Chasing the observational signatures of seed black holes at $z > 7$: candidate statistics

Rosa Valiante,[1][*] Raffaella Schneider,[1,2] Luca Graziani[1] and Luca Zappacosta[1]

[1]*INAF-Osservatorio Astronomico di Roma, via di Frascati 33, I-00040, Monteporzio Catone, Italy*
[2]*Dipartimento di Fisica, Universitá di Roma 'La Sapienza' P.le Aldo Moro 2, I-00185 Roma, Italy*

ABSTRACT

Supermassive black holes (SMBHs) of 10^9–10^{10} M$_\odot$ were already in place ∼13 Gyr ago, at $z > 6$. Super-Eddington growth of low-mass BH seeds (∼100 M$_\odot$) or less extreme accretion on to∼10^5 M$_\odot$ seeds have been recently considered as the main viable routes to these SMBHs. Here, we study the statistics of these SMBH progenitors at $z \sim 6$. The growth of low- and high-mass seeds and their host galaxies are consistently followed using the cosmological data constrained model GAMETE/QSODUST, which reproduces the observed properties of high-z quasars, like SDSS J1148+5251. We show that both seed formation channels can be in action over a similar redshift range $15 < z < 18$ and are found in dark matter haloes with comparable mass, ∼5×10^7 M$_\odot$. However, as long as the systems evolve in isolation (i.e. no mergers occur), noticeable differences in their properties emerge: At $z \geq 10$ galaxies hosting high-mass seeds have smaller stellar mass and metallicity, the BHs accrete gas at higher rates and star formation proceeds less efficiently than in low-mass seeds hosts. At $z < 10$ these differences are progressively erased, as the systems experience minor or major mergers and every trace of the BH origin gets lost.

Key words: stars: black holes – galaxies: active – quasars: general – quasars: supermassive black holes.

1 INTRODUCTION

The first billion years of the Universe must have been the stage of intense black hole (BH) activity, as shown by the large number of quasars (>170) discovered at $z > 5.6$ (e.g. Bañados et al. 2016; Jiang et al. 2016, and references therein).

The luminosity of the brightest ones ($> 10^{47}$ erg s^{-1}) is commonly interpreted as an indication of actively accreting supermassive black holes (SMBHs) of $>10^9$ M$_\odot$ in their hosting galaxies (Fan et al. 2001; De Rosa et al. 2011, 2014). To date, the most massive BH observed at $z > 6$ is the one powering the quasar SDSS J0100+2802 (Wu et al. 2015), while the most distant SMBH is located at $z = 7.08$ (Mortlock et al. 2011).

In the standard paradigm of BH evolution, a SMBH is expected to be the evolutionary product of a less massive BH seed, which subsequently grows via gas accretion and mergers. The BH seeds birth mass function and the physical mechanisms driving their formation and growth are among the most studied issues in the recent literature (see reviews by Volonteri 2010; Natarajan 2011; Latif & Ferrara 2016; Johnson & Haardt 2016; Valiante et al. 2017).

BHs forming as remnants of Population III (Pop III) stars have been predicted to have masses from few tens to few hundreds (or even thousands) solar masses (e.g. Madau & Rees 2001; Heger et al. 2003; Yoshida, Omukai & Hernquist 2008; Latif et al. 2013b; Hirano et al. 2015). They are also supposed to become supermassive by $z \sim$ 6–7 only if their growth is driven by uninterrupted, although unrealistic, gas accretion at the Eddington rate or, alternatively, by short (intermittent) periods of supercritical feeding (Haiman 2004; Volonteri & Rees 2005, 2006; Tanaka & Haiman 2009; Volonteri, Silk & Dubus 2015; Lupi et al. 2016; Pezzulli, Valiante & Schneider 2016; Pezzulli et al. 2017).

Alternative routes to the first SMBHs are provided by intermediate mass seeds ($>[10^3$–$10^4]$ M$_\odot$) forming in dense stellar cluster either via runaway collisions (Omukai, Schneider & Haiman 2008; Devecchi & Volonteri 2009; Devecchi et al. 2010, 2012; Alexander & Natarajan 2014; Katz, Sijacki & Haehnelt 2015; Sakurai et al. 2017) or throughout gas-driven core-collapse (Davies, Miller & Bellovary 2011; Lupi et al. 2014), and by massive BH seeds with $[10^4$–$10^6]$ M$_\odot$, the so-called direct collapse BHs (DCBHs). DCBHs are supposed to form via the rapid collapse of metal poor gas clouds in which star formation is prevented either by high velocity galaxy collisions, or by intense fluxes of H$_2$ photo-dissociating photons (e.g. Bromm & Loeb 2003; Lodato & Natarajan 2006; Spaans & Silk 2006; Volonteri, Lodato & Natarajan 2008; Inayoshi & Omukai 2012; Latif et al. 2013a,b, 2014a; Agarwal et al. 2014; Sugimura, Omukai & Inoue 2014; Visbal, Haiman & Bryan 2014; Inayoshi, Visbal & Kashiyama 2015; Regan et al. 2017).

[*] E-mail: rosa.valiante@oa-roma.inaf.it

In a previous work (Valiante et al. 2016, hereafter V16), we explored the relative role of Pop III remnant BHs and DCBHs in the formation pathways to the first SMBHs considering an Eddington-limited BH accretion scenario. We used a semi-analytic, data-constrained, cosmological model called GAMETE/QSODUST aimed at study the formation of high-redshift quasars and their host galaxies (Valiante et al. 2011, 2014, 2016). In V16, we also showed that the condition for the formation of DCBHs are very tight, so that they form in a small number of haloes (from 3 to 30 depending on the merger history) and in a very narrow redshift range ($16 < z < 18$). Still, the Eddington-limited growth of a SMBH of $>10^9$ M$_\odot$, as inferred for SDSS J1148+5251, relies on the formation of these few massive seeds.

The aim of this work is to investigate to what extent the properties of BH progenitors and their host galaxies can be used to disentangle the nature of the first BH seeds, before their memory is lost along their cosmological evolution. Here, we extend the analysis presented in V16, exploring additional physical properties of the BHs and their hosts. In a companion paper (Valiante et al., in preparation), we will make observational predictions for the most promising candidates. Our approach has the advantage of following the evolution of light and heavy seed progenitors along the same hierarchical history (i.e. the two BH seed formation channels are not mutually exclusive), starting from the epoch at which the first stars form ($z \sim 24$, in our model) down to $z \sim 6$, when the observed quasar is eventually assembled. Moreover, the properties of the BH seeds population are self-consistently related to the evolution of their host galaxies: mass, number, redshift distribution, and growth history are regulated by the build-up of the ultraviolet (UV) radiation field, by the metals and dust pollution of the interstellar medium (ISM) and intergalactic medium (IGM), and by the effect of stellar and active galactic nucleus (AGN) driven winds.

This paper is organized as follows. In Section 2, we briefly summarize the main features of GAMETE/QSODUST. In Section 3, we present the properties of the sample of accreting light and heavy seeds as predicted by our model. Finally, we discuss our results drawing the conclusions of the study in Section 4.

