF (LF) at $z \geq 4$, we have run a large set of simulations on a grid of halo merger histories extending over several orders of magnitude in mass at $z = 4$ using the galaxy formation model GALFORM (Cole et al. 2000; Parkinson, Cole & Helly 2008). Our model predictions are then compared with the observed properties of galaxies and AGNs in view of future observations which could shed some light in discriminating among different BH seeding and growth scenarios.

The paper is organized as follows. In Section 2, we introduce and describe the model, while in Section 3 we show how we calibrate the model to set the free parameters that govern star formation and BH growth. The results are presented in Section 4, where we also compare our findings with the most recent observations and theoretical models. Finally, in Section 5 we discuss and summarize our main results.

2 THE COSMIC ARCHAEOLOGY TOOL

In this section, we illustrate the CAT adopted in this work. First, we present the galaxy formation model GALFORM and how its DM halo merger tree algorithm was properly adapted to generate the sample of DM haloes. Second, we describe the main features of GQD that has been imported in CAT in order to follow the evolution of halo baryonic components (gas, stars, and nuclear BHs), along with the major improvements we introduced in this work.

In what follows we assume a Lambda cold dark matter (ΛCDM) cosmological model with the following parameters: $\Omega_\Lambda = 0.685$, $\Omega_m = 0.315$, $h = 0.674$, $\Omega_b = 0.05$ (Planck Collaboration VI 2020) so that the age of the Universe at the final redshift $z = 4$ is $t_H \sim 1.53$ Gyr.

2.1 Halo merger trees

GALFORM is a semi-analytic model of galaxy formation that reconstructs the hierarchical merger history of a given DM halo, also referred to as halo merger tree. We adopted the improved Monte Carlo algorithm developed by Parkinson et al. (2008), based on the Extended Press Schechter theory (EPS) and properly tuned to obtain an accurate agreement with N-body simulations, in particular with the results of the Millennium Simulation (MS, Springel et al. 2005). Here, we briefly summarize the merger trees reconstruction algorithm.

Starting from a target halo with a given mass at redshift z_0, GALFORM follows its evolution backward in time reconstructing its progenitors. The key point of this process is the conditional MF given by the EPS theory (Cole et al. 2000):

$$f(M_1|M_2)\,\mathrm{d}\ln M_1 = \sqrt{\frac{2}{\pi}} \frac{\sigma_1^2(\delta_1 - \delta_2)}{[\sigma_1^2 - \sigma_2^2]^{3/2}}$$
$$\times \exp\left[-\frac{1}{2}\frac{(\delta_1 - \delta_2)^2}{(\sigma_1^2 - \sigma_2^2)}\right]\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M_1}\right|\,\mathrm{d}\ln M_1,$$

(1)

where $f(M_1|M_2)$ represents the fraction of mass of haloes of mass M_2 at redshift z_2 that is contained in progenitor haloes of mass M_1 at an earlier redshift z_1. The values δ_1 and δ_2 are instead the linear density thresholds for collapse at these two redshifts and $\sigma(M)$ represents the rms linear density fluctuation in spheres containing a mass M, extrapolated to $z = 0$, with $\sigma_{1/2} \equiv \sigma(M_{1/2})$. Starting from equation (1) it is possible to obtain the mean number of haloes of mass M_1 into which a halo of mass M_2 splits after a step up in redshift dz_1, that is

$$\frac{\mathrm{d}N}{\mathrm{d}M_1} = \frac{1}{M_1}\frac{\mathrm{d}f}{\mathrm{d}z_1}\frac{M_2}{M_1}\mathrm{d}z_1 \qquad (M_1 < M_2).$$

(2)

Hence, the halo is decomposed into its progenitors and the process is repeated on each new halo at previous redshift steps up to a final value z_{max}, building up a complete tree. Although the above algorithm produces merger trees with statistical properties in good agreement with those obtained through detailed N-body simulations, it should be noted that the classical EPS theory systematically underestimates the mass of the most massive progenitor haloes with increasing redshift (Cole et al. 2008). For this reason a perturbing function $G(\sigma_1/\sigma_2,\ \delta_2/\sigma_2)$ was introduced, leading to a modification of equation (2) as follows:

$$\frac{\mathrm{d}N}{\mathrm{d}M_1} \to \frac{\mathrm{d}N}{\mathrm{d}M_1}G(\sigma_1/\sigma_2,\ \delta_2/\sigma_2),$$

(3)

where

$$G(\sigma_1/\sigma_2,\ \delta_2/\sigma_2) = G_o\left(\frac{\sigma_1}{\sigma_2}\right)^{\gamma_1}\left(\frac{\delta_2}{\sigma_2}\right)^{\gamma_2}.$$

(4)

G_0, γ_1 and γ_2 are free parameter calibrated to reproduce the MS conditional MF.

We used the GALFORM algorithm to simulate the formation histories of DM haloes with masses in the range $[10^9-10^{14}]\,\mathrm{M}_\odot$ at $z_{min} = 4$. This allows us to explore both the low-mass end of halo population and haloes as large as $M_{halo} \sim 10^{13}\,\mathrm{M}_\odot$ at $z \geq 6$, which are assumed to host the first SMBHs that power the observed highly luminous quasars (Fan et al. 2003; Valiante et al. 2011, 2016). We divide this mass interval into 11 logarithmically spaced bins with size 0.5. For each bin, we consider a final halo of mass equal to the central bin value and we use it as a starting point for the GALFORM code to simulate 10 independent halo merger trees.

Once the total merger tree sample has been generated, the resulting redshift-dependent mass distributions of each mass bin are weighted according to the number density of DM haloes at redshift $z = 4$, as given by the Sheth and Tormen MF (Sheth, Mo & Tormen 2001). In this way, the normalized sample is representative of the halo population at redshifts $z \geq 4$. It is important to stress that the limited number of merger tree realizations simulated for each final halo mass might, in principle, limit the allowed variability in the evolution of the corresponding galaxy properties. This is especially true for the lowest final halo masses. By doubling the number of merger tree simulations we find that the redshift evolution of mean galaxy properties does not change, showing that the current sampling scheme provides enough statistical variance even for the lowest halo mass bins.

2.1.1 The DM mass resolution

In the ΛCDM cosmological model, where larger structures form hierarchically through successive mergers of smaller ones, the first stars are expected to form inside the so-called minihaloes, i.e. small DM haloes with $M_{halo} \sim 10^5-10^6\,\mathrm{M}_\odot$, at redshift $z \sim 20-30$. Here, we classify as minihaloes systems with virial temperature in the range $1200\,\mathrm{K} \leq T_{vir} \leq 10^4\,\mathrm{K}$, where T_{vir} can be expressed as a function of
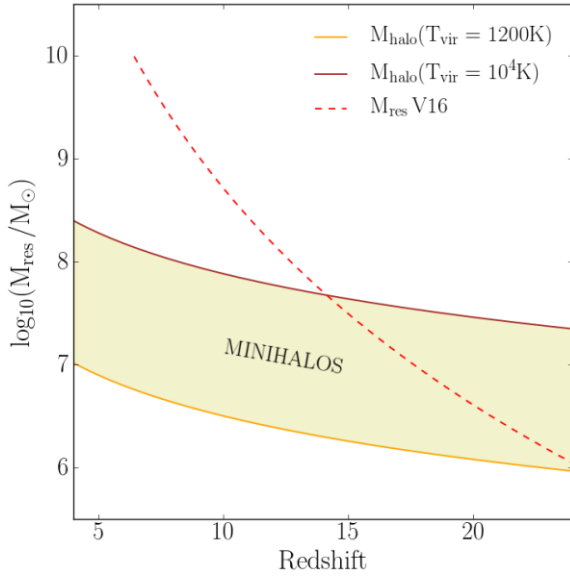
Figure 1. The mass resolution adopted when generating GALFORM merger trees (orange solid line) is compared to the merger tree mass resolution of Valiante et al. (2016, red dashed line), where a final halo mass of 10^{13} $M_\odot$ at $z = 6.4$ is assumed. For comparison, we also show the redshift-dependent minimum mass of atomic cooling haloes (brown solid line) so that the yellow shaded region illustrates the masses of DM *minihaloes* with 1200 K $\leq T_{vir} \leq 10^4$ K.

redshift z as (Bromm 2013):

$$T_{vir} \simeq 2 \times 10^3 \, \text{K} \left(\frac{M_{halo}}{10^6 \, M_\odot} \right)^{2/3} \left(\frac{1+z}{20} \right). \quad (5)$$

In these objects, where the gas temperature is below the threshold of 10^4 K for efficient cooling due to atomic hydrogen, the pristine gas can still cool down and fragment through the roto-vibrational emission of molecular hydrogen (H_2), giving birth to the first generation of Pop III stars. Due to their birth conditions, Pop III stars are expected to be massive and a large fraction of these will terminate their life as a stellar-mass BH, providing the first *light* BH seeds (Valiante et al. 2016). Hence, any model attempting to describe the build-up of SMBHs from growing BH seeds must be able to describe the formation of the first stars and BHs in high-z minihaloes. For this reason, while running GALFORM, we set a minimum DM halo mass of $M_{res} = M_{halo} (T_{vir} = 1200 \, \text{K})$. We assumed 1200 K as the minimum virial temperature for the onset of efficient H_2 cooling following the work of Haiman, Thoul & Loeb (1996).[6]

In Fig. 1, the mass resolution as a function of redshift adopted in CAT (solid orange line) is compared to the one adopted in GQD (red dashed line) for a final halo mass of 10^{13} $M_\odot$ at $z = 6.4$. The solid brown line shows the redshift-dependent minimum mass of *atomic-cooling haloes*, $M_{halo} (T_{vir} = 10^4 \, \text{K})$, so that the yellow shaded region represents the range of masses of the minihaloes population. It is clear that the increased resolution of CAT largely improves the statistics of minihaloes with respect to GQD and allows us to follow their evolution all the way down to $z = 4$. We obtain maximal improvements especially in the description of the most massive haloes since in

CAT the mass resolution at a given redshift is fixed, while in GQD it depends on the final halo mass.

2.1.2 Multiple mergers

In its original set-up, GALFORM merger trees are generated according to an adaptive number of steps that ensures their binarity (Cole et al. 2000). However, to follow the baryonic evolution, we re-grid all the merger trees, according to Valiante et al. (2016), on $N_t = 800$ time-steps logarithmically spaced in expansion factor between $z = 24$ and $z = 4$, so that the time interval between two simulation snapshots is ~ 0.5 Myr at $z \sim 20$, and ~ 4 Myr at $z \sim 4$. A consequence of this process is that the binary structure of the merger trees is no longer ensured since a given DM halo at a redshift step dz_i can be formed as a result of the merger of more than two haloes that were in place at a higher redshift dz_{i+1}, enabling the occurrence of what we call *multiple mergers*.

It is common to define as *major merger* between two DM haloes a merger event where the mass ratio μ between the least and most massive halo is higher than a chosen threshold value. Because of the loss of binarity, here we classify as *major mergers* interactions where the mass ratio between the first and the second most massive haloes among the merging ones is $\mu > 1/10$. This choice appears to be conservative, possibly leading to an underestimation of the number of major mergers, due to the lack of information about the order of mergers in multiple interactions occurring in a single time-step.

2.2 Baryonic evolution

In CAT, the evolution of the baryonic component is governed by the same physical prescriptions adopted in GQD, which we briefly summarize below. We will focus on BH evolution and on the new features introduced in this work, referring the reader to Valiante et al. (2014, 2016) for a more detailed description.

GQD was developed to investigate the formation history of the most extreme quasars observed at $z \geq 6$ and their host galaxies. To this aim, it follows the evolution of gas, stars, metals, and dust in each progenitor galaxy along a merger tree, tracking the process of star formation and the enrichment of the interstellar medium (ISM) due to asymptotic giant branch stars and Supernovae (SNe) by means of mass- and metallicity-dependent yields. The model follows a two-phase ISM environment where dust grains can both be destroyed by SN shocks expanding in the diffuse hot medium and can grow in mass by accreting gas-phase metals in warm dense gas (see de Bennassuti et al. 2014; Valiante et al. 2014, for details).[7] Mechanical feedback due to SN explosions and energy deposition associated with BH growth is also considered. The energy released by these processes couples with the gas, eventually removing a significant fraction of the galactic reservoir though energy-driven galactic scale winds (Valiante et al. 2012).

So far, GQD has been applied to study the formation of single $z > 6$ quasars, hosting $> 10^9$ $M_\odot$ SMBHs, in association with the evolution of their host galaxies (Valiante et al. 2011, 2014), BH seeds origin/properties (Valiante et al. 2016, 2018a, b; Sassano et al. 2021), earliest binary BHs formation (Valiante et al. 2020), and different gas accretion regimes (Pezzulli et al. 2016, 2017a, b). In this work, we

---

[6]Note, however, that in our model not all minihaloes are equally efficient at forming stars as the fraction of available cold gas depends on the strength of the illuminating UV background, the redshift, and the metallicity (see equation 6).

[7]The same physical prescriptions for metal and dust enrichment have been adopted in seminumerical models (Mancini et al. 2015, 2016) as well as in cosmological hydrodynamical simulations (Graziani et al. 2020) and provide good agreement with the observed dusty galaxy population at $z \geq 4$.

are instead interested in studying a population of galaxies and their nuclear BHs across a broad range of masses and populating a less biased region of the Universe down to $z = 4$.

2.2.1 Star formation and feedback

In the same spirit of GQD, CAT follows the evolution of the gas mass inside each galaxy during the whole DM halo assembly. Following halo virialization, the gas is accreted on to the newly collapsed halo and cools down to efficiently trigger star formation. The resulting fraction of gas mass is set by the balance between gas cooling and dynamical times. Inside each galaxy the star formation rate (SFR) is computed as

$$\text{SFR} = f_{\text{cool}} M_{\text{gas}} \epsilon_{\text{SF}} / \tau_{\text{dyn}}, \tag{6}$$

where M_{gas} is the available gas mass reservoir, ϵ_{SF} is the star formation efficiency per unit of time, and $\tau_{\text{dyn}} = [R_{\text{vir}}^3/(G\, M_{\text{halo}})]^{1/2}$ is the dynamical time of the system. The SF efficiency ϵ_{SF} represents a free parameter of the model and is calibrated as discussed in Section 3. f_{cool} quantifies the reduced cooling efficiency in mini-haloes, where it depends on the halo virial temperature, redshift, gas metallicity, and intensity of the illuminating LW radiation, as previously implemented by Valiante et al. (2016) and de Bennassuti et al. (2017). Conversely, in atomic cooling haloes we set $f_{\text{cool}} = 1$. As described in Valiante et al. (2016), we compute at each redshift z_{obs} the global LW cumulative background flux at the observed frequency ν_{obs} as

$$J(\nu_{\text{obs}}, z_{\text{obs}}) = \frac{(1 + z_{\text{obs}})^3}{4\pi} \int_{z_{\text{obs}}}^{z_{\text{max}}} \text{d}z\, c \left| \frac{\text{d}t}{\text{d}z} \right| \epsilon(\nu_z, z)\, \text{e}^{-\tau_{\text{H}_2}(\nu_{\text{obs}}, z_{\text{obs}}, z)}, \tag{7}$$

where $\epsilon(\nu_z, z)$ is the comoving emissivity in the LW band at redshift z, which is obtained summing over all the emitting sources, both stars, and accreting BHs. In equation (7), τ_{H_2} is the H_2 optical depth in the LW band (for a detailed description of its calculation see Valiante et al. 2016), while z_{max} is the maximum redshift from which an LW photon emitted by a source at $z > z_{\text{obs}}$ can reach the observer before being redshifted outside the LW band.

Finally, we account for the effects of photoheating feedback by suppressing star formation in haloes with virial temperatures below the temperature of the intergalactic medium (IGM), $T_{\text{vir}} < T_{\text{IGM}}$. We consider $T_{\text{IGM}} = Q_{\text{HII}} T_{\text{reio}} + (1 - Q_{\text{HII}}) T_{\text{HI}}$, where $T_{\text{reio}} = 2 \times 10^4$ K, $T_{\text{HI}} = 0.017(1 + z)^2$ and the filling factor of H II regions, Q_{HII}, is computed as in Valiante et al. (2016).

The abundance of gas inside each galaxy is also affected by mechanical feedback associated with SN explosions and BH accretion, whose released energy drives massive outflows of gas out of the galaxy. The total gas ejection rate $\dot{M}_{\text{ej}}$ is described as

$$\dot{M}_{\text{ej}} = \dot{M}_{\text{ej,SN}} + \dot{M}_{\text{ej,AGN}}, \tag{8}$$

where $\dot{M}_{\text{ej,SN}}$ and $\dot{M}_{\text{ej,AGN}}$ are the SN- and AGN-driven outflow rates, respectively, defined as

$$\dot{M}_{\text{ej,SN}} = \frac{2 E_{\text{SN}} \epsilon_{\text{w,SN}} R_{\text{SN}}(t)}{v_{\text{e}}^2}, \tag{9}$$

and

$$\dot{M}_{\text{ej,AGN}} = 2\, \epsilon_{\text{w,AGN}}\, \epsilon_r\, \dot{M}_{\text{accr}} \left(\frac{c}{v_{\text{e}}} \right)^2. \tag{10}$$

In equation (9) $R_{\text{SN}}(t)$ is the SN explosion rate, which depends on the SF history and on the nature of the stellar populations hosted by each

galaxy. For Pop III stars, we stochastically sample the initial mass function (IMF) in each SF episode and $R_{\text{SN}}(t)$ depends on the number of stars formed at each time in each galaxy. For Pop II stars, instead, we assume a fully sampled IMF and $R_{\text{SN}}(t) = 1.25x10^{-2}\, M_\odot^{-1}$ (see Section 2.3.1 for more details). In equation (9) E_{SN} represents the explosion energy per SN, assumed to be 2.7×10^{52} erg for Pop III stars and 1.2×10^{51} erg for Pop II stars.

The terms $\dot{M}_{\text{accr}}$ and ϵ_r, in equation (10), are instead the gas accretion rate and the AGN radiative efficiency, described in Section 2.3.4. In both expressions, $v_{\text{e}} = (2GM/R_{\text{vir}})^{1/2}$ is the escape velocity of the galaxy, and $\epsilon_{\text{w,SN}}$ and $\epsilon_{\text{w,AGN}}$ are free parameters representing the SN- and AGN-driven wind efficiencies, respectively. The adopted values are discussed in Section 3.

The ejection of metal-enriched gas and dust due to SN explosions and AGN activity enriches the IGM, increasing its metallicity and dust content. This leads to a corresponding increase in the initial gas metallicity and dust-to-gas mass ratio of DM haloes collapsing at later times, affecting their star formation history, as well as the formation of nuclear BHs, as will be outlined in the next section.