2 BH FORMATION AND EVOLUTION MODEL

Here, we briefly introduce the semi-analytic model GAMETE/QSODUST, aimed at studying the formation and evolution of high redshift quasars and their host galaxies at $z > 5$. We refer the reader to Valiante et al. (2011, 2014, 2016) for a full description of the code.

The model successfully reproduces the BH mass and the properties of the quasars host galaxies, such as the mass of gas, metals and dust and the star formation rate. In this analysis, we select the quasar SDSS J1148+5251, observed at $z = 6.4$, as our target. This is one of the best studied objects at $z > 6$ and it is powered by a SMBH of $(2–6) \times 10^9$ M$_\odot$ (Barth et al. 2003; Willott, McLure & Jarvis 2003). We summarized the other main observed properties of this quasar and its host galaxy in Valiante et al. (2011) and Valiante et al. (2014).

With the same model, we have investigated the nuclear BHs-host galaxies co-evolution histories for a sample of $z > 5$ quasars, reproducing their observed properties (Valiante et al. 2014).

2.1 Hierarchical merger histories

We developed a Monte Carlo algorithm, based on the extended Press–Schechter theory (EPS), to reconstruct several hierarchical merger histories of the 10^{13} M$_\odot$ host dark matter (DM) haloes. As

introduced in V16, we also extended the DM haloes mass spectrum to mini-haloes[1], which are resolved down to redshift $z \sim 14$ along the merger trees. These are the birth places of the first generation of stars (Pop III stars), and at $z \gtrsim 17$, they represent the dominant population among DM progenitors. Their number progressively decreases at lower redshift, down to $z \sim 14$, below which more massive haloes, with $T_{\rm vir} \geq 10^4$ K (called Ly α cooling haloes), dominate the halo population.

The use of a Monte Carlo approach based on the EPS theory to simulate the hierarchical evolution of a 10^{13} M$_\odot$ DM halo at $z \sim 6$ offers the following advantages: It enables to (i) resolve early star-forming mini-haloes at $z > 20$, (ii) run several independent hierarchical histories of this 'biased' region of the Universe, and (iii) sample a large parameter space, exploring the impact of poorly constrained physical processes. This combination of features cannot be found in current numerical simulations. In fact, simulations cannot resolve large and small scales at the same time: small scale simulations allow us to follow the physical processes accurately (for instance, the combination of feedback effects in star-forming mini-haloes) but suffer from poor statistics, while large-scale simulations provide a larger statistics but at the price of not capturing some fundamental physical processes, that are crucial to understand early BH seeding (see e.g. Habouzit et al. 2016, for a discussion).

However, the use of an analytic merger tree comes at the price of lacking the information on the halo spatial distribution. In Section 4, we will discuss the consequences of this limitation comparing our results with independent studies.

2.2 Evolution of progenitor galaxies

The evolution of each progenitor galaxy along the merger tree, is determined by star formation, BH growth and stellar+AGN feedback processes.

The star formation history (SFH) of galaxies along the merger tree is described as a series of quiescent episodes of star formation and/or major merger-induced bursts[2].

The nuclear BHs can grow via both mergers with other BHs and accretion of gas in an Eddington-limited regime (i.e. the gas accretion rate does not exceed the Eddington value). In major mergers, pre-existing BHs are assumed to coalesce, following their host galaxies. Conversely, in minor mergers only one of the two BH is assumed to settle in the centre of the galaxy. The less massive BH of the pair is left as a satellite (see e.g. Tanaka & Haiman 2009) and we do not follow its evolution. Gas accretion on to the central BH is described by the Bondi–Hoyle–Lyttleton (BHL) rate with an additional (multiplicative) free parameter, $\alpha_{\rm BH}$ usually adopted in both semi-analytic models and numerical simulations to account for the increased density in the inner regions around the BH (Di Matteo, Springel & Hernquist 2005)[3]. The different values adopted for

---

[1] Halos with virial temperature in the range $1200 \leq T_{\rm vir} < 10^4$ K.

[2] We define as major the mergers among haloes with DM mass ratio $\mu > 1/4$, where μ is the ratio of the less massive over the most massive halo of the system.

[3] In GAMETE/QSODUST, we assume that the gas mass of the galaxy is distributed within the virial radius of the DM halo following an isothermal sphere profile with a flat core. Under this assumption, the gas density at the Bondi radius is underestimated. In addition, the BHL accretion rate strongly depends on the sound speed of the local material, but we do not track the temperature of the cold gas, nor that of the accreting material around the BH. The BH accretion efficiency α enables us to overcome these limitations (see e.g. Booth & Schaye 2009, for a discussion).

this parameter are given in Valiante et al. (2011) and Valiante et al. (2014). In V16, a value $\alpha_{BH} = 50$ enables us to match the observed BH mass and estimated accretion rate of SDSS J1148+5251.

The progressive pollution of the ISM with metals and dust, produced by asymptotic giant branch (AGB) stars and supernovae (SNe), is computed consistently with the stellar evolutionary time-scales (lifetimes; see Valiante et al. 2014, for details.). In addition, we follow dust enrichment in a two-phase ISM (de Bennassuti et al. 2014; Valiante et al. 2014): Dust grains residing in the hot, diffuse gas can be destroyed by expanding SN shocks, while in the cold, dense clouds dust can grow in mass by accretion of the available gas-phase elements on to the grains surface.

The UV radiation from both stars and AGN is computed according to the emissivity properties of each source. For Pop III stars, we adopt the mass-dependent emissivities presented by Schaerer (2002) and compute the age-metallicity-dependent emissivities for Pop II stars using the model by Bruzual & Charlot (2003). Following Volonteri & Gnedin (2009), the emissivity of accreting BHs is inferred by modelling their spectral energy distributions (SEDs) with a multicolour disc spectrum plus a non-thermal power-law component (see V16 for details).

We compute the LW background within a comoving volume of $50 \, \text{Mpc}^3$ (our biased, high-density region at $z > 6$), following Haardt & Madau (1996) in the dark screen approximation and accounting for intergalactic absorption using the modulation factor given by Ahn et al. (2009). As discussed in Valiante et al. (2017), despite the lack of spatial information, in the highly biased volume that we simulate, the flux level that we predict for the LW background is comparable to the maximum local flux found by Agarwal et al. (2012). We will further discuss this point in Section 4.

The effect of stellar and BH mechanical feedback is included in the form of energy-driven winds, which are able to trigger galaxy-scale gas outflows (including metals and dust) polluting the IGM. We assume that a fraction $\epsilon_{w, SN} = 2 \times 10^{-3}$ and $\epsilon_{w, AGN} = 2.5 \times 10^{-3}$ of the energy released by SN explosions and BH accretion process, respectively, is responsible for launching gas outflows (see V16 for details).

Although they have similar wind efficiencies, in Valiante et al. (2012) we show that the SN-driven winds have a minor effect on the BH-host galaxy co-evolution history, which is instead regulated by the AGN feedback, as observed in SDSS J1148+5251 (Maiolino et al. 2012; Cicone et al. 2015).

2.3 Forming the first stars and BHS

As discussed in V16, the combined levels of chemical enrichment (*in situ* and/or infall driven) and of the UV flux to which the host galaxy is exposed, regulate star formation and set the conditions for DCBHs formation.

In GAMETE/QSODUST, we describe both *in situ* and external (or *ex situ*) pollution. The first dominates at high redshift and the second progressively becomes important and ultimately leads to a termination of Pop III star formation at $z \sim 16$ (on average, see Section 4 for a discussion).

Whether stars can form in minihaloes and what is the efficiency of star formation is dictated by the halo properties (virial temperature, redshift, and ISM metallicity) and by the photodissociating radiation illuminating the halo (see Appendix A in V16 for details).

Radiative feedback also determines the minimum mass of star-forming haloes. This threshold mass increases with redshift, first as a consequence of H_2 photodissociation (e.g. Omukai 2001; Machacek, Bryan & Abel 2001, and Appendix A in V16), then,

at lower redshifts, due to ionization which can lead to photoevaporation of less massive haloes (e.g. Barkana & Loeb 1999, Shapiro, Iliev & Raga 2004, Sobacchi & Mesinger 2013).

In addition, chemical feedback determines the type of stellar population that can form in a galaxy. We assume that Pop III stars in the mass range $[10–300] \, M_\odot$ form in metal poor haloes, as long as the ISM metallicity is lower than the critical threshold $Z_{cr} \sim 10^{-4} \, Z_\odot$ (Schneider et al. 2002, 2003, 2012). Conversely, second-generation stars (Pop II), with a mass spectrum shifted towards less massive stars, $[0.1–100] \, M_\odot$, form out of gas enriched above this critical value. The stellar initial mass function (IMF) is assumed to have the Larson functional form (Larson 1998): $\Phi(m) \propto m^{\alpha-1} e^{-m_{ch}/m}$, where $\alpha = -1.35$ and $m_{ch} = 20 \, (0.35)$ is the characteristic mass for the Pop III top-heavy (Pop II standard) IMF(de Bennassuti et al. 2014).

In low-efficiency star forming haloes, we stochastically sample the Pop III IMF. Following each star formation episode, we randomly select stars in the $[10–300] \, M_\odot$ mass range until we reach the total stellar mass formed in the burst. As a result only few less massive Pop III stars can form in mini-haloes, while the intrinsic top-heavy IMF is fully sampled if a total stellar mass larger than $10^6 \, M_\odot$ is produced. This approach enables us to avoid overestimating the contribution of Pop III stars to radiation emission and chemical enrichment.

2.3.1 Seeding prescription

Following this stochastic procedure to sample the Pop III IMF, we form light seeds when we extract massive stars in the range $[40–140] \, M_\odot$ and $[260–300] \, M_\odot$. We assume that only the most massive among BH remnants formed in each burst settles in the galaxy centre where it can accrete gas and grow.

We compute the time evolution of the cumulative LW emission, J_{LW}, coming from all the emitting sources (the progenitor galaxies and BHs of SDSS J1148+5251) along the merger trees (see section 2.2 in V16). We then assume that a DCBH can form when $J_{LW} > J_{cr}$, where $J_{cr} = 300 \times 10^{-21} \, \text{erg s}^{-1} \, \text{Hz}^{-1} \, \text{cm}^2 \, \text{sr}^{-1}$, in our reference model[4]. This critical level is exceeded at $z \sim 16–18$ (depending on the particular merger tree) so that a massive seed of $10^5 \, M_\odot$ is assumed to form in poorly enriched ($Z_{ISM} < Z_{cr}$) Lyα cooling haloes.

In situ or external metal pollution determines the Pop III to Pop II transition epoch, $z \sim 15$. Below this redshift, both light and heavy seeds no longer form.

3 RESULTS

Here, we investigate the statistical properties of BH seeds and their host galaxies. In V16, we discussed the evolution and properties of BH progenitors of a $z \sim 6$ SMBH, i.e. excluding satellite BHs emerging from minor galaxy merger events. The analysis presented here considers the entire high-z BH population that we find in the simulations, independently of whether they will contribute to the formation of the SMBH mass at $z \sim 6$.

---

[4] The value of this critical threshold ranges from few tens up to $10^3–10^4$ times $10^{-21} \, \text{erg s}^{-1} \, \text{Hz}^{-1} \, \text{cm}^2 \, \text{sr}^{-1}$ depending on the model or simulation adopted (see e.g. Shang, Bryan & Haiman 2010; Wolcott-Green, Haiman & Bryan 2011; Latif et al. 2014b, 2014a; Regan, Johansson & Wise 2014; Hartwig et al. 2015). We refer the reader to the recent review by Valiante et al. (2017) for a detailed discussion.
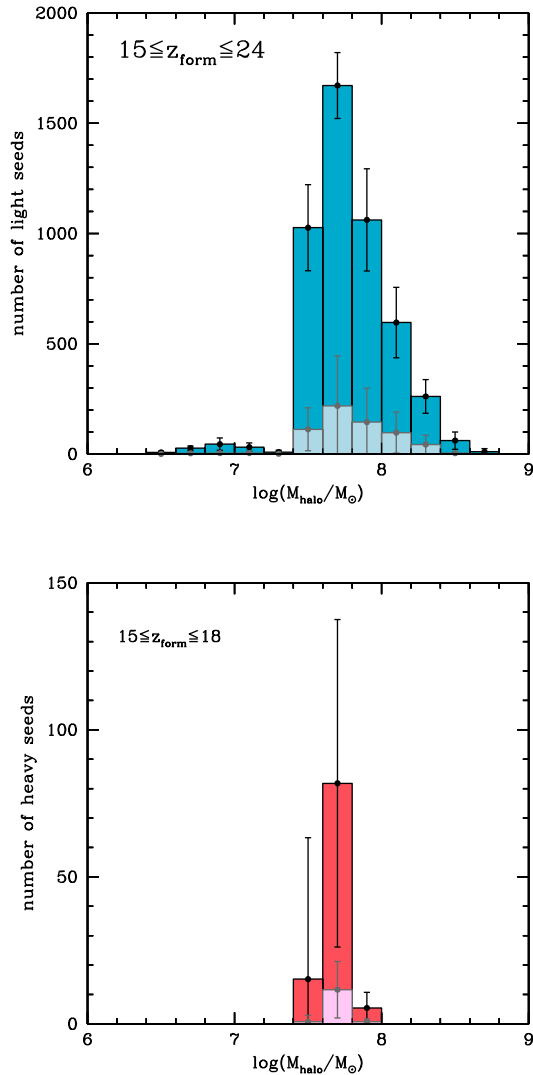
Figure 1. Mass distribution of DM haloes hosting light (upper panel) and heavy (lower panel) BH seeds. The total number of BH seeds is shown with darker colours. Lighter histograms show the subsample of seeds that directly classify as SMBH progenitors (see V16 for details). The average range of formation redshifts is also reported in each panel.