2.3 BH formation and evolution

SMBHs are supposed to grow via both gas accretion and mergers starting from their *seeds*, less massive progenitors whose nature has a crucial role in understanding the origin of SMBHs.

2.3.1 Light BH seeds

In CAT the mass distribution of *light BH seeds* depends on the Pop III stellar IMF, which is still highly uncertain (Bromm 2013). According to the most recent numerical simulations of metal-free star-forming regions hosted by minihaloes at $z \sim 20$–30, the Pop III mass distribution ranges from a few 10s to a few 100s $M_\odot$ (Greif et al. 2011; Hirano et al. 2014; Hosokawa et al. 2016; Stacy, Bromm & Lee 2016; Sugimura et al. 2020). Following Valiante et al. (2016), we assume that Pop III stars form according to a Larson IMF:

$$\Phi(m_*) \propto m_*^{\alpha-1}\, \text{e}^{-m_*/m_{\text{ch}}}, \tag{11}$$

where $\alpha = -1.35$, $m_{\text{ch}} = 20\, M_\odot$, and the possible range of stellar mass is $10\, M_\odot \leq m_* \leq 300\, M_\odot$. This choice is motivated by stellar archaeology studies and appears to best match the observed Galactic halo metallicity distribution function and the properties of C-enhanced and C-normal stars at [Fe/H] < -3 (de Bennassuti et al. 2014, 2017).

In our model, we stochastically sample the Pop III IMF untill we saturate the total stellar mass formed in each star formation episode. To consistently compute the BH remnants mass distribution, we assume that Pop III stars with masses in the range [40–140] $M_\odot$ and [260–300] $M_\odot$ collapse directly to BHs of comparable mass (Heger & Woosley 2002). Since these *light BH seeds* are expected to wander through the host galaxy, it is very unlikely that they will undergo mergers (Volonteri 2010), unless they form in binary systems (Sugimura et al. 2020). Moreover, dynamical effects such as three-body scattering or gravitational recoil following BH mergers could lead to the ejection of the merging objects from the host galaxy (Campanelli et al. 2007), especially inside smaller DM haloes (Dunn, Holley-Bockelmann & Bellovary 2020). For this reason, here we assume that only the most massive BH settles in the centre of the galaxy and is considered as its *light BH seed*.

Pop III star formation can be sustained until the gas metallicity of the star-forming region remains below a critical value Z $\leq$ Z$_{\text{cr}}$,

where we assume $Z_{cr} = 10^{-3.8} Z_{\odot}$ (Valiante et al. [2016](#)). Above this threshold value, metal-fine structure lines and dust cooling increase the cooling efficiency (Omukai [2001](#); Schneider et al. [2002](#), [2006](#), [2012](#); Omukai et al. [2005](#)), leading to a transition in the characteristic stellar masses. We therefore assume that above Z_{cr}, Pop II stars form in the mass range $0.1\,M_{\odot} \leq m_* \leq 100\,M_{\odot}$ according to a Larson IMF with $m_{ch} = 0.35\,M_{\odot}$ (de Bennassuti et al. [2014](#), [2017](#)).

2.3.2 Heavy BH seeds

The second viable scenario for BH seed formation implemented in CAT is the so-called Direct Collapse (DC) mechanism. Inside atomic-cooling haloes (where $T_{vir} \geq 10^4$ K), where metal and dust cooling is still inefficient ($Z \leq Z_{cr}$), if the abundance of molecular hydrogen is suppressed by LW photons (11.2–13.6 eV) inducing H_2 photodissociation, the gas collapses almost isothermally with no fragmentation, leading to the formation of a single supermassive star that becomes unstable, due to nuclear exhaustion or GR instabilities (Hosokawa et al. [2012](#); Inayoshi et al. [2014](#)), and forms a *heavy BH seed*, with mass in the range $[10^4–10^6]\,M_{\odot}$ (Latif et al. [2013](#); Ferrara et al. [2014](#); Becerra et al. [2015](#), [2018](#); Latif & Ferrara [2016](#)).

The importance of *heavy BH seeds* for the formation of high-redshift SMBHs strongly depends on their birth rate that is still subject to large uncertainties (Inayoshi et al. [2020](#)). If one neglects the effects of dynamical heating associated with structure formation (Wise et al. [2019](#)) or with major mergers (Mayer et al. [2010](#)), the abundance of *heavy BH seeds* in the family tree of SMBHs depends on the adopted value of Z_{cr} and on the critical value of the LW flux (J_{cr}) above which H_2 remains photodissociated. The latter condition is usually expressed as $J_{LW} \geq J_{cr}$, where J_{LW} is the cumulative flux into the LW energy band in units of 10^{-21} erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}. The value of J_{cr} is still very uncertain and depends on (i) the total spectral energy distribution of the radiation background created by the various sources (Agarwal & Khochfar [2015](#)), (ii) the efficiency of H_2 self-shielding, and (iii) the increase of the free electron fraction due to the presence of intense ionizing radiation which increases the H_2 formation rate (Inayoshi & Tanaka [2015](#)). As a result, values of J_{cr} ranging in a wide interval between ~30 and ~10^4 have been proposed in the literature (see Woods et al. [2019](#) and Inayoshi et al. [2020](#) and references therein). In addition, it has been recently suggested that the strong gas accretion rates may favour supermassive star formation event at higher metallicities than usually assumed (Chon & Omukai [2020](#)), through the so-called supercompetitive accretion scenario. This has been shown to increase the number of *heavy BH seeds* by a factor ranging from 2 (Sassano et al. [2021](#)) to 4 (Regan et al. [2020](#)).

Following Valiante et al. ([2016](#)), here we adopt as threshold values $Z_{cr} = 10^{-3.8} Z_{\odot}$ and $J_{cr} = 300$ to identify regions in atomic-cooling haloes where *heavy BH seeds* can form. If the conditions $Z < Z_{cr}$ and $J_{LW} \geq J_{cr}$ are satisfied, we set in the centre of the galaxy a *heavy BH seed* with a mass of $10^5\,M_{\odot}$.

2.3.3 BH mergers

Once formed, BH seeds are expected to grow via both gas accretion and coalescences with other BHs, eventually forming the SMBHs that power high-redshift AGNs (see Volonteri [2010](#); Johnson & Haardt [2016](#); Inayoshi, Visbal & Haiman [2020](#), for complete reviews).

Following Valiante et al. ([2011](#)), we assume that two BHs coalesce only during major halo–halo mergers, i.e. if the mass ratio of their interacting host DM haloes is $\mu > 1/10$ (as defined in Section 2.1.2;

Tanaka & Haiman [2009](#)). In our model, both the host galaxies and their nuclear BHs merge within the characteristic time interval of the simulation ($\Delta t \sim 0.5$–4 Myr) and the merger product settles in the nuclear region of the final galaxy. Conversely, in minor mergers ($\mu < 1/10$), only the most massive BH is assumed to migrate in the centre of the newly formed galaxy. The least massive one is instead considered as a satellite, wandering in the outskirts of the main galaxy (e.g. Callegari et al. [2009](#); Tamfal et al. [2018](#)), and its subsequent evolution is no longer followed within the model.

Although oversimplified, our assumption is based on the common expectation that Keplerian BH binaries form promptly in interacting galaxies of 'similar' mass and shrink to sub-pc separations (comparable to the primary BH influence radius) on relatively short time-scales (within about a million years in the most optimistic cases; e.g. Mayer et al. [2007](#); Tanaka & Haiman [2009](#)). We will return to this point in Section 5.

2.3.4 BH accretion

Nuclear BHs are assumed to grow by accreting gas from the surrounding medium. The growth is regulated by the processes of star formation and mechanical feedback, which both lead to a depletion of gas inside the host galaxy. Following the original GQD model, in our *reference* model we assume that nuclear BHs accrete gas according to the Bondi–Hoyle–Lyttleton (BHL) accretion rate (Hoyle & Lyttleton [1941](#); Bondi [1952](#)), given by

$$\dot{M}_{BHL} = \alpha \frac{4\pi G^2 M_{BH}^2 \rho_{gas}(r_A)}{c_s^3}. \tag{12}$$

In the above equation c_s is the sound speed, which is estimated assuming a gas temperature $T_{gas} = T_{vir}$, and $\rho_{gas}(r_A)$ is the gas density evaluated at the Bondi radius, i.e. the radius of gravitational influence of the BH, $r_A = 2GM_{BH}/c_s^2$. Following Valiante et al. ([2011](#)), the gas density distribution is approximated as a singular isothermal sphere with a flat core:

$$\rho_{gas}(r) = \frac{\rho_{norm}}{1 + (r/r_{core})^2}, \tag{13}$$

where $r_{core} = 0.012\,R_{vir}$ and ρ_{norm} represents a normalization constant that ensures that, at each time-step, the gas is distributed within the halo virial radius.

The parameter α in equation (12) does not appear in the original BHL formula and it is usually introduced in numerical simulation as a correction factor to take into account the enhanced gas density in the inner regions around the central BH. In fact, due to the lack of resolution of the simulations, the actual BHL accretion rate tends to be strongly underestimated (Di Matteo et al. [2012](#); Schaye et al. [2015](#)). As will be discussed in Section 3, the α parameter is a free parameter of the model.[8]

In our *reference* model, we assume that the BH accretion rate, $\dot{M}_{BH}$, cannot exceed the Eddington limit, so that

$$\dot{M}_{BH} = (1 - \epsilon_r)\,\dot{M}_{accr} = \min(\dot{M}_{BHL}, \dot{M}_{Edd}), \tag{14}$$

where

$$\dot{M}_{Edd} = \frac{L_{Edd}}{\epsilon_r\,c^2}, \tag{15}$$

---

[8]Note that since the Bondi radius is always much smaller than the core radius, the parameter α quantifies the unknown gas density enhancement around the black hole.

ϵ_r is the radiative efficiency, i.e. the efficiency at which the accreting gas is converted into radiated luminosity,

$$L_{Edd} = \frac{4\pi c G M_{BH} m_p}{\sigma_T} \qquad (16)$$

is the Eddington luminosity, c is the speed of light, and σ_T is the Thomson cross-section. Here, we assume that $\epsilon_r = 0.1$ (Shakura & Sunyaev 1973), unless otherwise specified. The bolometric luminosity of the accreting BH can be expressed as

$$L_{bol} = \epsilon_r \dot{M}_{accr} c^2. \qquad (17)$$

2.4 Model variants

As will be discussed in detail in Sections 4 and 5, our *reference* model, presented in the previous section, shows the best agreement with different independent theoretical and observational constraints in reproducing the BH mass and luminosity distribution at $z \gtrsim 4$. However, we also decided to explore two additional model variants, in order to understand how the evolution of the AGN population depends on the assumed accretion paradigm. In fact, these alternative models do not change the BH seeding prescriptions but consider different scenarios for their mass growth, as described below.

2.4.1 Exceeding the Eddington limit

In the CAT *reference* model BHs are assumed to experience spherical accretion according to the BHL rate, with a maximum allowed limit at the Eddington rate (see Section 2.3.4). The BH mass dependence of both rates may represent a strong limitation to early BH growth, particularly in the case of *light* BH seeds.

Provided that the gas reservoir can be efficiently replenished, through large-scale accretion or galaxy mergers, BH seeds at high redshift may quickly reach accretion rates with Eddington ratios $\eta_{Edd} \equiv \dot{M}_{BHL}/\dot{M}_{Edd} \to 1$. In order to check whether the restriction $\eta_{Edd} \leq 1$ provides a limitation to *light* BH seeds mass growth, we have explored an alternative model, which we dubbed *super-Edd*, where BHs are allowed to grow at super-Eddington rates (see Mayer 2019 and Inayoshi et al. 2020 for a thorough presentation of the main supercritical accretion models applied to BH seeds growth). Here, we adopt the optically thick, geometrically slim disc solution developed by Abramowicz et al. (1988), where part of the generated heat remains trapped within the accreting flow and is advected into the BH, leading to a low radiative efficiency. In this model variant, BHs accrete gas according to the BHL formula (see equation 12), so that

$$\dot{M}_{BH} = \dot{M}_{BHL}, \qquad (18)$$

but their bolometric luminosity is computed according to the fitting formula for the radiative efficiency proposed by Madau et al. (2014) which, in turn, is based on the numerical solution obtained by Sądowski (2009):

$$\frac{L_{bol}}{L_{Edd}} = A(a) \left[\frac{0.985}{\dot{M}_{Edd}/\dot{M}_{accr} + B(a)} + \frac{0.015}{\dot{M}_{Edd}/\dot{M}_{accr} + C(a)} \right], \qquad (19)$$

where $A(a)$, $B(a)$, and $C(a)$ are three functions of the BH spin parameter $a = a_{BH}$

$A(a) = (0.9663 - 0.9292a)^{-0.5639}$,

$B(a) = (4.627 - 4.445a)^{-0.5524}$,

$C(a) = (827.3 - 718.1a)^{-0.7060}$.

Note that in the above expression, the Eddington rate is defined as $\dot{M}_{Edd} \equiv 16 L_{Edd}/c^2$, i.e. it is a factor 1.6 larger than the definition

given by equation (15). Hence, in this model variant, the radiative efficiency is computed as

$$\epsilon_r = \frac{L_{bol}}{\dot{M}_{accr}c^2}. \qquad (20)$$

To allow a better comparison with the *reference* model, the spin parameter has been assumed to be $a_{BH} = 0.572$ for all the BHs. This ensures that in the limit $\dot{M}_{accr} << \dot{M}_{Edd}$ the radiative efficiency $\epsilon_r \to 0.1$, the same value adopted in the *reference* model.

With the above assumptions, large accretion rates up to $\dot{M}_{accr} \sim 100 \dot{M}_{Edd}$ lead to a radiated luminosity that remains only slightly super-Eddington, $L_{bol} \leq 5L_{Edd}$. Therefore, this enables super-Eddington accretion while – at the same time – the effects of AGN feedback is still limited, reducing the energy injected in the surrounding medium, and favouring BH growth.

2.4.2 Merger-driven BH accretion

As shown by Pezzulli et al. (2016, 2017b), super-Eddington accretion through a radiatively inefficient slim disc can be triggered by gas-rich galaxy mergers at high redshift. By implementing these prescriptions in GQD, it was found that – in highly biased regions, such as those that will later host bright quasars – episodic intense accretion is capable to grow $\sim 100 M_\odot$ *light* BH seeds to masses $\geq 10^4 M_\odot$, comparable to the mass of *heavy* BHs, by $z \sim 20$. Since these biased regions are growing their mass at a faster pace, it is interesting to investigate the impact of this scenario on the general galaxy and AGN populations at $z \geq 4$.

Here, we adopt a simplified description, which we refer to as *merger-driven* model, where we assume that, following a major merger, the gas inside the newborn galaxy is able to quickly loose angular momentum and the nuclear BH undergoes a period of enhanced accretion, with a rate given by

$$\dot{M}_{BH} = \frac{\epsilon_{BH} M_{gas}}{\tau_{accr}}, \qquad (21)$$

where M_{gas} is the gas mass inside the newly formed galaxy, ϵ_{BH} is the BH accretion efficiency, and τ_{accr} is the accretion time-scale. We assumed an efficiency $\epsilon_{BH} = 1/3 \, \epsilon_{SF}$ and a constant value of $\tau_{accr} = 10$ Myr, which are comparable to the maximum dynamical time-scale of the bulges found by Pezzulli et al. (2016) in their simulations. To be consistent with this merger-driven gas inflow scenario, the SFR in the newly formed galaxy is also enhanced by

$$SFR = \frac{\epsilon_{SF} M_{gas}}{\tau_{dyn,SF}}, \qquad (22)$$

where we assume a fixed star formation time-scale $\tau_{dyn,SF} = \tau_{accr}$. The enhancement in BH accretion and SF is assumed to terminate when either the ratio between the gas and BH masses, M_{gas}/M_{BH}, becomes lower than 10 or after a time interval $\Delta t_{burst} = \tau_{dyn}/100$. At that point, both star formation and BH accretion turn back to the quiescent mode, where the SFR is described by equation (6) and the BH accretion rate is described according to the BHL formula expressed by equation (12), but assuming $\alpha = 1$ (see Table 1 for the set of assumptions and parameters characterizing this *merger-driven* growth model).

3 MODEL CALIBRATION

As discussed in Section 2, our model presents four parameters which are tuned to regulate the evolution of the baryonic component: the star formation efficiency (ϵ_{SF}), the SN and AGN wind efficiencies ($\epsilon_{w,SN}$;

Table 1. Adopted set of parameters characterizing our CAT *reference* model and the *super-Edd* and *merger-driven* model variants (see Section 3): the star formation efficiency ϵ_{SF}, the SN and AGN wind efficiencies $\epsilon_{w, SN}$, $\epsilon_{w, AGN}$, and the BH accretion parameter α.

Model	ϵ_{SF}	$\epsilon_{w,SN}$	$\epsilon_{w,AGN}$	α	ϵ_r
reference	0.05	1.6×10^{-3}	2.5×10^{-3}	90	0.1
super-Edd	0.05	1.6×10^{-3}	2.5×10^{-3}	40	Equation (20)
merger-driven	0.05	1.6×10^{-3}	2.5×10^{-3}	1	Equation (20)

$\epsilon_{w, AGN}$), and the BH accretion parameter (α). In Valiante et al. ([2011](#), [2016](#)), these free parameters were tuned to reproduce the SMBH mass and the properties of the host galaxy of SDSSJ1148+5251 at $z = 6.4$ (Fan et al. [2002](#)), a well-known source, assumed to represent a prototypical example of luminous quasars at $z > 6$. In this work, we target the observed properties of high-redshift quasars, as well as a realistic cosmic star formation history down to $z \geq 4$. Here, we discuss the main observables used to calibrate the reference model, checking the consistency between predictions and observations. Although not reported, the same calibration has been performed for the model variants described in Section 2.4, see Table [1](#) for a summary of the adopted parameters.

3.1 Cosmic star formation and stellar mass density

We first tuned the model predictions on the observed total SFR density (SFRD), i.e. the total stellar mass formed per unit time and comoving volume. Measurements of the cosmic SFRD are mainly inferred from galaxy UV luminosity, currently limited to bright sources (UV magnitudes $M_{UV} \leq -17.7$), which are also affected by dust extinction. To compare the model predictions with observations, we therefore consider the contribution from sources with UV luminosity above the observed threshold ($M_{UV} \leq -17.7$).[9] To apply this luminosity cut, we convert the SFR of each galaxy to an intrinsic UV luminosity following the relation (Madau & Dickinson [2014](#)):

$$L_{UV} = \left(\frac{\text{SFR}}{M_\odot \, \text{yr}^{-1}} \right) 7.14 \times 10^{27} \, \text{erg s}^{-1} \, \text{Hz}^{-1} \quad (23)$$

Fig. [2](#) shows that the *reference* model (dashed red line) is in good agreement with the observed SFR density and stellar mass density evolution, from $z = 4$ out to $z = 8$. Note that CAT predicts a large number of currently undetected sources ($M_{UV} > -17.7$), which dominate the SFR density and stellar mass density at $z > 4$ (solid red lines). The impact of dust extinction on the galaxy luminosity might increase even further the contribution coming from undetected sources, particularly at high redshift, where faint galaxies account for the majority of the total SFRD.

3.2 Mass and luminosity of quasars at $z > 5$

In order to check whether our model, once calibrated, properly reproduces the properties of the observed population of high-redshift quasars, here we compare the mass and bolometric luminosity of the most massive systems predicted by CAT at $z > 5$ with the values inferred from quasar observations at similar redshifts. The results of the *reference* model at $z = 5$, 6, and 7 are shown, respectively, as yellow, orange, and dark red points in Fig. [3](#). The empty black data points show a sample including all $z > 5.8$ quasars

for which BH masses have been derived via Mg II line single epoch virial estimator.[10] The diagonal dotted lines indicate the Eddington luminosity as a function of the BH mass.

We find that our simulated sample is composed mainly of low-mass objects, $\lesssim 10^{7.5} \, M_\odot$, shining close to $0.1 \, L_{Edd}$ regardless of their redshift. More massive BHs show instead higher Eddington ratios, especially at $z = 6$ and 7, where a large amount of gas is available for accretion. The comparison with observational data at $5.8 < z < 7.5$ (empty squares) shows that, despite our limited statistics due to the low number density of such objects, the most massive BHs predicted by CAT at similar redshifts populate the observed range of quasar masses and luminosities.

In conclusion, Figs [2](#) and [3](#) show that the parameter values assumed in our *reference* model (see Table [1](#)) guarantee a good agreement with the observed quasar population at $z > 5$ and lead, at same time, to a galaxy population characterized by an evolution of the SFR density and stellar mass density in agreement with the available data.

Finally, a similar analysis has been carried out also for the two model variants, and the corresponding model parameters are reported in Table [1](#).

4 RESULTS

The main motivation of our study is to investigate if, when, and how the nature of BH seeds and their mass growth affect the population of BHs at $z > 4$. In this section, we first analyse the redshift evolution of the BH MF predicted by the *reference* model and discuss how it changes in the two model variants. Secondly, we investigate whether the differences across models are revealed in the predicted LFs, exploring at the same time their accordance with observational constraints obtained by current and future facilities.

4.1 BH mass function

In Fig. [4](#), we show the BH MF predicted by the *reference* model at different redshifts, ranging from $z = 18$ down to $z = 4$. Mass bins populated by *light* and *heavy* (i.e. at least with one DCBH progenitor) *seed* descendants are shown in magenta and red points, respectively. Here, as in the rest of the paper, binned data points are shown with the corresponding 1σ Poisson error bars. The mass bin which includes newly formed and/or not grown heavy BH seeds has been highlighted in orange and it is not considered when fitting the distributions. In fact, the high number density of $10^5 \, M_\odot$ BHs reflects the adopted seeding prescription, while a more realistic heavy seed birth MF (e.g. Ferrara et al. [2014](#)) and/or the inclusion of the so-called medium-weight seed formation channel (Sassano et al. [2021](#)) would probably result in a lower number density, spread across the intermediate-mass range (10^3–$10^5 \, M_\odot$).

It is immediately evident that the distribution is characterized by two well-defined regions: BH descendants of *light* BH *seeds* are

---

[9]The same correction is adopted to compare the cosmic stellar mass density predicted by the model with observations.

[10]We adopted the virial BH mass relation involving the Mg II emission line full width at half-maximum and continuum luminosity at 3000 Å, using the calibration from Shen & Liu ([2012](#)). The compilation of sources considered here has been assembled collecting data from the following works: Willott et al. ([2010a](#)); De Rosa et al. ([2011](#)); Wu et al. ([2015](#)); Mazzucchelli et al. ([2017](#)); Reed et al. ([2017](#)); Shao et al. ([2017](#)); Chehade et al. ([2018](#)); Eilers, Davies & Hennawi ([2018](#)); Eilers et al. ([2020](#)); Bañados et al. ([2018](#)); Wang et al. ([2018](#), [2020](#)); Shen et al. ([2019](#)); Reed et al. ([2019](#)); Onoue et al. ([2019](#)); Pons et al. ([2019](#)); Andika et al. ([2020](#)); Yang et al. ([2020](#)).
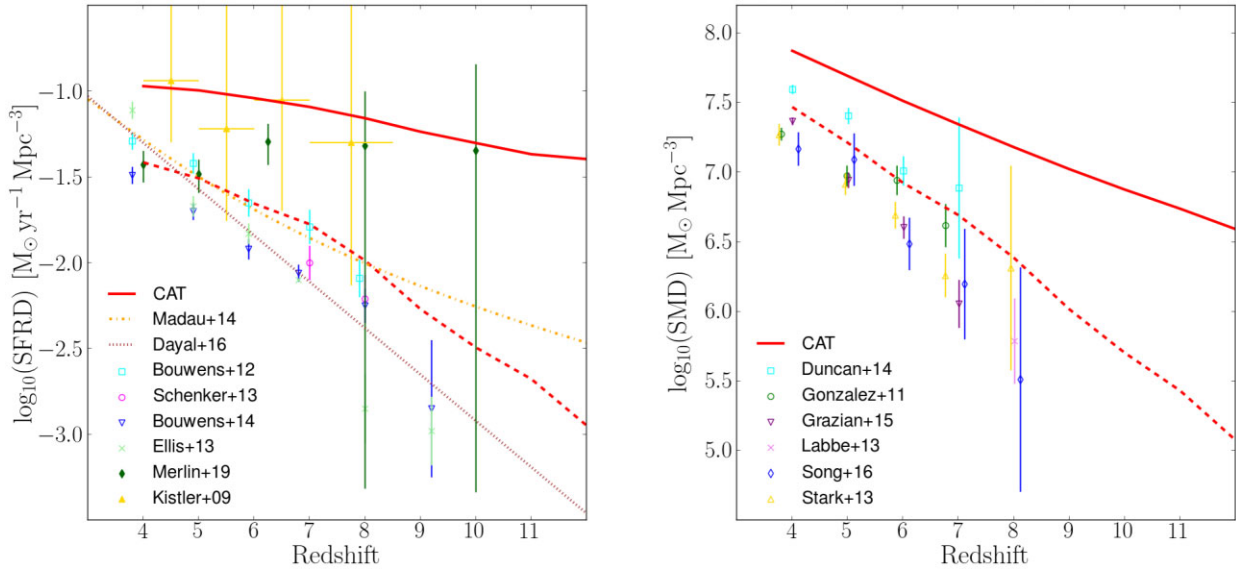
Figure 2. Redshift evolution of the comoving SFRD (left-hand panel) and stellar mass density (right-hand panel) predicted by CAT, when the entire galaxy population is considered (red solid lines) and when only galaxies with $M_{\rm UV} < -17.7$ are accounted for (red dashed lines). In each panel, the model predictions are compared with different observational results, as indicated in the legends Kistler et al. (2009), González et al. (2011), Bouwens et al. (2012), Labbé et al. (2013), Stark et al. (2013), Schenker et al. (2013), Ellis et al. (2013), Duncan et al. (2014), Bouwens et al. (2014), Oesch et al. (2014), Grazian et al. (2015), Song et al. (2016), and Merlin et al. (2019). In the right-hand panel, the yellow dash–dotted line shows the empirical SFRD by Madau & Dickinson (2014) and the brown dotted line shows the model prediction by Dayal, Ward & Cockell (2016).
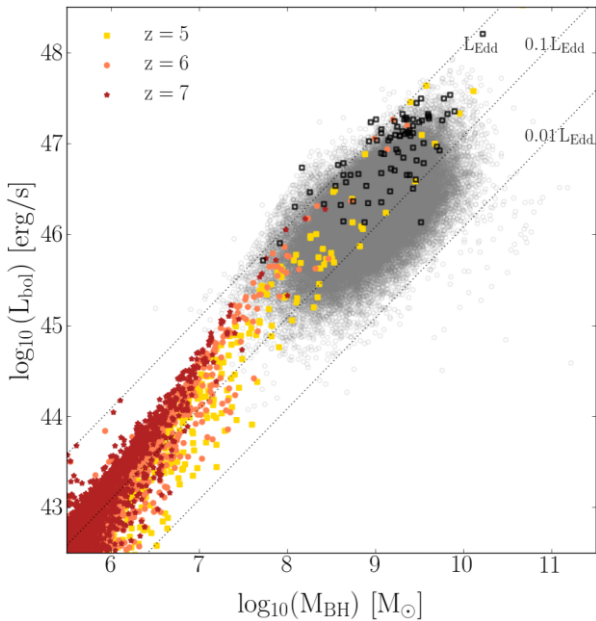
Figure 3. Bolometric luminosity as a function of the BH mass for high-redshift AGNs. The coloured data points represent the quasar sample of the CAT reference model at $z = 5$ (yellow), 6 (orange), and 7 (red). The black empty squares represent a collection of observed quasars at redshift $5.8 < z < 7.5$, while the grey data show a large sample of quasars in the range $0.6 < z < 2$ drawn from the SDSS-DR7 quasar catalogue by Shen et al. (2011) and for which Mg II-based BH masses have been derived. The dotted lines mark the position of BHs with $L_{\rm bol} = L_{\rm Edd}$, $0.1L_{\rm Edd}$, and $0.01L_{\rm Edd}$.

confined below $10^5\,{\rm M_\odot}$, while BHs grown from *heavy* BH *seeds* dominate the high-mass end of the MF, i.e. above $10^5\,{\rm M_\odot}$. These two populations appear well separated across all the redshifts explored by our model. Early on ($z \geq 18$) only *light* BH seeds are formed, some

of which have already grown up to $\sim 10^{3.5}\,{\rm M_\odot}$. At $14 \leq z \leq 17$ the first *heavy* BH seeds start to form, but they do not significantly grow above their formation mass scale ($10^5\,{\rm M_\odot}$). At these same redshifts, the distribution of *light* BH seed descendants has already reached its maximum extension in mass, well below $10^4\,{\rm M_\odot}$. At $z < 14$ *heavy* BH seeds descendants continue to grow, and since their formation becomes progressively rarer, their mass distribution shifts to larger masses, covering the mass range 10^6–$10^{10}\,{\rm M_\odot}$ at $z \lesssim 6$. Hence, the MFs of *light* and *heavy* BH descendants remain completely segregated in mass as a consequence of the inefficient growth of *light* BH seeds. Some of these may be involved in major mergers contributing to the growth of the *heavy* seed descendants, others may instead participate in minor mergers, eventually becoming satellites (or wandering) BHs, which, in turn, explains why the highest mass bin of the distribution shifts to lower values with increasing time (decreasing redshift).

The gap in the BH number density just below the minimum adopted mass of *heavy* BH seeds indicates thus that in the CAT *reference* model *light* BH seeds fail to grow efficiently all the way down to $z \sim 4$. This could be a consequence of the host environmental properties of this class of seeds, namely the abundance of gas available to fuel BH growth, or the adopted mass accretion model. For this reason, we explored the redshift evolution of the BH MFs predicted by two model variants: the *super-Edd* model and the *merger-driven* model (see Section 2.4 and Table 1 for more details).

Fig. 5 illustrates the results for the CAT *super-Edd* model. To aid the comparison with the *reference* model, the magenta and red dashed lines indicate the best fit to the MFs of *light* and *heavy* BH seeds predicted by the *reference* model (as in Fig. 4). We find a significantly faster growth of *heavy* BH descendants, especially in the first phase after their formation, at redshifts $16 \lesssim z \lesssim 10$. At lower redshift this accelerated growth seems to slow down quickly and, despite the early build-up of the global MF, the maximum BH masses reached are ~ 1 dex below the results of the reference model. The rapid consumption of gas and the associated feedback in the
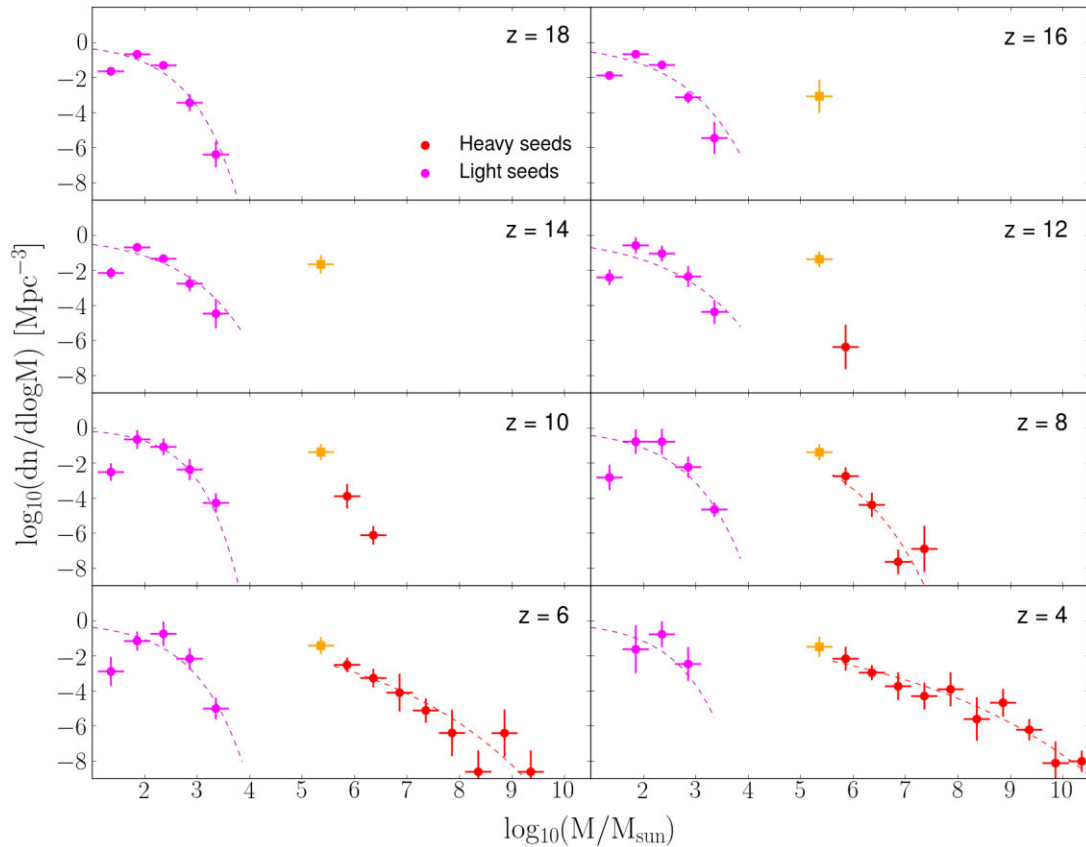
Figure 4. The BH MF predicted by the *reference* model at different redshifts, ranging from $z = 18$ down to $z = 4$. In each panel, we show the separate contributions of BHs descendants of *light* and *heavy seeds* with magenta and red points, respectively. The 1σ Poisson error bars are shown in each data point. Where possible, the best fits of the two distributions are also shown with dashed lines. The orange data point represents the mass bin populated by newly formed and/or not grown heavy BHs seeds and it is not considered to produce best-fitting curves since its occupation is closely related to the adopted seed birth mass.

early BH growth phase might be the reason for this behaviour. In fact, super-Eddington accretion on to massive seeds leads to a quick depletion of gas from the host galaxy, which affects the subsequent BH evolution at later times. Hence, this *super-Edd* model well reproduces the global star formation history but struggles in reproducing the formation of the billion solar mass BHs observed to power quasars at $z > 6$.

However, the *super-Edd* results show the same characteristic feature of the *reference* model observed in Fig. 4, i.e. a persisting gap between the *light* and *heavy* BH seeds distribution. This implies that the Eddington limit imposed to the BH accretion does not represent the leading cause of the stunted growth of *light* BH seeds descendants. Instead, it could be due to a limited content of gas to fuel BH growth inside their hosts, or it could be a consequence of the adopted BHL accretion rate. In order to discern between these two possibilities, we investigate the predictions of the *merger-driven* model, shown in Fig. 6. In this figure, we indicate again with magenta data points BH mass bins contributed only by *light* BH seeds descendants. However, for BH masses $\geq 10^5$ M$_\odot$, the mass bins are now populated also by BHs grown from *light* BH seeds, in addition to those coming from *heavy* BH seeds progenitors and we show this mixed population in violet. In the previous *reference* and *super-Edd* models, instead, the high-mass end of the BH MF is entirely populated by BHs with at least one *heavy* BH seed progenitor. As a comparison, we also show with magenta and red dashed lines the best fit to the MFs of *light* and *heavy* BH seeds in the *reference* model, respectively.

The *merger-driven* model shows a very rapid and early growth of *light* seeds, which reach BH masses as high as 10^6 M$_\odot$ already at $z \sim 18$. The resulting distribution evolves shifting towards higher mass values, with minor changes in the global shape. This leads to a BH MF at $z \lesssim 6$ that continuously ranges between $\sim 10^4$ and 10^{10} M$_\odot$. At the low-mass end, the original gap has now disappeared since *light* BH seeds and their descendants can efficiently grow, competing with *heavy* seeds in SMBHs growth. The high-mass end of the distribution, instead, reaches values consistent with the reference model, predicting however a slightly higher number density of objects in the range $\sim 10^6$–10^8 M$_\odot$.

It is important to note that the peak in the BH number density is shifted towards higher masses at later times (lower redshifts). This reflects the progressive depletion of low-mass objects as BH seeds are able to efficiently increase their masses through accretion and mergers, while newly formed seeds become rarer. The decreased number density of low-mass objects is clearly noticeable at $z \lesssim 6$.