In what follows, histograms and data points will show quantities obtained by averaging over 10 different merger tree realizations, with 1σ errorbars, unless otherwise stated.

3.1 Properties of BH seeds and their hosts

Fig. 1 shows the mass distribution of DM haloes hosting BH seeds formed along the merger history of the halo hosting SDSS J1148+5251. On average, light and heavy seeds are predicted to form in haloes with comparable mass and at comparable redshifts. The average number of light BH seeds forming in mini-haloes with $[10^6$–$10^7]\,M_\odot$ (at $z > 20$, see V16), is small. This is a consequence of the LW feedback, which strongly limits the efficiency of star formation in these haloes. Hence, the bulk of light seeds form at redshift $15 < z_{\rm form} < 20$ in metal-poor ($Z \leq Z_{\rm cr}$) Lyα-cooling haloes. On the other hand, the combined effect of chemical and radiative feedback enables the formation of heavy seeds (bottom panel) only in a relatively small number of haloes, with masses in the range $[5 \times 10^7$–$10^8]\,M_\odot$ and at redshifts in the range $15 < z_{\rm form} < 18$.

3.2 BH seed hosts evolving in 'isolation'

In what follows, we analyse the properties of BHs as long as their host galaxies evolve in 'isolation', namely from the epoch of the BH seed formation down to the redshift of the first merger with a companion galaxy. During this phase, the galaxies interact only with the external medium, via gas infalls and outflows, without experiencing any major or minor merger. We will refer this class of objects as *isolated seeds* hosts. The aim of this analysis is to bring out the distinguishing features of light and heavy seeds, which may help us discriminating their observational signatures.

Fig. 2 shows the distribution of isolated light and heavy seeds (hereafter ILS and IHS, respectively) as a function of (rest-frame) time. Labels in the two figures indicate the total number of galaxies hosting IHS and ILS, averaged over 10 different merger trees.[5] These represent 80(98) per cent of light (heavy) seed hosts at $z > 10$.

The majority of IHS (ILS) hosts merge with another galaxy within the first $\sim$50 Myr, and only a small fraction of systems, 14(7) per cent, remains isolated for more than 150 Myr, as indicated by the tail of the distributions extending to $\Delta t \sim 250(400)$ Myr.

During their evolution, ILS and IHS maintain different BH masses and ILS are less massive than IHS. This is shown in Fig. 3 where the histograms represent the mass distribution of ILS and IHS at $z = 16.5$, 15, 13, and 10. The redshift range has been chosen in order to encompass a period of time during which the evolution of both IHS and ILS can be compared. At $z = 16.5$, we find newly planted DCBHs together with a number of evolving DCBHs formed at earlier times, so that we can capture the host galaxies properties in both phases. Below $z = 10$, it is increasingly difficult to consider light and heavy seeds as two independent populations, due to mergers along their cosmological evolution (see Section 3.3).

The difference in the mass spectrum of the two populations grows with time. In fact, at $z < 15$ when the mass growth is mostly driven by gas accretion, the BHs grow $\propto M_{\rm BH}^2$ when gas accretion occurs at the Bondi rate, and $\propto M_{\rm BH}$ when it occurs at the Eddington rate. Hence, by $z \sim 10$ the mass of IHS is $[10^7$–$10^8]\,M_\odot$, i.e. about two orders of magnitude larger than the upper mass limit of ILS at the same redshift. The statistics also include a small fraction ($\sim$1 per cent) of ILS, which grow almost as fast as their most massive counterparts, reaching a mass $>10^4\,M_\odot$, up to $\sim$2 $\times$ $10^5\,M_\odot$ at redshift 12–13 (see the last 2 mass bins in the third upper panel, from the left-hand side). Although they are not representative of the average ILS population, these BHs are interesting objects for our study, as their emissivity can directly compete with that of IHS (Valiante et al., in preparation).

In Fig. 4, we show the distribution of DM halo mass (panel A), stellar mass (panel B), ISM metallicity (panel C), star formation rate (panel D), BH accretion rate (BHAR; panel E), and Eddington ratio $\lambda_{\rm Edd}$ (panel F) at the same redshifts as in Fig. 3. Vertical dashed lines in all panels show average values. We discuss the properties of ILS (upper panels) and IHS (lower panels) in the following.

---

[5] In an isolated halo a heavy BH seed grows only via gas accretion, while a light seed can also merge with other Pop III BHs that may form in subsequent starbursts occurring in the same halo (see V16 for details). For the purpose of the present analysis, we consider as the effective light seed only the first Pop III BH formed in the halo, when counting the number of ILS and defining the duration of the isolated phase, Δt.
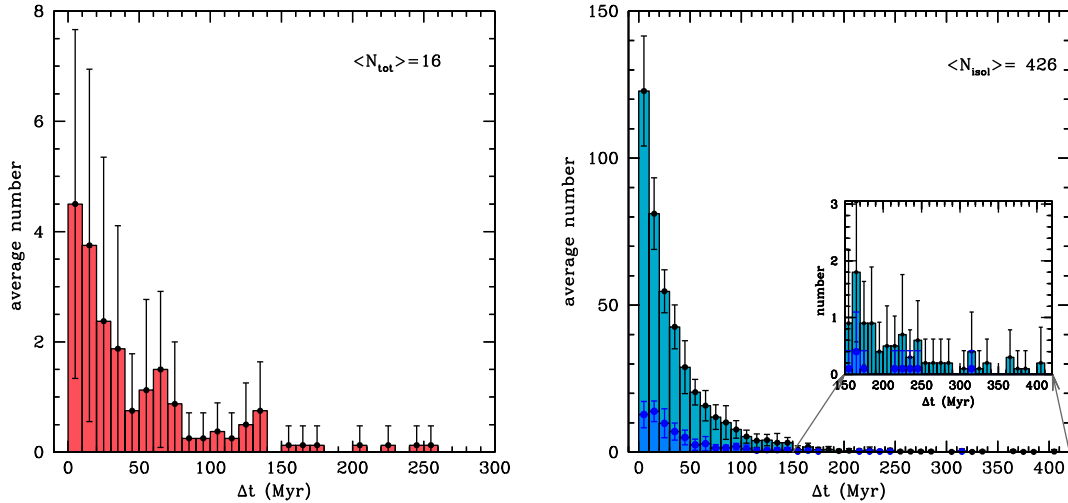
Figure 2. Distribution of IHS (left-hand panel) and LHS (right-hand panel) as a function of (rest-frame) time. The labels indicate the average number of IHS and ILS formed in 10 different merger trees. In the right-hand panel, the superimposed bluer histogram shows the number of ILS with masses $>10^3 \, M_\odot$ and the small inserted plot enlightens the long-living seed tail, at $\Delta t > 150 \, \mathrm{Myr}$.
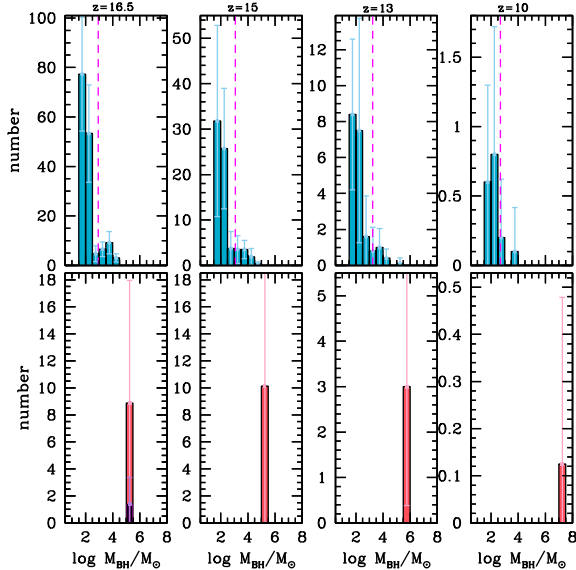
Figure 3. Mass distribution of accreting light (upper panels) and heavy (lower panels) BH seeds at four different redshifts: $z = 16.5, 15, 13,$ and 10. The dashed lines in the upper panels indicate the average mass of ILS at the corresponding z.

3.2.1 DM halo mass distribution

As it can be seen in panel A, at $z = 16.5$ both ILS and IHS reside in Lyα cooling haloes with mass $\sim 5 \times 10^7 \, M_\odot$, with only a small fraction (< 15 per cent) of ILS being hosted in more massive ones. The violet histogram in the first bottom panel on the left-hand side shows the newly planted heavy seeds.[6] At lower redshifts, the two populations of accreting seeds continue to reside in small DM haloes, with very similar average masses. In particular, at $z = 10$ all seeds are found to lie in the same $(1–5) \times 10^9 \, M_\odot$ DM halo mass bin.

3.2.2 Stellar mass distribution

In terms of stellar mass (panel B), ILS host galaxies are, on average, more massive than the ones hosting IHS. The average stellar masses differ by a factor of ~ 5–10 at all redshifts. Such a delayed growth of the stellar component in IHS hosts reflects the environmental conditions required to form the seeds as well as the more efficient BH feedback effect. In ILS host galaxies, star formation starts at redshift $z > 20$, with Pop III stars and their remnant BHs forming down to $z \sim 15$. On the other hand, heavy seeds begin to form at $z \sim 18$ in Lyα cooling haloes that are the descendants of sterile mini-haloes, where star formation has been suppressed by radiative feedback (see V16 for a detailed discussion). Following the formation of the DCBH, BH accretion and feedback prevent efficient star formation to occur, as a large fraction of the available gas mass is either accreted on to the BH or ejected by the BH-driven wind. This is reflected in the higher BHAR versus star formation rate depicted in panels D and E.
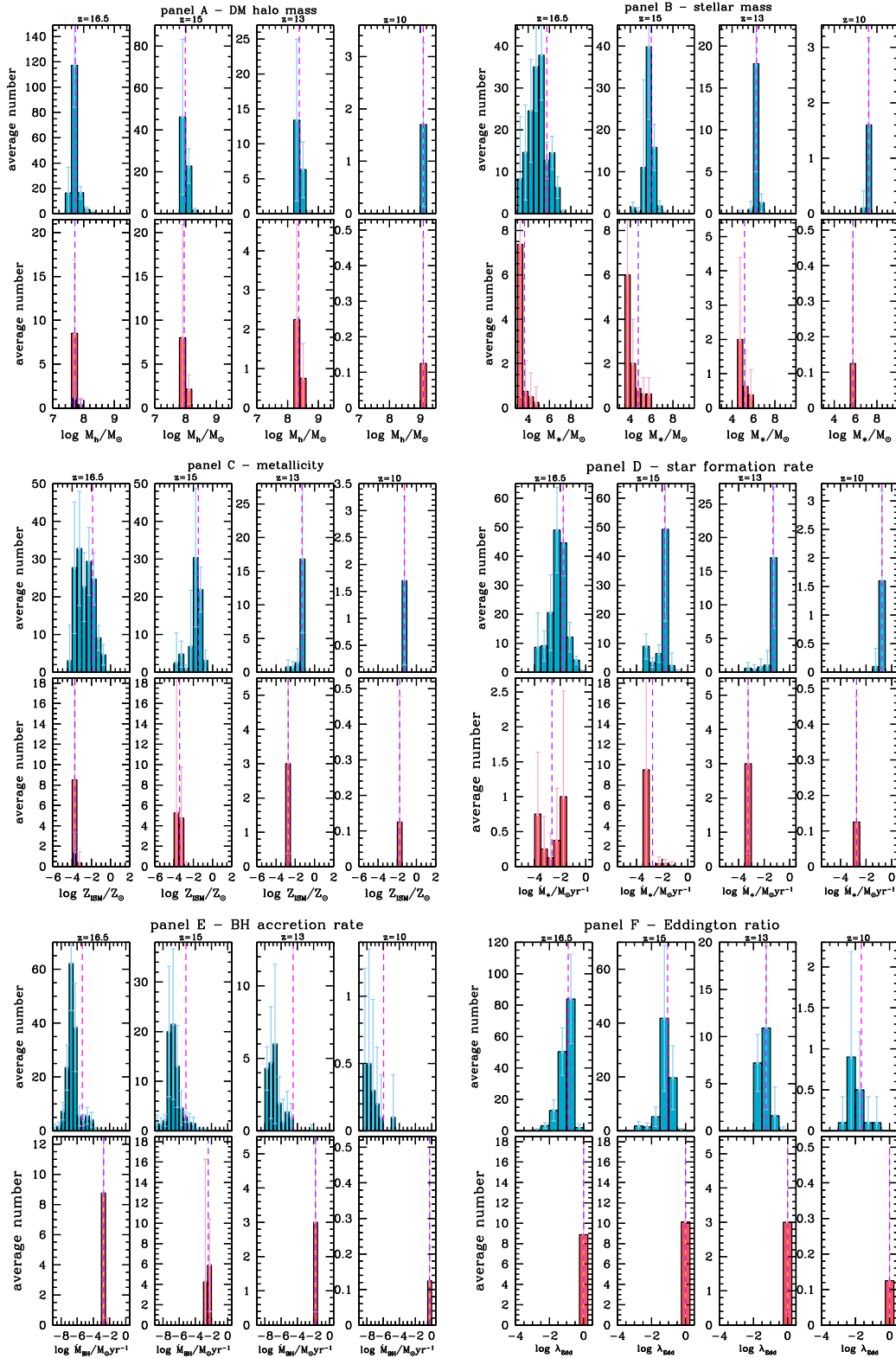
3.2.3 ISM metallicity distribution

For the same reasons as discussed above, the ISM metallicity of IHS hosts is always lower (by about the same factor as the stellar mass) than the metallicity found in the hosts of ILSs, as shown in panel C.