In order to emphasize the contribution to the MF of BHs grown only from *light* BH seeds in the *merger-driven* model, in Fig. 7 we show the MF of *heavy* seed descendants at $z = 4$, 6, 8, and 10, along with their percentage of occupation of each BH mass bin. It is clear that, in this model variant, BHs descending only from *light* seeds represent a large fraction of the BH population. In particular, besides largely dominating the entire MF at $z \gtrsim 10$, BHs grown from *light* seeds represent ~ 50 per cent of the SMBH population ($M_{\rm BH} > 10^5$ M$_\odot$) down to redshift $z \sim 6$.
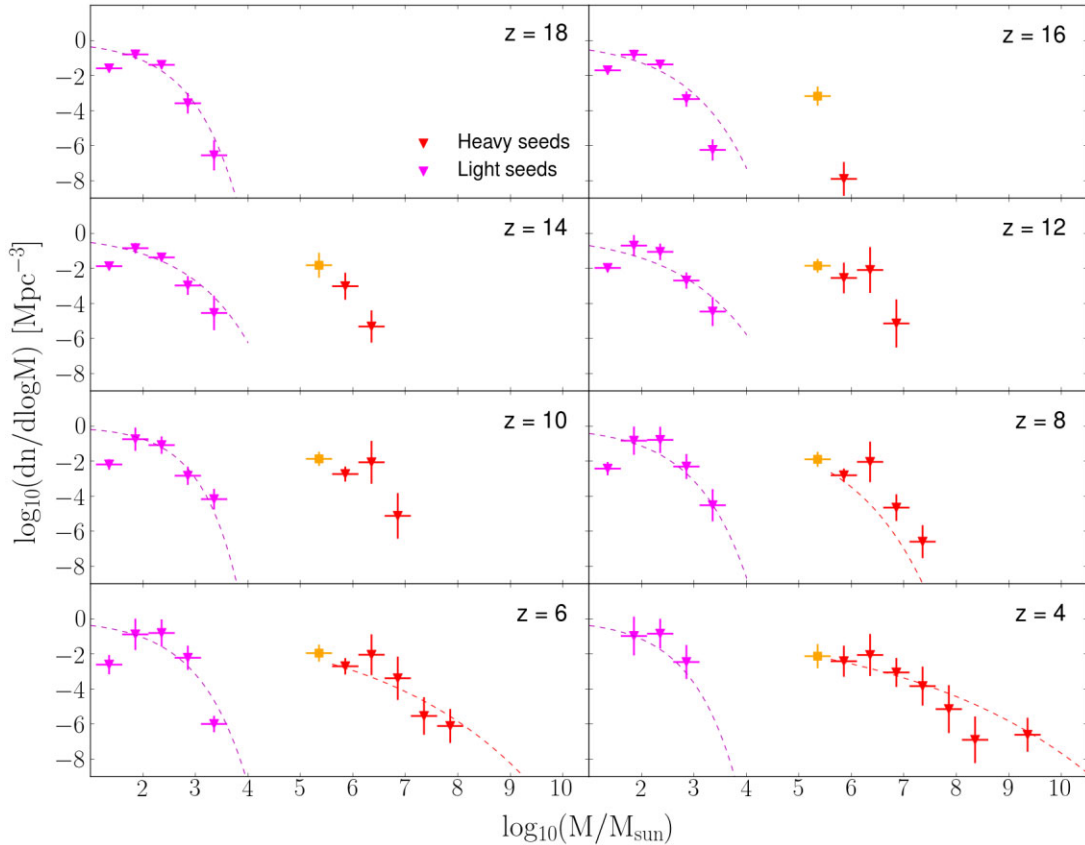
Figure 5. Same as Fig. 4 but for the *super-Edd* model (see Section 2.4.1). As a reference, we show with magenta and red dashed lines the best fits of the *light* and *heavy* BH seed descendant distributions obtained for the *reference* model.

Hence, these results lead us to the conclusion that the main limiting factor to the growth of a BH seed is the mode of gas accretion rather than the gas mass reservoir around the BH. The successful growth of *light* seeds predicted by the *merger-driven* scenario suggests that the key element affecting the efficiency of BH growth is the way in which the gas reaches the central regions, and therefore the physical mechanisms driving BH mass accretion.

However, we found that the large-scale mechanism of merger-driven infall of gas towards the central regions around the nuclear BH is not sufficient, alone, to ensure an efficient growth of the smaller seeds. A BH accretion model that allows super-Eddington growth, as the *slim disc* model assumed in the *merger-driven* variant, is also needed to fill up the gap in the MF. In fact, assuming an Eddington-limited growth in the merger-driven scenario, $\dot{M}_{BH} \leq \dot{M}_{Edd}$, we open-up again a clear gap in the mass distribution below 10^5 M$_\odot$, similarly to that shown in Figs 5 and 6. In addition, such a model predicts an overall reduction of the BH growth due to the low BH duty cycle assumed in the merger-driven scenario.

This suggests that the low-mass end of the BH MF at $z \leq 6$ may provide important indications on the mass scale and growth rate of BH seeds at high redshift, while, at the high-mass end, SMBHs with masses $\geq 10^6$ M$_\odot$ are relatively insensitive to the nature of their BH progenitors. Still, the depopulation of the low-mass region (below $\sim 10^5$ M$_\odot$) predicted by the *merger-driven* model at $z \lesssim 6$ might resemble the clear gap in the MF predicted by a Bondi-like accretion mechanism, worsening the chances to clearly identify the distinctive features of different evolutionary scenarios.

The question is then whether current or future surveys may be able to discriminate among the various models, as it will be discussed in the next section.

4.2 The AGN LFs

CAT model allows us to estimate the accretion rate of each nuclear BH during its evolution through cosmic time, as determined by the environmental conditions present in its host galaxy. Hence, we are able to infer the luminosity of all active BHs (see Sections 2.3.4, 2.4.1, and 2.4.2) and to reconstruct their luminosity distribution, i.e. the AGN luminosity function. Here, we compare the bolometric LF predicted by CAT with several independent estimates based on multiband observational data. These comparisons represent a fundamental benchmark for our BH evolution model, in order to determine which of the considered accretion scenarios better reproduce the observed trends. At the same time, they also allow us to check whether current or future surveys may be able to detect signatures of BH seeds populations and of their growth mode. A comparison between CAT predictions and the findings of the most recent AGN UV and X-ray surveys is discussed in Section 4.2.2.

4.2.1 Bolometric LF

In Fig. 8, we show the bolometric LF predicted by all models in Table 1 at $z = 4$, 5, and 6. In the *reference* and *super-Edd* models a clear gap in the LF is visible below a threshold luminosity of
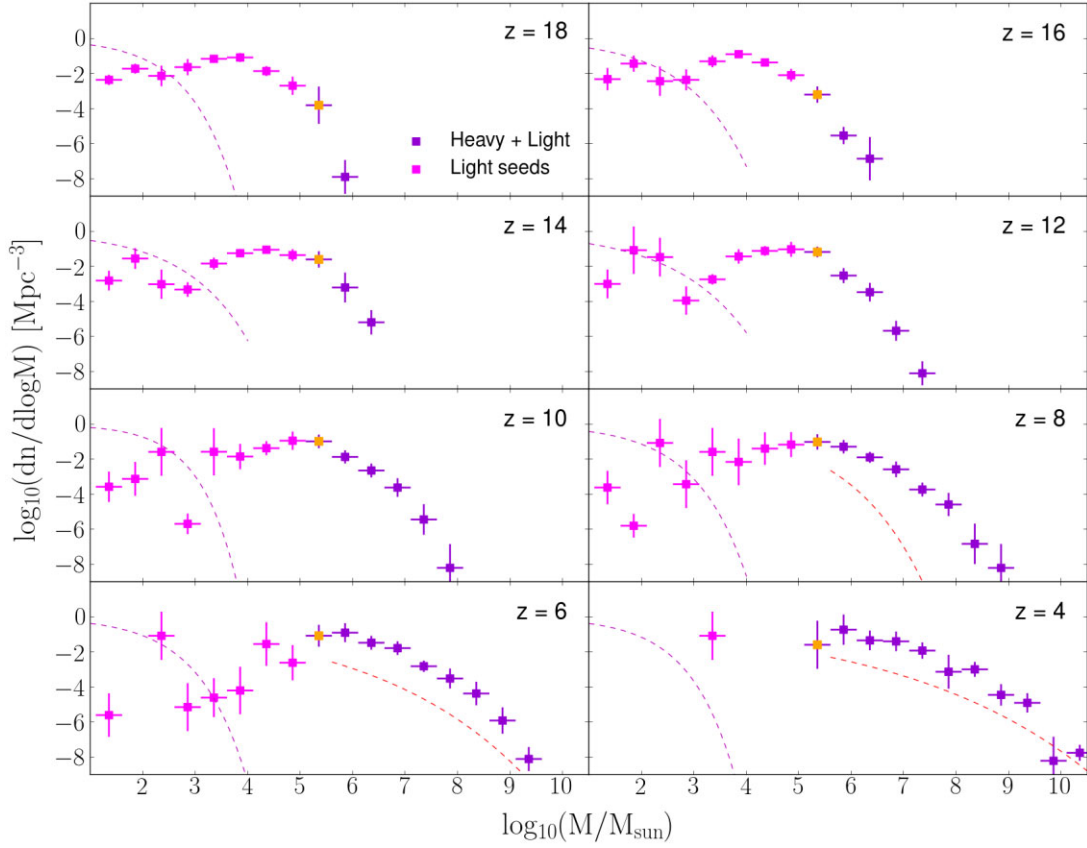
Figure 6. Same as Fig. 5 but for the *merger-driven* accretion model (see Section 2.4.2). Here, we show with violet data points the mass bins above $10^5\,M_\odot$ since, unlike the previous models, they are now populated by both *light* and *heavy* BH seed descendants. Distributions (data points with error bars) are compared with best fits obtained for the *reference* model (dashed lines).
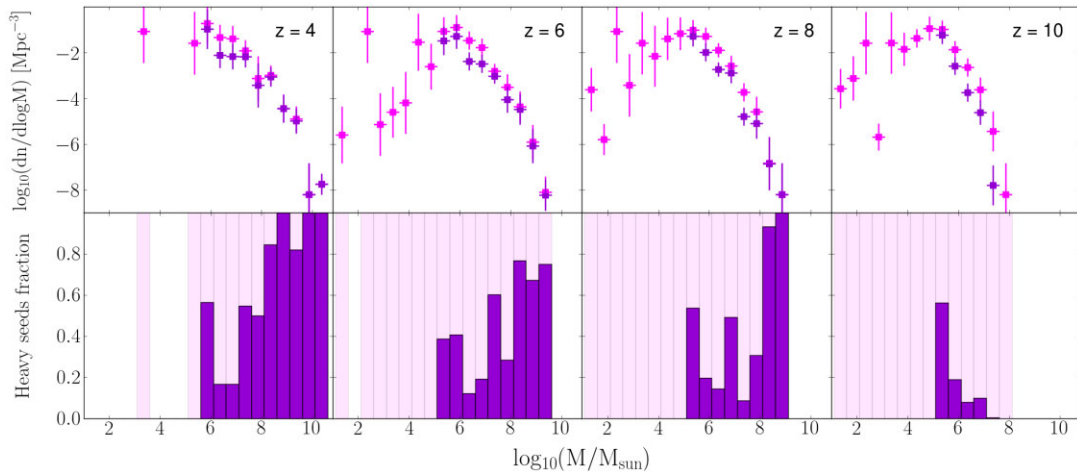

Figure 7. The figure shows, for the *merger-driven* model, the relative contribution to the final BH MF of BHs descending from at least one *heavy* or only *light* BH seeds. Top row: the total BH MF (magenta points, same as in Fig. 6) and the contribution of only *heavy*-BH seeds descendants (violet) at $z = 4, 6, 8$, and 10 (from left to right). Bottom row: the corresponding percentage of *heavy* BH seed descendants in different mass bins. BHs descending only from *light*-BH seeds (represented by lighter magenta regions of the histograms) clearly provide, in this model, a significant contribution to the global MF, even at the high-mass end.

$\sim 10^{41}\,\mathrm{erg\,s^{-1}}$ at all redshifts. This reflects the behaviour of the BH MFs shown in Figs 4 and 5, and it is a consequence of the failed growth of *light* BH seeds that characterize these two models, as previously discussed. While *heavy* seeds descendants continue to grow, populating the bright-end of the LF, the empty region in the luminosity distribution widens with decreasing redshift. In fact, as already described in Section 4.1, *light* seeds descendants are progressively more involved in galaxy and BH merger events, eventually becoming satellite BHs or merging with *heavy* descendants.
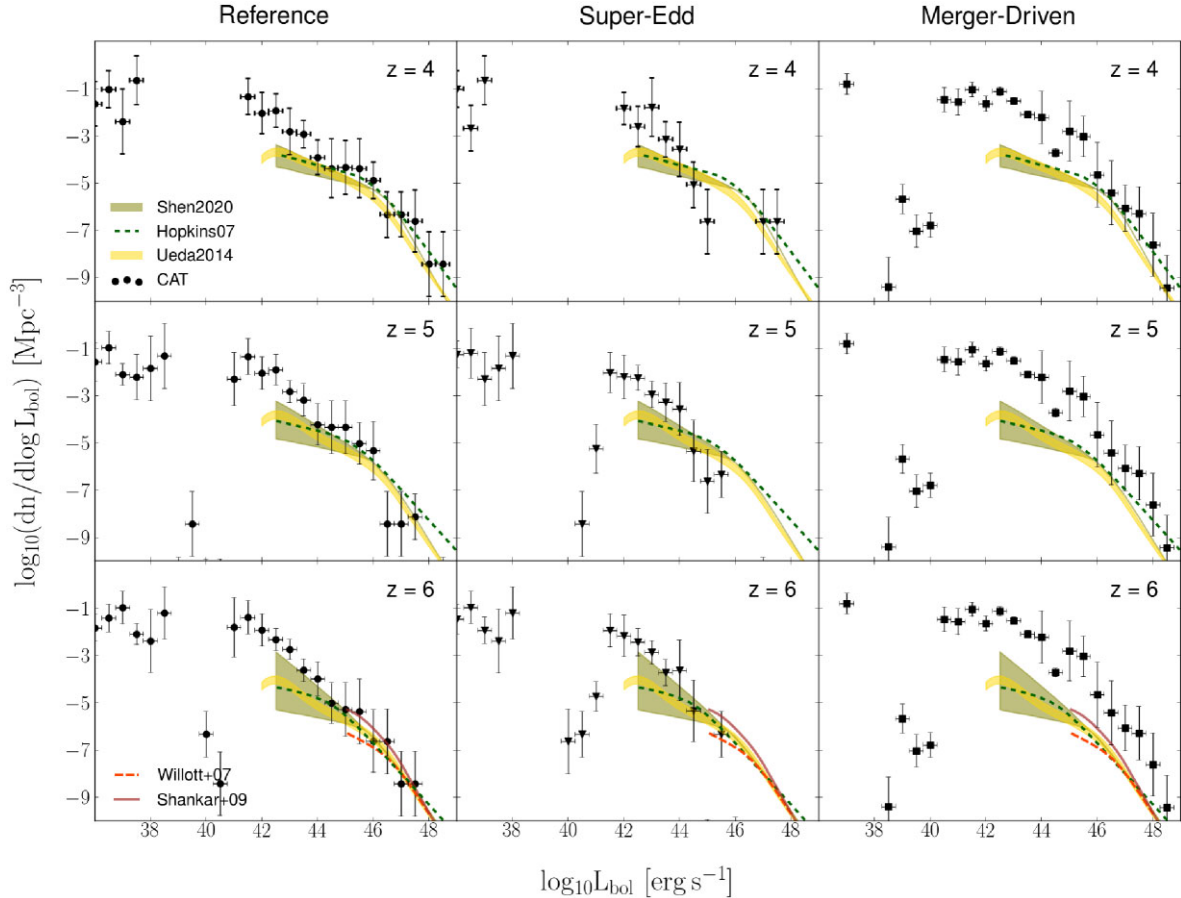
Figure 8. Bolometric LF of AGNs at $z = 4$, 5, and 6 (from top to bottom). The black data points show the luminosity distribution of CAT AGN population. Each column shows the results obtained in one of the three models examined: the *reference model* (left-hand panels), the *super-Edd* model (central panels), and the *merger-driven* model (right-hand panels). Our results are compared with the predictions of Hopkins, Richards & Hernquist (2007, green dashed lines), Willott et al. (2010b, red dashed line), Shankar, Weinberg & Miralda-Escudé (2009, brown solid line), and Shen et al. (2020, green shaded area). We also show as a comparison the constraints on the bolometric LF derived from the X-ray data by Ueda et al. (2014) (yellow shaded area, see the text).

In the *merger-driven* model instead the efficient growth of light BH seeds leads to a continuous luminosity distribution at all redshifts. However, at $z \sim 6$ the mass distribution shown in Fig. 6 results in an LF which peaks around $\sim 10^{40-42}\,\mathrm{erg\,s^{-1}}$ and then quickly drops down at fainter luminosities. This behaviour, as mentioned before, reflects the progressively lower number of lighter seeds, due to both the involvement in galaxy mergers or their efficient growth towards higher masses.

Note that this third model is characterized by brief periods of enhanced accretion on to BHs. If we infer the LF at a single given time, we might undersample rare objects with high luminosity that are active only for a short interval of time. Therefore, following Griffin et al. (2019), when computing the LF for the *merger-driven* model, we average the LF over a time window Δt_w of $\sim 50\,\mathrm{Myr}$ around each redshift of interest. To reconstruct the global luminosity distribution, each active object is then assigned a weight

$$w = t_Q / \Delta t_w, \tag{24}$$

where t_Q is the time spent in the enhanced accretion mode during the considered time window.

The model predictions are compared to the results obtained by Willott et al. (2010b), based on SDSS and CFHQS data, and with the LF evolutionary models proposed by Shankar et al. (2009) and

Hopkins et al. (2007). We also show the LF constraints presented by Shen et al. (2020), which are based on a large compilation of AGNs observations in different energy bands, from the rest-frame IR to the X-rays. Finally, we compare our findings with the results of the X-ray studies by Ueda et al. (2014), applying the bolometric correction proposed by Duras et al. (2020). The shaded region shows the range of uncertainties associated with obscuration effects, with the upper and lower bounds illustrating the LF obtained with and without obscuration correction, respectively.[11]

The LFs predicted by the *reference* model are remarkably consistent with the observed bright-end distributions ($\gtrsim 10^{42}\,\mathrm{erg\,s^{-1}}$) at all the considered redshift, $z = 4$–6. The number of sources is instead slightly overpredicted at fainter luminosities, $\sim 10^{40-42}\,\mathrm{erg\,s^{-1}}$, where, however, the observations should be largely affected by obscuration effects.

---

[11] Here, the obscuration correction has been applied assuming a fraction of Compton-thick AGNs with column density $24 < N_H < 26$ four times higher that the one considered in Ueda et al. (2014). This is in agreement with the recent work of Ananna et al. (2019), which found a much higher number density of Compton-thick objects with respect to the original work of Ueda et al. (2014).

The distribution of low-luminosity sources predicted by the *super-Edd* model (middle panels of Fig. 8) is similar to that described in the *reference* case. Conversely, the number density of brighter sources, above $\sim 10^{44}\,\mathrm{erg\,s^{-1}}$, dramatically drops, especially at $z = 5$ and 6, as a consequence of the lack of massive/luminous AGNs predicted in the model.

This result suggests that the accelerated growth of BH seeds at early times, driven by the super-Eddington accretion (see Fig. 5), strongly affects the BH evolution at later times. In particular, the enhanced accretion rate and the associated feedback significantly deplete the gas reservoir available around the BH. On the other hand, new infalling material from the external medium is mainly consumed by star formation that is quite efficient in massive haloes hosting growing heavy seeds. Hence, the growth in mass of the initial seed remains shortly stuck and it will undergo major accretion episodes only as a consequence of gas-rich mergers, failing to reproduce both masses and luminosities of the extreme quasars observed at $z \sim 6$.

In the *merger-driven* model (right-hand panels of Fig. 8) the bright-end of the LF ($L_{bol} \gtrsim 10^{46}\,\mathrm{erg\,s^{-1}}$) is consistent with the observational data at $4 \leq z \leq 5$. At $z = 6$ the LF is instead overestimated as a consequence of the early mass growth of BHs driven by halo mergers. At the faint-end of the distribution, the *merger-driven* model overpredicts the number of objects at all redshift within the $10^{41}\,\mathrm{erg\,s^{-1}} \lesssim L_{bol} \lesssim 10^{45}\,\mathrm{erg\,s^{-1}}$ luminosity range. Compared to the *reference* and *super-Edd* models, this is due to the population of efficiently grown *light* seed descendants that were unable to significantly grow in mass in the classic Bondi-like accretion scenario. At even fainter luminosities, below $\sim 10^{40}\,\mathrm{erg\,s^{-1}}$, the LF drops down similarly to what we observe in the *reference* model, but as a result of the exact opposite process. In fact, while in the BHL accretion scenario the lack of sources below this threshold is due to the low mass of *light* seed descendants, here most of them grows efficiently, depopulating this luminosity region at low redshift.

It is worth considering that if the bursts of accretion driving BH growth take place on very short time-scales, $\lesssim 2\,\mathrm{Myrs}$, the t_Q parameter in equation (24) will be overestimated, explaining, at least partially, the higher number density of sources predicted by the merger-driven model compared to observations, as will be highlighted in detail in Section 4.2.2.

Note that, for the *merger-driven* scenario, we considered a simplified version of the original model proposed by Pezzulli et al. (2016) mainly to investigate the impact of the accretion model on the overall shape of the BH luminosity and mass distributions.

A more sophisticated description of BH growth triggered by gas-rich galaxy mergers will be developed in future works to further explore its comparison to observations.

It is important to emphasize that the agreement of the *reference* model with observations of the bright end of the LF is an important result, as it implies that CAT is capable to reproduce reliable mass and luminosity distributions of the population of massive ($\gtrsim 10^7\,\mathrm{M_\odot}$) nuclear BHs down to $z \sim 4$, despite the BH seeding and growth model originally aimed at reproducing the evolution of the most extreme SMBHs ($\sim 10^9\,\mathrm{M_\odot}$) at $z \geq 6$. The above result suggests that the mechanisms which drive the formation and evolution of these classes of massive BHs across cosmic times remain the same, regardless of the mass of the final object and of its host galaxy.

However, as already observed, if we extrapolate the empirical LF at luminosities $\lesssim 10^{42}\,\mathrm{erg\,s^{-1}}$, the *reference* model seems to predict an excess of faint AGNs compared to observations. This may be a hint that the accretion mechanism assumed in CAT leads to an overgrowth of less massive BHs. In order to properly reproduce the faint end of the BH LF, in fact, similar semi-analytic models artificially shut off

BH growth below a given halo mass (e.g. Piana et al. 2021), assuming that this is caused by the impact of SN feedback (Anglés-Alcázar et al. 2017; Habouzit et al. 2017). Hence, an improved treatment of the accretion process of less massive nuclear BHs may be required to provide a better agreement with the empirical relations at low luminosities, as will be discussed in Section 5.

Nevertheless, we should note that the empirical LF predicted at $z \gtrsim 4$ are constrained mainly by luminous sources with $L_{bol} \gtrsim 10^{43}\,\mathrm{erg\,s^{-1}}$. The distributions at lower luminosities are thus highly uncertain, especially since the population of fainter AGNs might be heavily obscured, as suggested by several studies (Matsuoka et al. 2018; Giallongo et al. 2019). The improved agreement at faint luminosities observed in the UV and X-ray LFs between CAT predictions and several observations (see Section 4.2.2) supports the idea that these tensions might be due, at least partially, to observational limits.

Overall, this comparison suggests that the *reference* model provides the closest match to current observations, which however are able to probe only the bright-end of the predicted LF, dominated by *heavy* BH seeds descendants. Yet, at the very faint-end, below the gap, the LF is due to the more numerous population of inefficiently growing *light* BH seeds which remains completely invisible, as they lie at luminosities that are four orders of magnitudes lower than the faintest luminosity probed by current data. We also find that a significantly different evolutionary scenario, such as the *merger-driven* model, might lead to comparable results for the bright-end of the AGN luminosity distribution. The main differences show up instead in the low-luminosity region, making therefore very challenging to discern between different accretion models through observational campaigns.

A more detailed comparison of the model predictions with current and forthcoming observational facilities, in particular with the most promising deep sky surveys, is carried out in Section 4.2.2.

The improved sensitivity of the forthcoming generation of surveys might nevertheless push down the observational limit towards the threshold of $10^{41}\,\mathrm{erg\,s^{-1}}$, which is less than two orders of magnitude below the actual empirical constraints. In this range of luminosity the number density of accreting BHs is maximally sensitive to the assumed model for BH growth, as shown in Fig. 8, despite host contamination might becoming significant for such faint sources. Still, even at these luminosities, different scenarios for BH fuelling and accretion might result in similar luminosity distributions, characterized by a decreasing number of sources below $\sim 10^{40}\,\mathrm{erg\,s^{-1}}$.

4.2.2 BH UV and X-ray LF

In order to compare more extensively CAT predictions with current and future observational facilities, and in particular with the most promising deep sky surveys, we derived, for the three model variants, the BH luminosity distribution predicted in the UV and X-ray energy bands. The bolometric luminosity of each active BH can be converted into a *B* band and X-ray luminosity adopting specific bolometric corrections, such as in Duras et al. (2020):

$$\frac{L_{bol}}{L_B} = 5.13 \quad ; \qquad \frac{L_{bol}}{L_X} = a\left[1 + \left(\frac{\log(L_{bol}/L_\odot)}{b}\right)^c\right], \qquad (25)$$

where L_B and L_X are, respectively, the luminosity in the B (4400 Å) and (2–10) keV band, and $a = 10.96$, $b = 11.93$, and $c = 17.79$ are parameters calibrated on a population of both type 1 and type 2 quasars. Then, the *B*-band luminosity has been converted into a
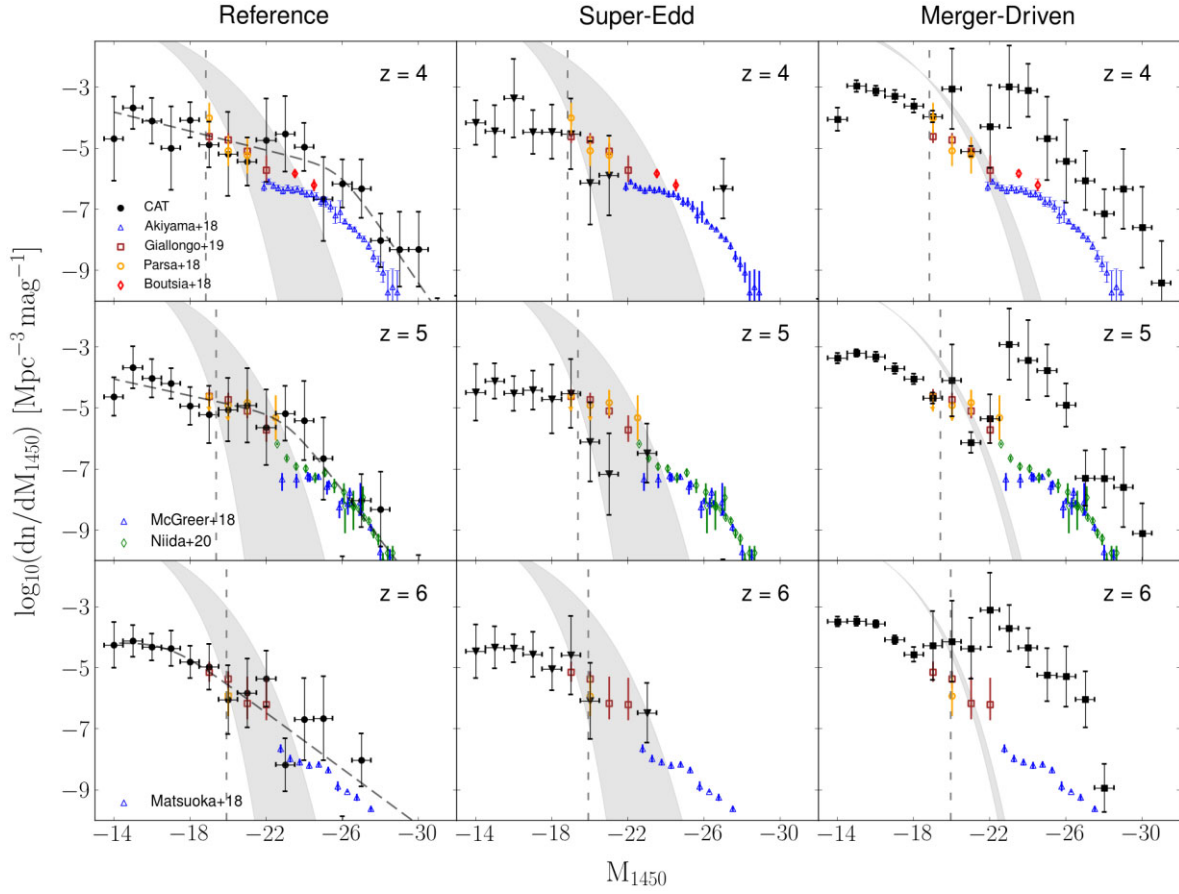
Figure 9. The AGN UV LFs at $z = 4$, 5, and 6 (from top to bottom). Each column shows the results obtained in one of the three CAT models: the *reference* model (left-hand panels), the *super-Edd* model (central panels), and the *merger-driven* model (right-hand panels). The black data points represent the predicted AGN LFs, properly corrected for obscuration as described in the text. The black dashed line shows the best fit for the *Reference* model. The grey shaded area encloses the region between the intrinsic and the dust-corrected galaxy UV luminosity function predicted by each model at the corresponding redshift. The coloured data represent the observational constraints on he AGN LF obtained by Parsa, Dunlop & McLure (2018, orange circles), Boutsia et al. (2018, red diamonds), Giallongo et al. (2019, brown squares), by the SHELLQs survey (Matsuoka et al. 2018), the CFHT Legacy Survey (McGreer et al. 2018), and the Hyper Suprime-Cam Wide Survey (Akiyama et al. 2018; Niida et al. 2020) (blue triangles and green diamonds). The vertical black dashed lines show the *JWST* luminosity limit at each redshift predicted by Griffin et al. (2020).

UV luminosity at 1450 Å assuming that $L_\nu \propto \nu^{-0.44}$ (Dayal et al. 2019).

However, for a proper comparison with observational data we must account for obscuration effects. In fact, the intrinsic LFs predicted by CAT have to be corrected for the fraction of obscured AGNs in each energy band. Following Merloni et al. (2014), we assume the fraction of obscured AGNs in the UV band to be a decreasing function of the intrinsic X-ray luminosity of the source. Hence, we compute the observable fraction as

$$f_{\rm obs} = 1 - 0.56 - \frac{1}{\pi} \arctan\left(\frac{43.89 - \log L_X}{0.46}\right). \qquad (26)$$

We show the resulting LFs, which we call obscured AGN UV LFs, in Fig. 9. Here, the predictions obtained for the three different models at redshift $z = 4$, 5, and 6 are represented by the black data points with error bars. The black long-dashed lines represent the best fit to the distributions at each redshift, which are well represented by broken power laws that flatten below a characteristic luminosity. For comparison, in each panel we also show the corresponding galaxy UV LFs predicted by CAT model at the same redshift. In order to take into account dust extinction, we corrected the galaxy UV

LFs as

$$L_{\rm UV,obs} = L_{\rm UV} \exp[-\Sigma_{\rm gas}\, \mathcal{D}\, k_{\rm UV}], \qquad (27)$$

where $\Sigma_{\rm gas}$ is the gas surface density, $\mathcal{D}$ the dust-to-gas mass ratio, and $k_{\rm UV}$ is the extinction coefficient per unit mass in the energy band of interest. The value of $k_{\rm UV}$ has been inferred considering the extinction curve of the Small Magellanic Cloud (Weingartner & Draine 2001). We assumed here a simple screen model where the optical depth is computed considering the contribution of all the gas and dust mass inside the galaxy. This will possibly result in an overestimation of the impact of dust obscuration, if compared to more sophisticated two-phase dust extinction models (see e.g. Mancini et al. 2016). Therefore, in Fig. 9, we show as a grey shaded area the region enclosed between the best fit of the intrinsic and dust-corrected galaxy UV LFs. In the *reference* and *super-Edd* models, we observe that dust extinction might heavily affect the galaxy UV luminosity. In the merger-driven scenario, instead, the lower abundance of gas due to the more competitive mechanisms of star formation and BH accretion translates into a smaller difference between the intrinsic and dust-corrected LFs.

However, for all the models and at all redshifts, the UV LF of AGNs dominates at the bright end, i.e. for magnitudes $M_{1450} \leq -22$. At fainter luminosities, up to $M_{1450} \sim -20$, the AGN LF still dominates only if the UV emission from the host galaxy is heavily reduced by dust extinction. Finally, at even fainter magnitudes the UV LF is dominated by the stellar emission from the host galaxies.

In the same figure, CAT model predictions are compared to the observations by the SHELLQs survey (Matsuoka et al. 2018), the CFHT Legacy Survey (McGreer et al. 2018), and the Hyper Suprime-Cam Wide Survey (Akiyama et al. 2018; Niida et al. 2020) (blue triangles and green diamonds). The figure also shows the results of the analysis by Parsa et al. (2018) and Giallongo et al. (2019), who estimated the AGN UV LF using X-ray data, resulting in an LF which should be less affected by obscuration. We also report the results of the COSMOS spectroscopic survey conducted by Boutsia et al. (2018).

Among the three models, the *reference* one provides the best agreement with observational constraints. In particular, at fainter magnitudes the predicted LF well trace the observations of Giallongo et al. (2019), Parsa et al. (2018), and Boutsia et al. (2018). In the bright end of the LF, instead, the distribution predicted by the *reference* model shows a more pronounced scatter due to the lower statistics, especially at $z \sim 6$. However, if we fit the obtained data with a broken power law, we find again a close agreement with the empirical data, despite a slight overprediction in the number density of bright sources at $z = 4$.

The agreement with the data is worse for the *super-Edd* model, which in fact, despite obtaining results similar to the *reference* model at $M_{UV} \gtrsim -19$, fails to reproduce the observed distributions at higher luminosities as pointed out also for the bolometric LF (Section 4.2).

Lastly, the *merger-driven* model shows an overall distribution very similar to that predicted by the *reference* model but shifted towards higher luminosities, which results in a higher number of sources at bright magnitudes $M_{1450} \lesssim -23$ compared to the observed one.

Interestingly we observe that at lower luminosities in all the three model variants the predicted LFs appear to be in better agreement with the LFs obtained through X-ray selection techniques. This suggests, as already outlined by Matsuoka et al. (2018) and Giallongo et al. (2019), that the apparent tension between different observational results might be due to an increasing incompleteness at fainter magnitudes or to a higher fraction of obscured AGNs towards lower luminosities.

It is important to note that in none of the models the UV LFs show the clear gap that was present in the bolometric LFs. This is because in the *reference* and *super-Edd* models, the gap appears at UV magnitudes $M_{1450} \geq -10$, which are several orders of magnitudes below the sensitivity limits of current observational facilities. Even with the sensitivity of *JWST*, whose luminosity limit is shown with vertical dashed lines in Fig. 9 (Griffin et al. 2020), it will be impossible to observe such faint sources, unless with the help of gravitational lensing. In addition, at magnitudes $M_{1450} \gtrsim -22$ deblending techniques need to be applied to discriminate the AGN emission and the emission coming from star formation in the host galaxy.

Using equation (25), we have also computed the LFs predicted by the three CAT models in the [2–10] keV X-ray energy band. The results are shown in Fig. 10. The black data points and error bars represent the binned intrinsic LFs at redshift $z = 4$, 5, and 6 (from top to bottom) and different columns refer to the three CAT model variants that we have considered (*reference*, *super-Edd*, and *merger-driven* models from left to right). In each panel, the dashed lines represent the best fit to the model predictions and the dotted lines

show the X-ray LFs contributed by star formation in the host galaxies at the same redshift. The latter component has been estimated using the empirical relation proposed by Fornasini et al. (2018):

$$L_{\mathrm{XRBs}} = 10^{29.98}(1+z)^{0.62}\,M_* + 10^{39.78}(1+z)^{0.2}\,\mathrm{SFR}^{0.84}, \quad (28)$$

where M_* and SFR are, respectively, the galaxy stellar mass in solar units and the SFR in solar masses per year. In Fig. 10, CAT predictions are compared with the X-ray LFs obtained by Fiore et al. (2012), Ueda et al. (2014), and Miyaji et al. (2015), which are some of the most complete studies of the AGN X-ray emission up to redshift $z \sim 4$–6. We also show the LF obtained by Vito et al. (2018) through a wide sample of AGNs at redshift $3.6 > z > 6$.

Since each of these works accounts differently for the fraction of absorbed and obscured AGNs, we decided to show with a shaded region, for the *reference* and the *merger-driven* models, how the best-fitting distribution of the CAT XLF changes considering absorption for the Compton-thin AGN population. We assume here the fraction of unabsorbed quasars $\psi_{\mathrm{X, unabs}}$ as a function of the X-ray luminosity as proposed by Ueda et al. (2014):

$$\psi_{\mathrm{X,unabs}} = 1 - \min\left[\psi_{\max}, \max[\psi^* - \beta\,(\mathrm{Log}\,L_X - 43.75)\,,\ \psi_{\min}]\right], \quad (29)$$

where $\psi_{\max} = 0.84$, $\psi_{\min} = 0.2$, $\psi^* \simeq 0.73$, and L_X is the [2 − 10] keV luminosity in erg s^{-1}. It is important to note, however, that the distributions corrected for quasar absorption represent only an upper limit for the unobscured AGN XLF, since we do not consider the contribution of the population of heavily obscured Compton-thick AGNs, which might be relevant especially at lower luminosities. In Fig. 10, the luminosity limits reported by Griffin et al. (2020) for the *Athena* and *Lynx* observatories are also shown, at each redshift, with dashed and dotted grey vertical lines, respectively.

In the X-ray band, the *reference* model provides again a very close agreement with the empirical data. The predicted LF well matches the observations in the entire luminosity range explored, $\sim 10^{42}$–10^{46} erg s^{-1}, and predicts similar values for the break magnitude of the double power-law distribution at all redshifts.