3.2.4 Star formation rate

At $z = 16.5$ only ~ 30 per cent IHS hosts are star forming (see panel D). The remaining ~ 70 per cent do not form stars as a consequence of the strong LW flux to which the haloes are exposed. They are either catched at DCBH formation, or in the subsequent stages, before metal enrichment enables gas cooling and Pop II star formation, overcoming the effects of radiative feedback.[7] At $z \leq 15$ efficient BH accretion and feedback (see also panels E and F) strongly limit star formation in IHS hosts, enlarging the separation between the average star formation rate of the two populations. As a result, at

---

[6] In our model, we stop planting heavy seeds already at redshift $z \sim 15.5$, on average (see V16), as a consequence of the fast chemical enrichment.

[7] It is worth noting that enrichment of these systems is a consequence of metal-rich (and dust-rich) gas accretion.

Figure 4. The distribution of DM halo mass (panel A), stellar mass (panel B), ISM metallicity (panel C), SFR (panel D), BHAR (panel E), and Eddington ratio $\lambda_{\rm Edd}$ (panel F) at $z = 16.5, 15, 13$, and 10. Upper, blue histograms show ILS, while lower, red, histograms show IHS. Vertical dashed lines in all panels show average values.
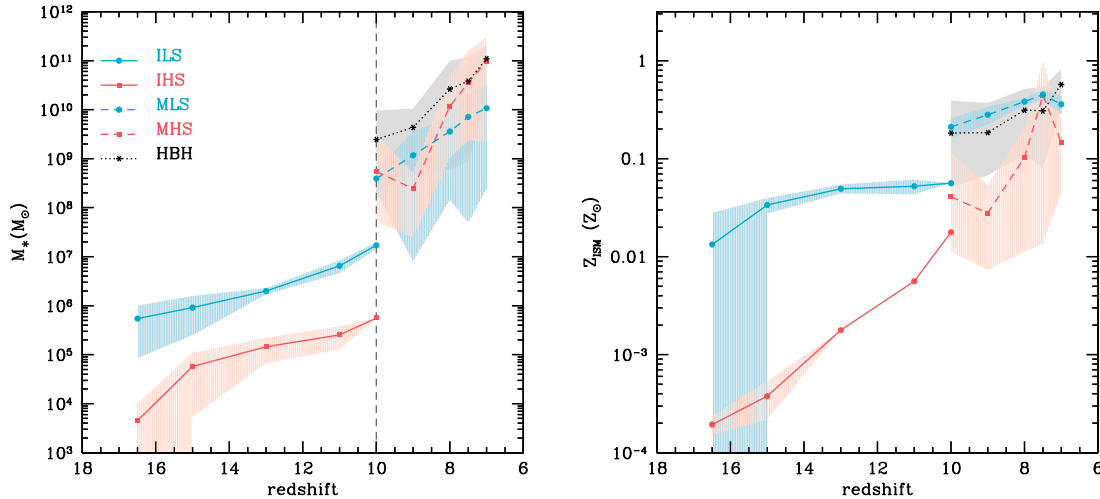
Figure 5. The redshift evolution of the average stellar mass (left-hand panel) and metallicity (right-hand panel) of galaxies hosting ILS (blue circles with solid line), IHS (red squares with solid line), MLS (blue circles with dashed line), MHS (red squares with dashed line), and HBH (black stars with dotted line). The shaded regions show the 1σ dispersion, while the vertical dashed line indicates $z = 10$, below which only a minor fraction of IHS (ILS) hosts evolve in isolation. See the text for details.

$z = 10$ the star formation rate of IHS hosts is about two orders of magnitude lower than in ILS hosts.

3.2.5 BH accretion rate and Eddington ratio

The last two lower panels of Fig. 4 show the distribution of BHAR, $\dot{M}_{\rm BH}$, and Eddington ratios, $\lambda_{\rm Edd}$. The average BHAR of heavy seeds is always one order of magnitude (or more) larger than that of light seeds (panel E) and heavy seeds grow close to or at the Eddington rate, with $\lambda_{\rm Edd} \sim 1$ (panel F).

We conclude that at $z \geq 10$ galaxies that host heavy BH seeds have smaller stellar mass and metallicity and form stars at a smaller rate than galaxies that hosts light BH seeds. In addition, heavy BH seeds grow at a higher rate.

3.3 BH seed evolution at $z < 10$

At $z \leq 10$ merger events with normal, star-forming galaxies (i.e. not hosting nuclear BHs) or with other AGN progressively erase the differences in the properties discussed above, so that any trace of the BH origin is lost. This appears clear in Fig. 5, where we show the average stellar mass and metallicity as a function of redshift. ILS and IHS hosts at $z \geq 10$ are shown by data points connected by solid lines. At $z < 10$ less than 2(20) per cent of IHS (ILS) hosts are found to evolve in isolation. Post-merger systems are divided into three different classes, according to the nature of the merging pairs: *Merged heavy (light)* seed hosts, hereafter MHS (MLS), are hosted in haloes assembled via the coalescence of an IHS (ILS) host with either a normal galaxy with no nuclear BH or with another IHS (ILS) host; *hybrid BHs* (HBH) hosts are formed from the coalescence of IHS and ILS hosts. The figure shows that at $z < 10$ differences between these three classes of systems become statistically insignificant, with the only exception of MHS metallicity, which remains 1 dex smaller than that of MLS and HBH down to $z \sim 8$.

4 DISCUSSION

As discussed in Section 2.1, we employed a Monte Carlo algorithm based on the EPS theory, with the aim of simulating the build-up

of one of the first SMBHs at $z > 6$ starting from Pop III forming minihaloes (see V16 for details). The main limitation of our Monte Carlo approach is the lack of spatial information. This implies that metal enrichment and UV emission from neighbouring haloes, which affect the transition from Pop III to Pop II stars and the number density of DCBHs (see Agarwal et al. 2012; Visbal et al. 2014; Habouzit et al. 2016; Regan et al. 2017, and references therein), can be modelled only in an average way.

To support the robustness of our results, in this section we compare them with hydrodynamical simulations and with hybrid models run on *N*-body codes (see e.g. Valiante et al. 2017, for an extensive discussion).

A few preliminary considerations are necessary. Large-scale numerical simulations are often designed to describe 'average' regions of the Universe, while our model is aimed at simulating a highly biased region, within the virial radius of a $10^{13}\,{\rm M}_{\odot}$ DM halo at $z \sim 6.4$. Small-scale simulations (box sizes up to few Mpc) are able to capture the scales and typical halo masses at which physical processes are important for early BH seeding or Pop III star formation (such as inhomogeneous metal enrichment and the clustering properties of radiation sources). However, their predicted seed BHs are not directly connected with the assembly of the first SMBHs at $z > 6$ (see e.g. Agarwal et al. 2012, 2014; Wise et al. 2012; Crosby et al. 2013; Griffen et al. 2016; Habouzit et al. 2016; Habouzit, Volonteri & Dubois 2017). These small-scale simulations produce enough seeds to match the number density of SMBHs at $z = 6$. It thus becomes a matter of how many of these seeds are able to form in galaxy progenitors of SMBH hosts at $z = 6$, and how fast these seed BHs are able to grow. On the opposite side, simulations with sufficiently large volumes to capture the formation of more massive DM haloes (and SMBHs at $z \sim 6$) do not have the mass resolution required to follow the formation of the first stars in mini-haloes (e.g. Khandai et al. 2012; Feng et al. 2014; Di Matteo et al. 2016; Rong et al. 2017; Simha & Cole 2017). Even the most recent, high-resolution, zoom-in hydrodynamical simulations, which follow the growth of SMBHs at $z \sim 6$ (e.g. Sijacki, Springel & Haehnelt 2009; Barai et al. 2018), do not have the adequate physical prescriptions required to describe the role of Pop III star formation and of different BH seeds in the final SMBH assembly.