The X-ray LF produced by the *super-Edd* model has a shape similar to the *reference* one, but systematically underpredicts the number of bright sources above 10^{43} erg s^{-1}.

Finally, the *merger-driven* model predicts a peculiar luminosity distribution. Below $\sim 10^{44}$ erg s^{-1}, the X-ray LF has a shape similar to the *reference* model but with a slightly larger amplitude. At higher luminosities the distribution is dominated instead by the large number of systems undergoing the bursty post-merger accretion phase, leading to a large scatter and to an overestimation of the bright-end of the X-ray LF. However, we have to be careful in comparing with empirical data the luminosity of systems during such brief phases of merger-driven accretion. In fact, if these bursts of accretion take place on time-scales much shorter than the typical time-step of our model, the time interval t_Q (presented in equation 24) during which the BH undergoes an enhanced accretion will be overestimated, leading to a higher number of sources in the bright-end of the AGN LF. In addition, these rapidly accreting systems are supposed to be highly affected by obscuration, which would furtherly shorten the duration of the observable burst of luminosity.

A gap in the X-ray LFs appears for both the *reference* and the *Super-Edd* models, just below a luminosity of $\sim 10^{41}$ erg s^{-1}, as a consequence of the inefficient accretion of *light* BH seeds. Starting from the same luminosity, we observe also in the *merger-driven* variant a declining number of sources towards the faint-end of the distribution. In contrast with the previous case, this is due to the
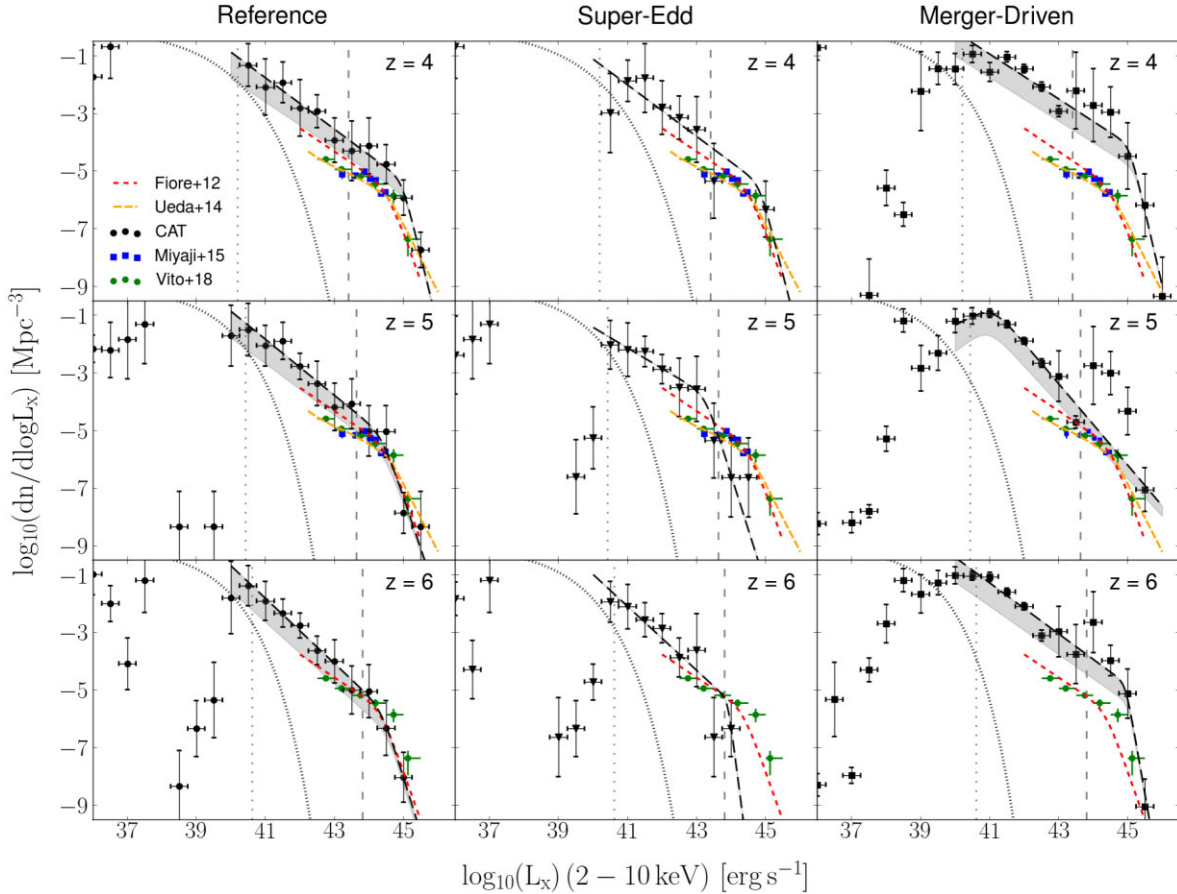
Figure 10. The AGN LFs in the $[2 - 10]$ keV X-ray energy band at $z = 4$, 5, and 6 (from top to bottom). Each column shows the results obtained in one of the three models examined: the *reference model* (left-hand panels), the *super-Edd* model (central panels), and the *merger-driven* model (right-hand panels). In the left- and right-hand panels, the grey shaded regions show how the best-fitting distributions change assuming the fraction of absorbed AGNs proposed by Ueda et al. (2014). In each panel, the black dotted line represents the X-ray LF associated with star formation in the host galaxies at the same redshift, which we have estimated using the relation proposed by Fornasini et al. (2018, see the text). The coloured data represent the observational results by Fiore et al. (2012, red dashed line, $z = 4.5, 6$), Ueda et al. (2014, orange dashed line, $4 < z < 5$), Miyaji et al. (2015, blue squares, $3 < z < 5.8$), and Vito et al. (2018, green circles, $3.6 < z < 6$). The dashed and dotted grey vertical lines represent the luminosity limits estimated by Griffin et al. (2020) for, respectively, *Athena* and *Lynx*.

very efficient growth of light seed descendants coupled with the increasingly less probable seed formation at lower redshift, as we already observed in Fig. 4. Nevertheless, the *merger-driven* model shows a significantly smoother decline with respect to the marked gap observed in the *reference* scenario. These different predictions might potentially be used to discern between different accretion mechanisms with future surveys exploring such faint luminosities, as will be discussed in the next section.

4.2.3 Predictions for future surveys

Distinctive features in the AGN X-ray LFs, characterizing the underlying model of accretion, might represent a key element to discern between different possible scenarios for early BH growth. According to CAT model predictions, such differences should affect primarily the evolution of lower mass BHs, requiring observational constraints at very faint luminosities, $L_X \lesssim 10^{41}$ erg s^{-1}. It is interesting to note that, as shown in Fig. 10, sources with an X-ray luminosity of $\sim 10^{41}$ erg s^{-1} might still be observable with the next-generation X-ray observatory *Lynx*, which will explore the faint end of the LFs, constraining the evolution of *light* BH seeds and their dominant accretion mode. *Athena* observations will be restricted

instead to brighter X-ray sources, constraining the distribution above 10^{43} erg s^{-1} (Griffin et al. 2020).

Still, it will be very challenging to investigate the AGN LFs sufficiently in detail to discern between different BH evolutionary scenarios. In fact, from Fig. 10 we observe that at luminosities $\leq 10^{41}$ erg s^{-1} the X-ray emission produced by stellar binaries formed in the host galaxies becomes comparable to the one contributed by AGNs. Therefore, the investigation of the X-ray luminosity distribution at such faint luminosities has to be carried out with methods that carefully take into account potential contamination from star formation in the host galaxies, in order to reliably discern between different BH growth modes. Moreover, the presence of a clear gap in the observed AGN number density below a given luminosity, as predicted in our *reference* model, might be covered, in future observations, by large scatters in the theoretical scaling relations or observational parameters assumed to estimate the empirical LFs.

In order to investigate the observational capability of the *Athena* and *Lynx* missions at even higher redshifts, in Fig. 11 we compare their forecast sensitivity in the redshift range $z \in [6, 7]$ and $[7, 8]$ with the AGN X LFs predicted by the CAT *reference* and *merger-driven* models at $z \sim 7$ and 8, accounting for obscuration. Similarly to the LFs shown at $z \leq 6$, a major difference between the two models is noticeable at the faint-end of the LFs, where the *reference*
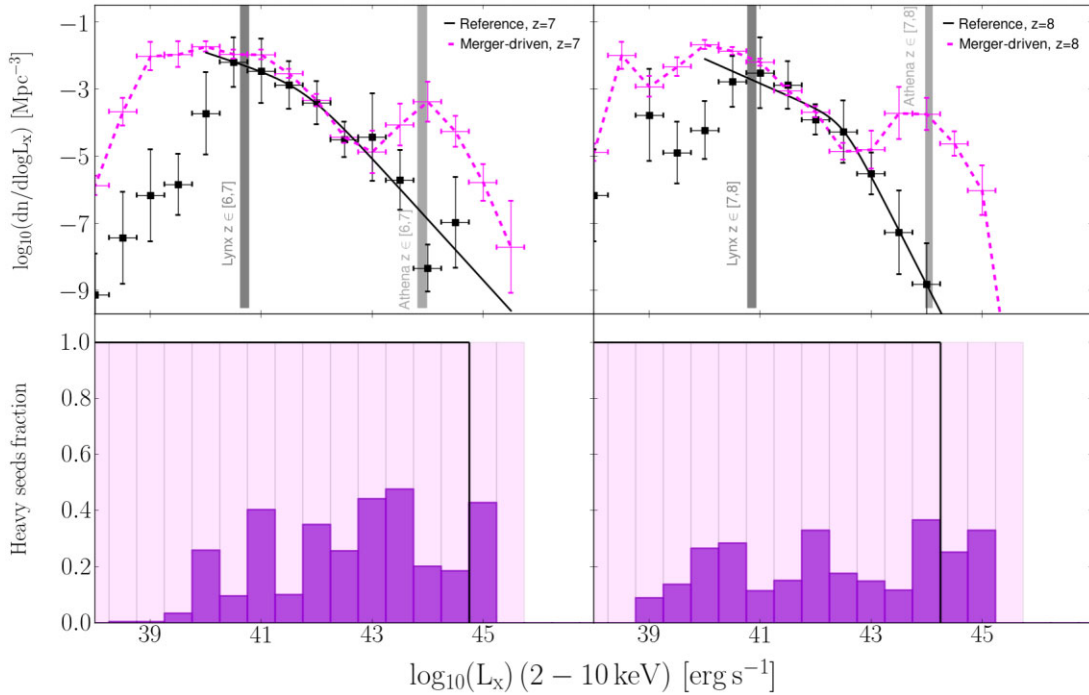
Figure 11. Top panels: AGN X-ray LFs predicted by CAT at $z \sim 7, 8$. The black data points are the results of our *reference* model, with the black solid curves representing the best fit of the bright-end distribution at each redshift. As a comparison, we also show the results of the *merger-driven* model with dashed magenta lines and error bars. The sensitivity limits proposed by Griffin et al. (2020) for a mission like Lynx and for Athena in the redshift ranges $z \in [6, 7]$, [7, 8] are shown, respectively, as dark grey and light grey vertical thick lines. Bottom panels: percentage of heavy BH seed descendants in each luminosity bin, for the *reference* (black lines) and *merger-driven* (violet histograms) models. In the *reference* model, all the AGNs with $\log_{10}(L_X) > 38$ have at least one heavy seed progenitor.

model shows a steeper decrease in the number density of fainter objects, due to the inefficient growth of light seeds descendants. This feature, however, appears just below the forecast sensitivity limit of *Lynx* and might be challenging to identify even with such deep observations. Despite that, our predictions suggest that a mission with sensitivity comparable to *Lynx* would have the potential to unveil a large population of AGNs, even at $z \geq 6-8$. This would open up the possibility to explore BHs with mass $\sim 10^5-10^6 \, \mathrm{M_\odot}$, which dominate the AGN population at $z \leq 15$, regardless of the assumed BH accretion model.

At these high redshifts, the forecast sensitivity of *Athena* would enable to explore only the bright end of the distribution, at $L_X > 10^{44} \, \mathrm{erg \, s^{-1}}$. While this hampers the possibility of detecting the dominant BH population at $z \sim 7-8$, an interesting feature appears $L_X \gtrsim 10^{43} \, \mathrm{erg \, s^{-1}}$, where the *merger-driven* model predicts a much larger number density of sources compared to the *reference* one. This is a consequence of the growing rate of galaxy mergers with increasing redshift, which causes a larger fraction of AGNs to experience enhanced, super-Eddington accretion, increasing their luminosity. Such a distinctive feature could be potentially observable by *Athena*. Therefore, the observation of a slower decline of the AGN X-ray LF at higher redshifts might be a hint that early BH evolution is strongly driven by short period of enhanced accretion occurring during galaxy mergers. An important caveat here is that the predicted luminosity (and observability) of such rapidly accreting BHs might be affected by the short time-scale of the process and by additional gas obscuration in the nuclear regions, as discussed in Section 4.2.2.

In the bottom panel of Fig. 11, we also show the percentage of BHs with at least one heavy seed progenitor in each luminosity bin. We observe that, in contrast with the results of the *reference* model,

where all the AGNs are predicted to descend from at least one heavy BH seed, in the *merger-driven* scenario, more than 50 per cent of AGNs descend from light BH seeds, even at the brightest X-ray luminosities. When BH growth can exceed the Eddington limit and it is triggered by galaxy mergers, the fraction of AGNs descending from heavy BH seeds is subdominant and decreases with redshift.

4.2.4 BH–galaxy scaling relations

In addition to the BH mass and luminosity distribution, we explored CAT predictions for the BH–galaxy scaling relations. In particular, it is interesting to understand if and how these relations are affected by the different BH growth scenario assumed in this work.

In Fig. 12, we show the M_*–BH relation for our sample of galaxies at $z = 5$, 6, and 7. The left and right columns represent, respectively, the results obtained in the *reference* and *merger-driven* models. We compare CAT predictions with several empirical relations based on AGN and galaxy observations in the local Universe, investigating both the unobscured (Reines & Volonteri 2015; Greene et al. 2016; Shankar et al. 2016; Suh et al. 2020) and obscured (Baron & Ménard 2019) AGN population. We also assume as a reference the predictions obtained from the empirical model recently presented by Zhang et al. (2021) at the redshifts of interest.

The CAT *reference* model shows again a clear gap around $M_{BH} \sim 10^4 \, \mathrm{M_\odot}$, which splits the two populations of galaxies hosting a *light* or a *heavy* BH seed descendant. Note that, in less massive galaxies, below $M_* \lesssim 10^9 \, \mathrm{M_\odot}$, the stellar mass is largely independent of the nature of the nuclear BH seed. The population of massive haloes with $M_* \gtrsim 10^9 \, \mathrm{M_\odot}$ hosts instead more massive BHs, showing a correlation between the two quantities. Remarkably, the galaxy
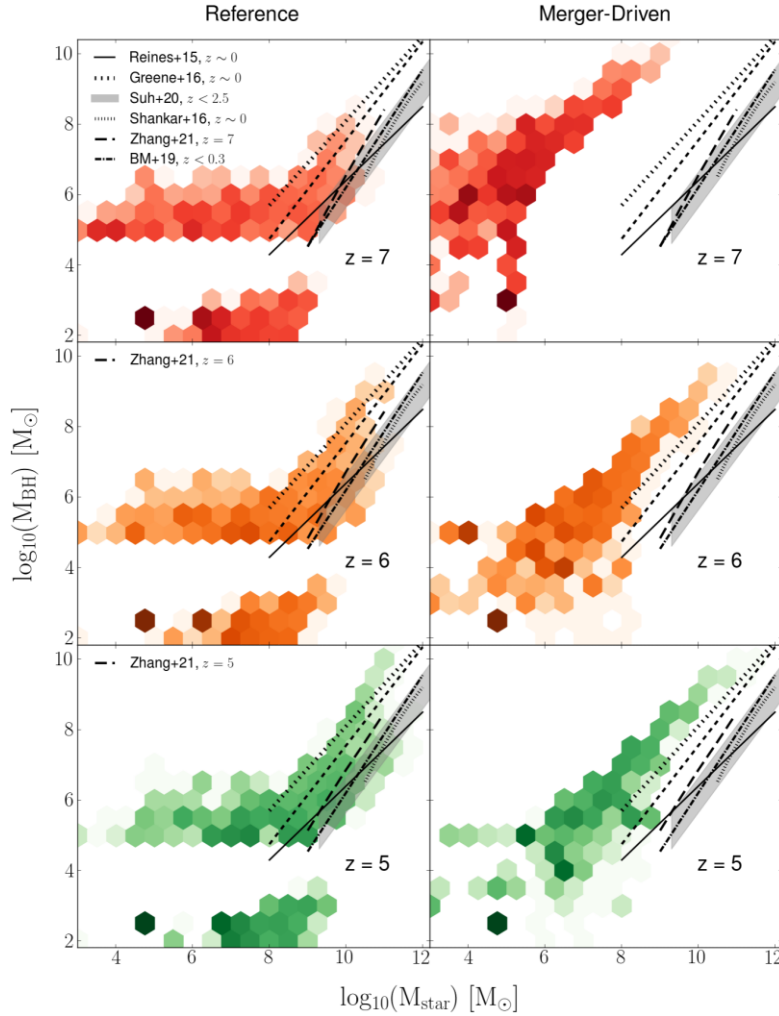
Figure 12. BH mass as a function of the galaxy stellar mass for the CAT galaxy population at different redshifts, $z = 5, 6, 7$. The left- and right-hand panels represent, respectively, the results obtained in the *reference* and *merger-driven* models. CAT results are compared with the predictions of the empirical models proposed by Reines & Volonteri (2015, solid and dashed lines for, respectively, AGNs and elliptical galaxies), Shankar et al. (2016, dotted line), Greene et al. (2016, loosely dotted line), Baron & Ménard (2019, dash–dotted line), Suh et al. (2020, grey shaded area), and Zhang et al. (2021, long dashed line).

population obtained by the CAT *reference* model closely reproduces the relations proposed by the empirically constrained models assumed as comparison.

The *merger-driven* model shows instead a continuous relation between the mass of the galaxy and that of its nuclear BH with an increasing scatter for lower mass galaxies. It is interesting to note that, also in this model, the galaxy population lies on a slope which is very similar to the ones predicted by numerical models. The very early growth of BHs, which characterize this model variant, is clearly noticeable at $z = 7$, where massive BHs populate less massive galaxies compared to the *reference* model. This is probably a natural consequence of the more competitive BH accretion model assumed, which affects the efficiency of star formation especially at early times. However, at lower redshift galaxies quickly increase their stellar mass, leading to a final distribution at $z \sim 5$ which presents an offset of $\lesssim 1$ dex compared to the prediction of empirical models.