The effects of metal enrichment and chemical feedback are reflected both in the Pop III star formation rate density (SFRD) and in the mass distribution of Pop III star forming haloes. With the above caveat in mind, we have compared our results with the recent ENZO Renaissance simulation presented by Xu, Wise & Norman (2013) and Xu et al. (2016a,b) ('rare peak' case) and with the GADGET-2 simulation by Chon et al. (2016). We predict a Pop III SFRD that is about 40–100 times higher than that in the 'rare peak' simulation by Xu et al. (2013, 2016b), and closer to the level obtained by Chon et al. (2016, ; assuming $h = 0.73$, see their fig. 2). If compared with the 'rare peak' simulation presented by Xu et al. (2013, 2016b), the higher SFRD that we find in our model reflects the conditions in the highly biased region that we are considering. If we estimate the overdensity of our comoving volume ($50\,\mathrm{Mpc}^3$) following the definition of Xu et al. (2016a), $\langle\delta\rangle = \langle\rho\rangle/(\Omega_M\rho_{cr}) - 1$, we find $\langle\delta\rangle \sim 5$, a factor of ~ 7 higher than that of the 'rare peak' simulation ($\langle\delta\rangle \sim 0.68$, see e.g. Xu et al. 2016a).

In addition, we find that Pop III stars are preferentially hosted in minihaloes [$6 < \log(M_{halo}/M_\odot) \gtrsim 7.5$] in the redshift range $20 \leq z < 24$ and in Ly α cooling systems [$7.5 < \log(M_{halo}/M_\odot) < 8.5$] in the redshift range $15 < z < 20$. The recent Renaissance simulations presented by Xu et al. (2013, 2016a,b) show a similar mass range for haloes hosting Pop III stars, independently of the density of the simulated region.[8] Moreover, the fraction of haloes containing Pop III stars drops to zero at halo masses $< 7 \times 10^6(10^7)\,\mathrm{M}_\odot$ in their normal (high-density) regions (O'Shea et al. 2015).

In much smaller boxes, simulations show that DM haloes with masses $< 2 \times 10^6\,\mathrm{M}_\odot$ do not host Pop III stars (see e.g. Wise et al. 2012). We find a similar lower limit in V16. In addition, Wise et al. (2012) suggest that a single pair-instability supernova is sufficient to enrich its host halo to a metallicity level as high as $10^{-3}\,\mathrm{Z}_\odot$, driving the transition to a second generation of stars in the halo already at $z > 15$, similarly to what we find in our biased region.

It is worth noting that, while we find similar trends in the typical mass range of Pop III star-forming haloes, we cannot compare the absolute number of haloes per mas bin predicted by our model with that obtained by Xu et al. (2013, 2016a). As it appears evident by comparing the results of the two ENZO simulations presented by those authors, a different number of haloes (total and hosting Pop III stars) is predicted in different simulated regions (high versus low density), even in boxes with comparable volumes.

For what concerns the effect of radiative feedback (LW radiation) on DCBHs formation, we have compared the evolution of the LW flux at which haloes are exposed to with the results found by studies based on hybrid models run on N-body simulations in Valiante et al. (2017, see Fig. 3). We found that in the highly biased volume that we simulate, the flux level that we predict for the background is comparable to the maximum local flux level found by Agarwal et al. (2012). To investigate the impact of the lack of spatial information on our DCBH seeding prescription, we can derive the maximum local J_{LW} from our Pop III and Pop II galaxies at fixed (physical) distances. For an escape fraction of 1, we find that the J_{LW} fluctuations are less important than the global background for

distance scales $\geq 7\,\mathrm{kpc}$. Conversely, the local J_{LW} could be significantly (at least an order of magnitude) higher than our global LW background within $\sim 1\,\mathrm{kpc}$ from the emitting source. Chon et al. (2016) find that 2 out of 42 DC host halo candidates (i.e. 5 per cent) in their simulations successfully form a DCBH when illuminated by a close-by source within few kpc (1.27 and $6.29\,h^{-1}\,\mathrm{kpc}$). We find a similar (average) fraction of atomic cooling haloes hosting heavy seeds.

Finally, we recall that although there is a general consensus between models on the fundamental role of the LW radiation and of metal pollution (both *in situ* and *ex situ*) in setting the environment where DCBHs form, the value of the critical LW radiation intensity required is still highly uncertain, and there is a large spread in the number density of DCBHs (or DCBH host candidates) derived in different models (see e.g. Agarwal et al. 2012, 2014; Dijkstra, Ferrara & Mesinger 2014; Habouzit et al. 2016). Yet, with a critical LW threshold $J_{cr} = 300$, we predict an average DCBH occurrence ratio of 5 per cent, in agreement with recent numerical simulations by Chon et al. (2016).

In light of the discussion presented above, we believe that our approach is well motivated for the purpose of our investigation, and that it leads to robust results when compared with independent studies based on different techniques.

5 CONCLUSIONS

In this work, we have analysed the evolution of heavy and light BH seeds, progenitors of a $z > 6$ SMBH, and their host galaxies in a cosmological context, as predicted in V16. The aim of this study is to characterize the properties of the birth/growth environment of these two different BH seeds.

Our main findings are summarized as follows:

(i) On average, light and heavy BH seeds form within a similar redshift range, $15 < z_{form} < 18$, in DM haloes having comparable average mass, $\sim 5 \times 10^7\,\mathrm{M}_\odot$. Even after their formation, they continue to reside in haloes with similar masses (10^8–$10^9\,\mathrm{M}_\odot$), as long as the systems evolve in isolation.

(ii) About 80(98) per cent of light (heavy) seed hosts are found to evolve in isolation (i.e. no minor or major mergers with other haloes) down to $z \sim 10$. Only a small number of galaxies hosting growing light (heavy) BH seeds remain isolated for longer than 200(400) Myr.

(iii) At $z \geq 10$, galaxies hosting heavy BH seeds are characterized by a factor of 5–10 smaller stellar mass and metallicity than the ones hosting light seeds. In addition, heavy BH seeds are accreting gas more efficiently (Eddington ratio close to 1) than their lighter counterparts. As a result of the efficient growth and feedback, the host galaxies of heavy BH seeds experience a less intense star formation activity. Their star formation rate is about two orders of magnitude lower than in galaxies that host light BH seeds.

(iv) At $z < 10$, the fraction of isolated systems dramatically decreases to less than 2(20) per cent, for heavy (light) seed hosts, as merger events occurs. The differences in the galaxy properties of different systems are progressively erased so that any trace of the BH origin is lost.