The above results show that even very different paradigms for BH growth, as the ones considered in the *reference* and *merger-driven* models, lead to similar properties for the galaxy and BH populations at $z \sim 4$–5. Deep sky observations at higher redshift will be hence

crucial to better understand the nature of the first BH seeds and their co-evolution with the host galaxy.

5 DISCUSSION

In this section, we first compare the results of CAT with independent numerical and semi-analytical studies. We then discuss the main caveats of the model and how we plan to address these in the future.

5.1 Comparison with previous studies

The formation and evolution mechanisms of SMBHs at high redshift have been the focus of several studies in the last few years. Although very challenging, observational constraints on the BH MF at redshift $z \gtrsim 4$ have been proposed by different works, often relying on the correlations between the BH mass and the properties of the SMBH host (Merloni & Heinz 2008; Shankar et al. 2009, 2010; Willott et al. 2010b). However, the scaling relations between the BH mass and the host galaxy are mainly determined in the local Universe, while their evolution in redshift is still largely uncertain. Hence, in order to reconstruct the BH MF at higher redshift, these works usually

rely on the AGN luminosity distribution as a tracer of the accretion history of SMBHs, assuming the local BH distribution as a boundary condition. Unfortunately, this requires some assumptions on the efficiency of the BH accretion process, such as duty cycle, radiative efficiency, obscured AGN fraction, etc., leading to discrepancies between different results (see Kelly & Merloni 2012, for a detailed discussion). This point has to be carefully taken into account in the comparison between the observational constraints and the intrinsic MF obtained by theoretical models as CAT. Important efforts on the SMBHs evolution have been carried out also through large-scale cosmological simulations. This approach starts usually with a given cosmological framework and follows the baryonic evolution by zooming in particularly dense regions where the massive nuclear BHs are supposed to form. This class of simulations has shown to be able to characterize properly the evolution of a wide range of massive BHs. However, it still hardly succeeds in reproducing the most extreme sources that we observe at high redshift, which reach masses above $10^9 \, \mathrm{M_\odot}$ already at $z \geq 7$ (Mortlock et al. 2011; Yang et al. 2020; Wang et al. 2021). This is probably due to the large volume needed to resolve the rare overdense regions where the assembly of the most massive SMBHs can take place (Tenneti et al. 2018, 2019). The limited volume of this class of simulations leads indeed to a significant underestimation of the number of rare and bright AGNs at very high redshifts, which consequentially affects the predicted mass and LFs (Amarantidis et al. 2019).

In Fig. 13, we compare the binned MF at redshift $z = 4$ obtained by CAT *reference* model with the results of some of the most recent and important large-scale cosmological simulations, namely the Illustris (Sijacki et al. 2015), IllustrisTNG (Weinberger et al. 2017), Horizon-AGN (Volonteri et al. 2016), SIMBA (Davé et al. 2019; Thomas et al. 2019), and EAGLE (McAlpine et al. 2017, 2018). For completeness, we also show the observational constraints obtained from Merloni & Heinz (2008) and Shankar et al. (2009). We can see that the MF predicted by numerical simulations covers a smaller range of BH masses with respect to our semi-analytical approach. While at the high-mass end of the distribution this is due to the limited simulation volume, below $\sim 10^6 \, \mathrm{M_\odot}$ the nuclear BH population cannot be properly modelled due to the resolution limits of the simulations, which dictate the seeding prescription. In fact, since large-scale simulations are not able to well resolve lower mass galaxies, they are forced to seed more massive galaxies with $M_* \gtrsim 10^9 \, \mathrm{M_\odot}$ with nuclear BHs with mass $\sim 10^{5-6} \, \mathrm{M_\odot}$ (Habouzit et al. 2021).

The figure shows that numerical simulations provide consistent results at the high-mass end of the distribution, above $10^8 \, \mathrm{M_\odot}$, while they differ mostly at lower masses, probably as a result of the different sub-grid physics and seeding prescriptions adopted. The BH MF obtained in our *reference* model is in broad agreement with the results of the numerical simulations, especially in the BH mass range $[10^6 – 10^8] \, \mathrm{M_\odot}$, where it is consistent with the observationally constrained MF of Merloni & Heinz (2008). At higher mass values, where the contribution of the most massive and rarer systems becomes relevant, both CAT and numerical simulations seem instead to slightly overestimate the number of SMBHs with respect to Merloni & Heinz (2008). Fitting our binned SMBH MF with a power law, though, a better agreement is found with the observational constraints obtained by Shankar et al. (2009) in the high-mass range, above $\sim 10^9 \, \mathrm{M_\odot}$.

In Fig. 13, for comparison, we also show the best fit of the BH MF obtained by CAT *merger-driven* model. As anticipated in the previous sections, this model variant predicts a significantly larger number density of BHs with masses $\lesssim 10^9 \, \mathrm{M_\odot}$ compared to the *reference* model and to the results of numerical simulations and observational

studies. This is probably a consequence of the importance that *light* seeds acquire in this alternative scenario, where they are able to efficiently grow across cosmic time and contribute in the building up of the entire BH MF. It has to be pointed out, however, that the MF predicted by our *merger-driven* scenario at $z \sim 4$ seems in tension with the estimated local BH mass density inferred from the $M_{\mathrm{BH}} – M_*$ relation (see e.g. the recent work of Shankar et al. 2020). This suggests that a more refined modelling of the BH accretion process, as the one originally proposed by Pezzulli et al. (2016), which assumes a distinguished treatment for the galaxy bulge and disc environments, would be required in this scenario. That, in fact, might ensure a better consistency with the AGN mass and luminosity distributions, maintaining at the same time a good accordance with the global constraints.

In the bottom panel of Fig. 13, we also compare our results at $z = 5$ with similar semi-analytical studies. We show in particular the results of different SMBH evolutionary models studied by Ricarte & Natarajan (2018a), considering both burst and steady mode accretion. Although their focus is on massive BH seeds, mainly DCBHs or extremely fast-growing Pop III remnants, their results in the mass range $[10^7 – 10^{10}] \, \mathrm{M_\odot}$ are in very good agreement with CAT predictions. Interestingly, their results fall just between the best fit of our *reference* and *merger-driven* models, as we would expect for a mixed accretion scenario as the one considered there.

CAT predictions are also compared with the results of the *Delphi* semi-analytic model proposed by Piana et al. (2021). The two models implement similar approaches in investigating the BH mass distribution, following the evolution of both stellar and direct-collapse BH progenitors in a cosmological framework. However, CAT and *Delphi* present considerably different prescriptions for BH seed formation and growth. In fact, in the latter, BH seeding is less sensitive to the galaxy environmental conditions, since *all* DM haloes are initially seeded with a stellar ($150 \, \mathrm{M_\odot}$) or a DCBH seed ($10^{3-4} \, \mathrm{M_\odot}$), depending only on the initial incident LW radiation. The newly formed BHs are then assumed to accrete a given fraction of the available gas, independently of their mass, without exceeding the Eddington limit. Despite these differences, we find consistent results for the BH MF at $z \sim 5$ in the whole range of masses above $\sim 10^{6.5} \, M_\odot$. Interestingly, they do not find any gap in the MF at lower BH masses. This is probably due to the different BH accretion mode adopted in *Delphi*, which does not depend on the actual BH mass but only on the available gas mass in the host halo, at odds with our *reference* and *super-Edd* models, where BHs are assumed to grow at the Bondi-Hoyle rate. In addition, in *Delphi* the initial mass assumed for *heavy* (direct collapse) BHs is different from what we adopt in CAT: while we consider an initial mass of $10^5 \, \mathrm{M_\odot}$, in the centre of the supposed mass range for this class of seeds (Latif et al. 2016; Becerra et al. 2018), Piana et al. (2021) rely on a more conservative value of $10^{3-4} \, \mathrm{M_\odot}$. This difference translates into a continuous MF in the range $\sim 10^3 – 10^{10} \, \mathrm{M_\odot}$ and possibly accounts for the observed flattening in the MF with respect to CAT predictions at $M_{\mathrm{BH}} \lesssim 10^{6.5} \, \mathrm{M_\odot}$. This comparison suggests that the initial mass of *heavy* BH seeds has a strong influence on the shape of the AGN mass and luminosity distributions. In particular, a smaller birth mass of heavy seeds might reduce the predicted gap in the MF and therefore have an impact on the observability of the distinctive features characterizing different accretion models. However, it is important to note that predicting the IMF of *heavy* BH seeds is very challenging as this likely depends on the adopted conditions for their formation (Ferrara et al. 2014; Bhowmick et al. 2021).
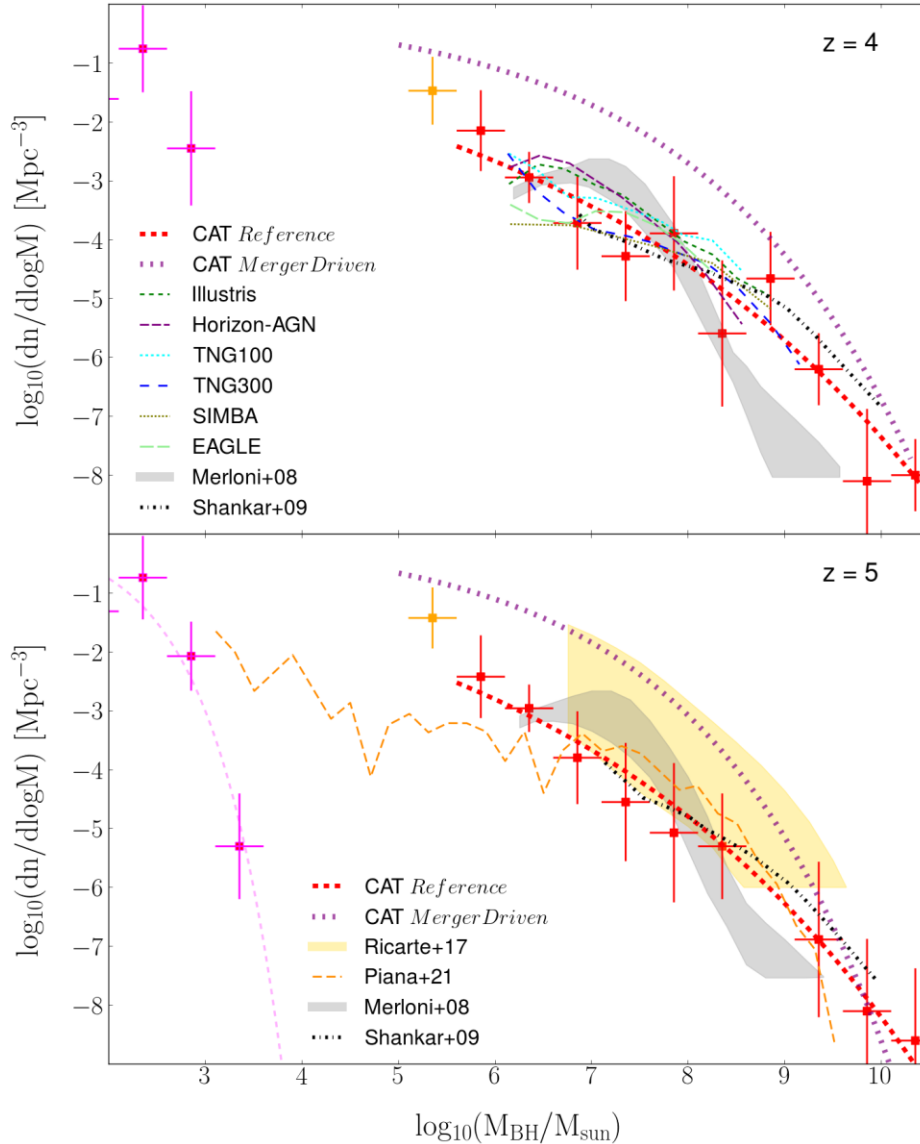
Figure 13. Comparison between the MF predicted by CAT and the results of numerical simulations (upper panel, $z = 4$) and semi-analytical models (lower panel, $z = 5$). As in Fig. 4, the data points show the binned distribution obtained by CAT *reference* model for *light* (magenta) and *heavy* (red, orange) BH descendants, with the red dashed line representing the best fit of the distribution of *heavy* seed descendants. The violet dotted lines show the best fit of the BH MF obtained by the *merger-driven* model. In the upper panel, as a comparison, we present with coloured dashed lines the results of large-scale cosmological simulations, namely: Illustris (dark green), IllustrisTNG100 (cyan), IllustrisTNG300 (blue), SIMBA (olive), and EAGLE (light green). In the lower panel we show the MF obtained by the semi-analytical models by Ricarte & Natarajan (2018a, yellow shaded region) and Piana et al. (2021, orange dashed line). The grey shaded area and the dash–dotted black lines represent instead the observational constraints proposed by Merloni & Heinz (2008) and Shankar et al. (2009), respectively.

5.2 Main caveats of the model

Despite the successes of CAT model predictions in reproducing a wide range of observational constraints (see Section 3) and the good agreement with independent studies discussed above, we plan to overcome some of the limitations that we have already anticipated in the previous sections and that we discuss below. The first one is the adopted seeding prescription. Although our model allows us to describe the formation of *light* and *heavy* BH seeds depending on the properties of their birth environment, we plan to expand our model following the approach of Sassano et al. (2021) to include the formation of *medium-weight* BH seeds by runaway stellar collisions in dense star clusters. This will allow us to track the formation and mass growth of three independent families of BH

seeds in a full cosmological context and to predict their observational signatures.

The paradigm assumed for the BH accretion in low-mass galaxies represents an additional crucial point. As observed in Section 4.2, the LFs obtained in the CAT *reference* model predicts an excess of faint sources, below $\sim 10^{42}$ erg s^{-1}, if compared to observations. Although empirical data are not strongly constrained at such low luminosities, this might be a consequence of considering an accretion model which is too efficient for BHs growing in low-mass galaxies. In similar semi-analytic models that are calibrated to reproduce the AGN LF at lower redshift, e.g. Piana et al. (2021), BH accretion is in fact artificially inhibited below a typical halo mass of $\sim 10^{12}$ M$_\odot$ to mimic the effect of SN feedback suggested by numerical studies

(Anglés-Alcázar et al. 2017; Habouzit et al. 2017). As shown in Fig. 13 this leads to a lower number density of BHs below $10^7 \, \mathrm{M_\odot}$ compared to CAT predictions. In future works, an improved prescription for BH accretion, assuming in particular an α_{BH} parameter (equation 12), which depends on the properties of the host galaxies, as proposed in Booth & Schaye (2009), will be crucial to better understand the nature of these discrepancies.

A further improvement of the model will be to include the effects of BH dynamics during galaxy mergers. As discussed in Section 2.3.3, in the current version of CAT, two BHs are assumed to merge only during major mergers of their DM host haloes, while in minor mergers the heaviest BH is assumed to migrate to the centre of the newly formed galaxy and the lightest one remains a satellite and we do not follow its evolution further. A more physical description of BH dynamics requires to take into account the impact of processes dominating on different spatial scales/cosmic epochs (Begelman, Blandford & Rees 1980; Armitage & Natarajan 2002, 2005; Colpi 2014): dynamical friction against background gas and stars regulates the sinking time-scale of BHs on kpc-to-pc scales, determining whether a bound system can form (e.g. Capelo et al. 2015; Pfister et al. 2017, 2019; Tamburello et al. 2017; Tamfal et al. 2018; Barausse et al. 2020; Bortolas et al. 2020, and references therein). On smaller scales, interactions with gaseous discs, stars, and other BHs (triple/multiple interactions) instead control the duration of the binary BH hardening phase (e.g. Bortolas et al. 2016; Bortolas, Mapelli & Spera 2018; Arca Sedda et al. 2019; Biava et al. 2019; Souza Lima et al. 2020, and references therein). We plan to study these aspects in future works, building on the first implementation of BH dynamics (triple BH interactions) in GQD recently proposed by Valiante et al. (2020).

6 SUMMARY AND CONCLUSIONS

In this work, we have used the CAT semi-analytical model to explore how different BH accretion and feedback prescriptions affect the formation and evolution of the first galaxies and their nuclear BHs from $z \sim 25$ to $z \sim 4$. In particular, our aim was to use CAT as a laboratory to test whether the nature of the first BH seeds and their growth mode may leave observable imprints on the BH mass and LFs. Note that, unlike the majority of numerical and semi-analytic models presented and used as a comparison in Section 5.1, CAT follows the formation of both *light* and *heavy* BH seeds, with a seeding prescription that depends on the physical conditions at their formation sites. This enables us to investigate how these two seed populations contribute to the statistical properties of the BH population at $4 \leq z \leq 6$.

We explored both a classic BHL accretion scenario, assumed as our *reference* model, and two model variants: the *super-Edd* model, where we removed the Eddington limit for BH accretion, and the *merger-driven* model, where – in addition – enhanced BH accretion episodes are triggered by galaxy mergers.

The most important results of this work are summarized below.

(i) The *reference* model shows the best agreement with observational data. The predicted BH luminosity function is in good agreement with several empirical constraints at $z = 4, 5, 6$, especially at higher luminosities. At the faint-end, close to the current observational limits, CAT slightly overpredicts the number of sources, suggesting an overstated growth for less massive BHs. The predicted BH MF is consistent with independent numerical models, as well as with different empirical constraints, although the agreement is limited to specific mass ranges as these empirical constraints do not always provide consistent results.

(ii) The *super-Edd* model seems instead to fail at reproducing the observations. The lack of an Eddington-limit leads to an accelerated early growth of BH seeds and thus to a quick depletion of gas inside their host galaxies. These conditions strongly affect the subsequent growth of nuclear BHs, which fail to reach both masses and luminosities of the most extreme quasars observed at $z \gtrsim 6$.

(iii) The *merger-driven* model predicts global trends for the mass and luminosity distributions very similar to the results of the *reference* model, despite not entirely consistent with empirical constraints and observational data. This tension might be also due to an intrinsic limit, since we implemented a simplified version of the original model developed by Pezzulli et al. (2016). In future works, refined prescriptions for BH accretion, as well as a more accurate treatment of the involved time-scales, will enable us to derive tighter constraints.

(iv) We find that the main difference between the *reference* and *merger-driven* accretion models lies in the evolution of the *light* BH seed population. In the first model, their stunted growth leads to a clear gap in the resulting mass and LFs, while in the second model they are able to grow in gas-rich galaxy mergers resulting into continuous BH mass and luminosity distributions where both *light* and *heavy* BH seeds can contribute to the same mass and luminosity bins.

(v) The signature of the BH seeds growth mode is imprinted in the BH luminosity function at the very faint end, in a luminosity regime that will be extremely challenging to test observationally. In the X-ray, a mission with a sensitivity comparable to the *Lynx* X-ray observatory might be able to probe the $z \sim 4$–6 LF at $L_X \lesssim 10^{41} \, \mathrm{erg \, s^{-1}}$, possibly unveiling precious hints on the BH accretion mechanism. Interestingly, we find that at $z \geq 6$–8 the forecast sensitivity of *Athena* could be enough to disentangle the signature of super-Eddington, *merger-driven* BH growth at $L_X > 10^{43} \, \mathrm{erg \, s^{-1}}$ by detecting a larger number of AGNs or a milder evolution at the bright end of the X-ray LF compared to the predictions of the *reference* model.

DATA AVAILABILITY

The simulated data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

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

Agarwal B., Khochfar S., 2015, MNRAS, 446, 160

Akiyama M. et al., 2018, PASJ, 70, S34

Amarantidis S. et al., 2019, MNRAS, 485, 2694

Ananna T. T. et al., 2019, ApJ, 871, 240

Andika I. T. et al., 2020, ApJ, 903, 34

Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017, MNRAS, 472, L109

Arca Sedda M., Berczik P., Capuzzo-Dolcetta R., Fragione G., Sobolenko M., Spurzem R., 2019, MNRAS, 484, 520

Armitage P. J., Natarajan P., 2002, ApJ, 567, L9

Armitage P. J., Natarajan P., 2005, ApJ, 634, 921

Bañados E. et al., 2016, ApJS, 227, 11

Bañados E. et al., 2018, Nature, 553, 473

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

Baron D., Ménard B., 2019, MNRAS, 487, 3404

Baugh C. M., 2006, Rep. Prog. Phys., 69, 3101

Becerra F., Greif T. H., Springel V., Hernquist L. E., 2015, MNRAS, 446, 2380

Becerra F., Marinacci F., Bromm V., Hernquist L. E., 2018, MNRAS, 480, 5029

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

Bellovary J., Volonteri M., Governato F., Shen S., Quinn T., Wadsley J., 2011, ApJ, 742, 13

Bhowmick A. K. et al., 2021, MNRAS, 507, 2012

Biava N., Colpi M., Capelo P. R., Bonetti M., Volonteri M., Tamfal T., Mayer L., Sesana A., 2019, MNRAS, 487, 4985

Boco L., Lapi A., Danese L., 2020, ApJ, 891, 94

Bondi H., 1952, MNRAS, 112, 195

Bonoli S., Mayer L., Callegari S., 2014, MNRAS, 437, 1576

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Bortolas E., Gualandris A., Dotti M., Spera M., Mapelli M., 2016, MNRAS, 461, 1023

Bortolas E., Mapelli M., Spera M., 2018, MNRAS, 474, 1054

Bortolas E., Capelo P. R., Zana T., Mayer L., Bonetti M., Dotti M., Davies M. B., Madau P., 2020, MNRAS, 498, 3601

Boutsia K., Grazian A., Giallongo E., Fiore F., Civano F., 2018, ApJ, 869, 20

Bouwens R. J. et al., 2012, ApJ, 754, 83

Bouwens R. J. et al., 2014, ApJ, 795, 126

Bromm V., 2013, Rep. Prog. Phys., 76, 112901

Bromm V., Loeb A., 2003, ApJ, 596, 34

Callegari S., Mayer L., Kazantzidis S., Colpi M., Governato F., Quinn T., Wadsley J., 2009, ApJ, 696, L89

Campanelli M., Lousto C., Zlochower Y., Merritt D., 2007, ApJ, 659, L5

Capelo P. R., Volonteri M., Dotti M., Bellovary J. M., Mayer L., Governato F., 2015, MNRAS, 447, 2123

Chehade B. et al., 2018, MNRAS, 478, 1649

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168

Cole S., Helly J., Frenk C. S., Parkinson H., 2008, MNRAS, 383, 546

Chon S., Omukai K., 2020, MNRAS, 494, 2851

Chon S., Hirano S., Hosokawa T., Yoshida N., 2016, ApJ, 832, 134

Chon S., Hosokawa T., Omukai K., 2021, MNRAS, 502, 700

Colpi M., 2014, Space Sci. Rev., 183, 189

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

Dayal P., Ward M., Cockell C., 2016, preprint (arXiv:1606.09224)

Dayal P., Rossi E. M., Shiralilou B., Piana O., Choudhury T. R., Volonteri M., 2019, MNRAS, 486, 2336

Dayal P. et al., 2020, MNRAS, 495, 3065

de Bennassuti M., Schneider R., Valiante R., Salvadori S., 2014, MNRAS, 445, 3039

de Bennassuti M., Salvadori S., Schneider R., Valiante R., Omukai K., 2017, MNRAS, 465, 926

De Lucia G., 2019, Galaxies, 7, 56

De Rosa G., Decarli R., Walter F., Fan X., Jiang L., Kurk J., Pasquali A., Rix H. W., 2011, ApJ, 739, 56

DeGraf C., Di Matteo T., Khandai N., Croft R., Lopez J., Springel V., 2012, MNRAS, 424, 1892

DeGraf C., Di Matteo T., Treu T., Feng Y., Woo J. H., Park D., 2015, MNRAS, 454, 913

Devecchi B., Volonteri M., 2009, ApJ, 694, 302

Devecchi B., Volonteri M., Rossi E. M., Colpi M., Portegies Zwart S., 2012, MNRAS, 421, 1465

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Di Matteo T., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, ApJ, 745, L29

Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, MNRAS, 467, 4243

Duncan K. et al., 2014, MNRAS, 444, 2960

Dunn G., Holley-Bockelmann K., Bellovary J., 2020, ApJ, 896, 72

Duras F. et al., 2020, A&A, 636, A73

Eilers A.-C., Davies F. B., Hennawi J. F., 2018, ApJ, 864, 53

Eilers A.-C. et al., 2020, ApJ, 900, 37

Ellis R. S. et al., 2013, ApJ, 763, L7

Fan X., Narayanan V. K., Strauss M. A., White R. L., Becker R. H., Pentericci L., Rix H.-W., 2002, AJ, 123, 1247

Fan X. et al., 2003, AJ, 125, 1649

Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, MNRAS, 455, 2778

Ferrara A., Salvadori S., Yue B., Schleicher D., 2014, MNRAS, 443, 2410

Fiore F. et al., 2012, A&A, 537, A16

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

Giallongo E. et al., 2019, ApJ, 884, 19

González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJ, 735, L34

Grazian A. et al., 2015, A&A, 575, A96

Graziani L., Schneider R., Ginolfi M., Hunt L. K., Maio U., Glatzle M., Ciardi B., 2020, MNRAS, 494, 1071

Greene J. E. et al., 2016, ApJ, 826, L32

Greif T. H., Springel V., White S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011, ApJ, 737, 75

Griffin A. J., Lacey C. G., Gonzalez-Perez V., Lagos C. d. P., Baugh C. M., Fanidakis N., 2019, MNRAS, 487, 198

Griffin A. J., Lacey C. G., Gonzalez-Perez V., Lagos C. d. P., Baugh C. M., Fanidakis N., 2020, MNRAS, 492, 2535

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

Haemmerlé L., Meynet G., Mayer L., Klessen R. S., Woods T. E., Heger A., 2019, A&A, 632, L2

Haiman Z., Thoul A. A., Loeb A., 1996, ApJ, 464, 523

Heger A., Woosley S. E., 2002, ApJ, 567, 532

Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, ApJ, 781, 60

Hirano S., Zhu N., Yoshida N., Spergel D., Yorke H. W., 2015, ApJ, 814, 18

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

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

Hosokawa T., Omukai K., Yorke H. W., 2012, ApJ, 756, 93

Hosokawa T., Hirano S., Kuiper R., Yorke H. W., Omukai K., Yoshida N., 2016, ApJ, 824, 119

Hoyle F., Lyttleton R. A., 1941, MNRAS, 101, 227

Huang K.-W., Ni Y., Feng Y., Di Matteo T., 2020, MNRAS, 496, 1

Inayoshi K., Tanaka T. L., 2015, MNRAS, 450, 4350

Inayoshi K., Omukai K., Tasker E., 2014, MNRAS, 445, L109

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

Inayoshi K., Visbal E., Haiman Z., 2020, ARA&A, 58, 27

Johnson J. L., Haardt F., 2016, Publ. Astron. Soc. Aust., 33, e007

Katz H., Sijacki D., Haehnelt M. G., 2015, MNRAS, 451, 2352

Kelly B. C., Merloni A., 2012, Adv. Astron., 2012, 970858

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kistler M. D., Yüksel H., Beacom J. F., Hopkins A. M., Wyithe J. S. B., 2009, ApJ, 705, L104

Labbé I. et al., 2013, ApJ, 777, L19

Latif M. A., Ferrara A., 2016, Publ. Astron. Soc. Aust., 33, e051

Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J. C., 2013, MNRAS, 436, 2989

Latif M. A., Schleicher D. R. G., Hartwig T., 2016, MNRAS, 458, 233

Lodato G., Natarajan P., 2006, MNRAS, 371, 1813

Lodato G., Natarajan P., 2007, MNRAS, 377, L64

Lupi A., Colpi M., Devecchi B., Galanti G., Volonteri M., 2014, MNRAS, 442, 3616

Lupi A., Volonteri M., Decarli R., Bovino S., Silk J., Bergeron J., 2019, MNRAS, 488, 4004

Lupi A., Haiman Z., Volonteri M., 2021, MNRAS, 503, 5046

Madau P., Dickinson M., 2014, ARA&A, 52, 415

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

Mancini M., Schneider R., Graziani L., Valiante R., Dayal P., Maio U., Ciardi B., Hunt L. K., 2015, MNRAS, 451, L70

Mancini M., Schneider R., Graziani L., Valiante R., Dayal P., Maio U., Ciardi B., 2016, MNRAS, 462, 3130

Matsuoka Y. et al., 2018, ApJ, 869, 150

Matsuoka Y. et al., 2019, ApJ, 872, L2

Mayer L., 2019, Formation of the First Black Holes. World Scientific Publishing Co. Pte. Ltd., Singapore, p. 195

Mayer L., Bonoli S., 2019, Rep. Prog. Phys., 82, 016901

Mayer L., Kazantzidis S., Madau P., Colpi M., Quinn T., Wadsley J., 2007, Science, 316, 1874

Mayer L., Kazantzidis S., Escala A., Callegari S., 2010, Nature, 466, 1082

Mayer L., Fiacconi D., Bonoli S., Quinn T., Roškar R., Shen S., Wadsley J., 2015, ApJ, 810, 51

Mazzucchelli C. et al., 2017, ApJ, 849, 91

McAlpine S., Bower R. G., Harrison C. M., Crain R. A., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 468, 3395

McAlpine S., Bower R. G., Rosario D. J., Crain R. A., Schaye J., Theuns T., 2018, MNRAS, 481, 3118

McGreer I. D., Fan X., Jiang L., Cai Z., 2018, AJ, 155, 131

Merlin E. et al., 2019, MNRAS, 490, 3309

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

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

Miyaji T. et al., 2015, ApJ, 804, 104

Mortlock D. J. et al., 2011, Nature, 474, 616

Natarajan P., Pacucci F., Ferrara A., Agarwal B., Ricarte A., Zackrisson E., Cappelluti N., 2017, ApJ, 838, 117

Niida M. et al., 2020, ApJ, 904, 89

Oesch P. A. et al., 2014, ApJ, 786, 108

Omukai K., 2001, ApJ, 546, 635

Omukai K., Tsuribe T., Schneider R., Ferrara A., 2005, ApJ, 626, 627

Omukai K., Schneider R., Haiman Z., 2008, ApJ, 686, 801

Onoue M. et al., 2019, ApJ, 880, 77

Pacucci F., Ferrara A., Volonteri M., Dubus G., 2015, MNRAS, 454, 3771

Pacucci F., Ferrara A., Grazian A., Fiore F., Giallongo E., Puccetti S., 2016, MNRAS, 459, 1432

Pacucci F., Natarajan P., Volonteri M., Cappelluti N., Urry C. M., 2017, ApJ, 850, L42

Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557

Parsa S., Dunlop J. S., McLure R. J., 2018, MNRAS, 474, 2904

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

Pezzulli E., Valiante R., Orofino M. C., Schneider R., Gallerani S., Sbarrato T., 2017a, MNRAS, 466, 2131

Pezzulli E., Volonteri M., Schneider R., Valiante R., 2017b, MNRAS, 471, 589

Pfister H., Lupi A., Capelo P. R., Volonteri M., Bellovary J. M., Dotti M., 2017, MNRAS, 471, 3646

Pfister H., Volonteri M., Dubois Y., Dotti M., Colpi M., 2019, MNRAS, 486, 101

Piana O., Dayal P., Volonteri M., Choudhury T. R., 2021, MNRAS, 500, 2146

Planck Collaboration VI, 2020, A&A, 641, A6

Pons E., McMahon R. G., Simcoe R. A., Banerji M., Hewett P. C., Reed S. L., 2019, MNRAS, 484, 5142

Reed S. L. et al., 2017, MNRAS, 468, 4702

Reed S. L. et al., 2019, MNRAS, 487, 1874

Regan J. A., Haehnelt M. G., 2009, MNRAS, 396, 343

Regan J. A., Downes T. P., Volonteri M., Beckmann R., Lupi A., Trebitsch M., Dubois Y., 2019, MNRAS, 486, 3892

Regan J. A., Haiman Z., Wise J. H., O'Shea B. W., Norman M. L., 2020, Open J. Astrophys., 3, 1

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

Reinoso B., Schleicher D. R. G., Fellhauer M., Klessen R. S., Boekholt T. C. N., 2018, A&A, 614, A14

Reinoso B., Schleicher D., Fellhauer M., Klessen R., Boekholt T., Vergara M., Alister Seguel P., 2019, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 61, 154

Ricarte A., Natarajan P., 2018a, MNRAS, 474, 1995

Ricarte A., Natarajan P., 2018b, MNRAS, 481, 3278

Sądowski A., 2009, ApJS, 183, 171

Sakurai Y., Yoshida N., Fujii M. S., Hirano S., 2017, MNRAS, 472, 1677

Salvaterra R., Haardt F., Volonteri M., Moretti A., 2012, A&A, 545, L6

Sassano F., Schneider R., Valiante R., Inayoshi K., Chon S., Omukai K., Mayer L., Capelo P. R., 2021, MNRAS, 506, 613

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

Schenker M. A. et al., 2013, ApJ, 768, 196

Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30

Schneider R., Omukai K., Inoue A. K., Ferrara A., 2006, MNRAS, 369, 1437

Schneider R., Omukai K., Bianchi S., Valiante R., 2012, MNRAS, 419, 1566

Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33

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

Shankar F., Crocce M., Miralda-Escudé J., Fosalba P., Weinberg D. H., 2010, ApJ, 718, 231

Shankar F. et al., 2016, MNRAS, 460, 3119

Shankar F. et al., 2020, Nat. Astron., 4, 282

Shao Y. et al., 2017, ApJ, 845, 138

Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, MNRAS, 495, 3252

Shen Y., Liu X., 2012, ApJ, 753, 125

Shen Y. et al., 2011, ApJS, 194, 45

Shen Y. et al., 2019, ApJ, 873, 35

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

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

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Song M. et al., 2016, ApJ, 825, 5

Souza Lima R., Mayer L., Capelo P. R., Bortolas E., Quinn T. R., 2020, ApJ, 899, 126

Springel V. et al., 2005, Nature, 435, 629

Stacy A., Bromm V., Lee A. T., 2016, MNRAS, 462, 1307

Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, ApJ, 763, 129

Sugimura K., Matsumoto T., Hosokawa T., Hirano S., Omukai K., 2020, ApJ, 892, L14

Suh H., Civano F., Trakhtenbrot B., Shankar F., Hasinger G., Sanders D. B., Allevato V., 2020, ApJ, 889, 32

Takeo E., Inayoshi K., Ohsuga K., Takahashi H. R., Mineshige S., 2018, MNRAS, 476, 673

Tamburello V., Capelo P. R., Mayer L., Bellovary J. M., Wadsley J. W., 2017, MNRAS, 464, 2952

Tamfal T., Capelo P. R., Kazantzidis S., Mayer L., Potter D., Stadel J., Widrow L. M., 2018, ApJ, 864, L19

Tanaka T., Haiman Z., 2009, ApJ, 696, 1798

Tenneti A., Di Matteo T., Croft R., Garcia T., Feng Y., 2018, MNRAS, 474, 597

Tenneti A., Wilkins S. M., Di Matteo T., Croft R. A. C., Feng Y., 2019, MNRAS, 483, 1388

Thomas N., Davé R., Anglés-Alcázar D., Jarvis M., 2019, MNRAS, 487, 5764

Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121

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

Valiante R., Schneider R., Salvadori S., Bianchi S., 2011, MNRAS, 416, 1916

Valiante R., Schneider R., Maiolino R., Salvadori S., Bianchi S., 2012, MNRAS, 427, L60

Valiante R., Schneider R., Salvadori S., Gallerani S., 2014, MNRAS, 444, 2442

Valiante R., Schneider R., Volonteri M., Omukai K., 2016, MNRAS, 457, 3356

Valiante R., Agarwal B., Habouzit M., Pezzulli E., 2017, Publ. Astron. Soc. Aust., 34, e031

Valiante R., Schneider R., Graziani L., Zappacosta L., 2018a, MNRAS, 474, 3825

Valiante R., Schneider R., Zappacosta L., Graziani L., Pezzulli E., Volonteri M., 2018b, MNRAS, 476, 407

Valiante R. et al., 2020, MNRAS, 500, 4095

Vito F. et al., 2018, MNRAS, 473, 2378

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

Volonteri M., 2010, A&AR, 18, 279

Volonteri M., Natarajan P., 2009, MNRAS, 400, 1911

Volonteri M., Lodato G., Natarajan P., 2008, MNRAS, 383, 1079

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

Wang F. et al., 2018, ApJ, 869, L9

Wang F. et al., 2020, ApJ, 896, 23

Wang F. et al., 2021, ApJ, 907, L1

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

Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296

Willott C. J. et al., 2010a, AJ, 139, 906

Willott C. J. et al., 2010b, AJ, 140, 546

Wise J. H., Turk M. J., Abel T., 2008, ApJ, 682, 745

Wise J. H., Regan J. A., O'Shea B. W., Norman M. L., Downes T. P., Xu H., 2019, Nature, 566, 85

Woods T. E. et al., 2019, Publ. Astron. Soc. Aust., 36, e027

Wu X.-B. et al., 2015, Nature, 518, 512

Yang J. et al., 2020, ApJ, 897, L14

Yue B., Ferrara A., Salvaterra R., Xu Y., Chen X., 2013, MNRAS, 433, 1556

Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Hopkins P. F., Yang J., Aird J., 2021, preprint (arXiv:2105.10474)

Zhu Q., Li Y., Li Y., Maji M., Yajima H., Schneider R., Hernquist L., 2020, preprint (arXiv:2012.01458)

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