We conclude that the probability to disentangle the origin of BH progenitors requires to target these systems at $z > 10$, when their own properties and the properties of their host galaxies still reflect/trace the conditions at BH seed formation.

The properties inspected here have a fundamental role in shaping the SED of accreting BHs and galaxies. Indeed, the results of the

---

[8] The Renaissance simulation has targeted three different subvolumes, with comoving sizes of 133.6, 220.5, and $220.5\,\mathrm{Mpc}^3$, designated as 'Rare peak', 'Normal', and 'Void' regions because they have a mean density that is higher, comparable, and lower than the cosmic average, respectively (e.g. Xu et al. 2013; O'Shea et al. 2015; Xu et al. 2016a). In all cases, haloes with mass up to few $10^9\,\mathrm{M}_\odot$ are formed.

statistical analysis presented here will be used in a companion paper (Valiante et al., in preparation) to characterize the luminosity and colours of BH progenitors and explore the prospects of discriminating the origin of BH seeds.

REFERENCES

Agarwal B., Khochfar S., Johnson J. L., Neistein E., Dalla Vecchia C., Livio M., 2012, MNRAS, 425, 2854

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

Ahn K., Shapiro P. R., Iliev I. T., Mellema G., Pen U.-L., 2009, ApJ, 695, 1430

Alexander T., Natarajan P., 2014, Science, 345, 1330

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

Barai P., Gallerani S., Pallottini A., Ferrara A., Marconi A., Cicone C., Maiolino R., Carniani S., 2018, MNRAS, 473, 4003

Barkana R., Loeb A., 1999, ApJ, 523, 54

Barth A. J., Martini P., Nelson C. H., Ho L. C., 2003, ApJ, 594, L95

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

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

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

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

Cicone C. et al., 2015, A&A, 574, A14

Crosby B. D., O'Shea B. W., Smith B. D., Turk M. J., Hahn O., 2013, ApJ, 773, 108

Davies M. B., Miller M. C., Bellovary J. M., 2011, ApJ, 740, L42

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

de Rosa G. et al., 2011, ApJ, 739, 56

de Rosa G. et al., 2014, ApJ, 790, 145

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

Devecchi B., Volonteri M., Colpi M., Haardt F., 2010, MNRAS, 409, 1057

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., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2016, MNRAS, 467, 4243

Dijkstra M., Ferrara A., Mesinger A., 2014, MNRAS, 442, 2036

Fan X. et al., 2001, AJ, 122, 2833

Feng Y., Di Matteo T., Croft R., Khandai N., 2014, MNRAS, 440, 1865

Griffen B. F., Dooley G. A., Ji A. P., O'Shea B. W., Gómez F. A., Frebel A., 2016, MNRAS, preprint (arXiv:1611.00759)

Haardt F., Madau P., 1996, ApJ, 461, 20

Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529

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

Haiman Z., 2004, ApJ, 613, 36

Hartwig T., Glover S. C. O., Klessen R. S., Latif M. A., Volonteri M., 2015, MNRAS, 452, 1233

Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288

Hirano S., Hosokawa T., Yoshida N., Omukai K., Yorke H. W., 2015, MNRAS, 448, 568

Inayoshi K., Omukai K., 2012, MNRAS, 422, 2539

Inayoshi K., Visbal E., Kashiyama K., 2015, MNRAS, 453, 1692

Jiang L. et al., 2016, ApJ, 833, 222

Johnson J. L., Haardt F., 2016, PASA, 33, e007

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

Khandai N., Feng Y., DeGraf C., Di Matteo T., Croft R. A. C., 2012, MNRAS, 423, 2397

Larson R. B., 1998, MNRAS, 301, 569

Latif M. A., Ferrara A., 2016, PASA, 33, e051

Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J., 2013a, MNRAS, 433, 1607

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

Latif M. A., Bovino S., Van Borm C., Grassi T., Schleicher D. R. G., Spaans M., 2014a, MNRAS, 443, 1979

Latif M. A., Schleicher D. R. G., Bovino S., Grassi T., Spaans M., 2014b, ApJ, 792, 78

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

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

Lupi A., Haardt F., Dotti M., Fiacconi D., Mayer L., Madau P., 2016, MNRAS, 456, 2993

Machacek M. E., Bryan G. L., Abel T., 2001, ApJ, 548, 509

Madau P., Rees M. J., 2001, ApJ, 551, L27

Maiolino R. et al., 2012, MNRAS, 425, L66

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

Natarajan P., 2011, preprint (arXiv:1105.4902)

O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12

Omukai K., 2001, ApJ, 546, 635

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

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

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

Regan J. A., Johansson P. H., Wise J. H., 2014, ApJ, 795, 137

Regan J. A., Visbal E., Wise J. H., Haiman Z., Johansson P. H., Bryan G. L., 2017, Nat. Astron., 1, 0075

Rong Y., Guo Q., Gao L., Liao S., Xie L., Puzia T. H., Sun S., Pan J., 2017, MNRAS, 470, 4231

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

Schaerer D., 2002, A&A, 382, 28

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

Schneider R., Ferrara A., Salvaterra R., Omukai K., Bromm V., 2003, Nature, 422, 869

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

Shang C., Bryan G. L., Haiman Z., 2010, MNRAS, 402, 1249

Shapiro P. R., Iliev I. T., Raga A. C., 2004, MNRAS, 348, 753

Sijacki D., Springel V., Haehnelt M. G., 2009, MNRAS, 400, 100

Simha V., Cole S., 2017, MNRAS, 472, 1392

Sobacchi E., Mesinger A., 2013, MNRAS, 432, 3340

Spaans M., Silk J., 2006, ApJ, 652, 902

Sugimura K., Omukai K., Inoue A. K., 2014, MNRAS, 445, 544

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

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 (V16)

Valiante R., Agarwal B., Habouzit M., Pezzulli E., 2017, PASA, 34, e031

Visbal E., Haiman Z., Bryan G. L., 2014, MNRAS, 445, 1056

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

Volonteri M., Gnedin N. Y., 2009, ApJ, 703, 2113

Volonteri M., Rees M. J., 2005, ApJ, 633, 624

Volonteri M., Rees M. J., 2006, ApJ, 650, 669

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

Volonteri M., Silk J., Dubus G., 2015, ApJ, 804, 148

Willott C. J., McLure R. J., Jarvis M. J., 2003, ApJ, 587, L15
Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
Wolcott-Green J., Haiman Z., Bryan G. L., 2011, MNRAS, 418, 838
Wu X.-B. et al., 2015, Nature, 518, 512
Xu H., Wise J. H., Norman M. L., 2013, ApJ, 773, 83
Xu H., Norman M. L., O'Shea B. W., Wise J. H., 2016a, ApJ, 823, 140

Xu H., Ahn K., Norman M. L., Wise J. H., O'Shea B. W., 2016b, ApJ, 832, L5
Yoshida N., Omukai K., Hernquist L., 2008, Science, 321, 669

This paper has been typeset from a T$_{\rm E}$X/L$^{\rm A}$T$_{\rm E}$X file prepared by the author.