The effect of baryonic streaming motions on the formation of the first supermassive black holes

Takamitsu L. Tanaka1*, Miao Li2 and Zoltán Haiman2
1Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany
2Department of Astronomy, Columbia University, 550 W. 120th Street, New York, NY 10027, USA

11 May 2014

ABSTRACT
Observations of quasars at redshifts $z > 6$ reveal that $10^9 M_\odot$ supermassive black holes (SMBHs) had already formed when the Universe was $\sim 0.9$ Gyr old. One hypothesis for the origins of these SMBHs is that they grew from the remnants of the first generation of massive stars, which formed in low-mass ($\sim 10^5 - 10^6 M_\odot$) dark matter minihaloes at $z > 20$. This is the regime where baryonic streaming motions—the relative velocities of baryons with respect to dark matter in the early Universe—most strongly inhibit star formation by suppressing gas infall and cooling. We investigate the impact of this effect on the growth of the first SMBHs using a suite of high-fidelity, ellipsoidal-collapse Monte Carlo merger-tree simulations. We find that the suppression of seed BH formation by the streaming motions significantly reduces the number density of the most massive BHs at $z > 15$, but the residual effect at lower redshifts is essentially negligible. The streaming motions can reduce by a factor of few the number density of the most luminous quasars at $z \approx 10 - 11$, where such objects could be detected by the James Webb Space Telescope. We conclude, with minor theoretical caveats, that baryonic streaming motions are unlikely to pose a significant additional obstacle to the formation of the observed high-redshift quasar SMBHs. Nor do they appreciably affect the heating and reionization histories of the Universe or the merger rates of nuclear BHs in the mass and redshift ranges of interest for proposed gravitational-wave detectors.

Key words: black hole physics, galaxies: formation, cosmology: theory, gravitational waves, quasars: supermassive black holes

1 INTRODUCTION
Recent surveys such as SDSS, CFHQS and UKIDSS have unveiled about 50 quasars at redshifts $z > 6$, when the Universe is less than 1 Gyr old (Fan et al. 2001, 2006; Lawrence et al. 2007; Willott et al. 2009). The most distant quasar discovered to date, ULAS J1120+0641, has a redshift of $z = 7.1$ ($\approx 0.8$ Gyr after the big bang), with a mass $\sim 2 \times 10^9 M_\odot$ (Mortlock et al. 2011).

It is still a mystery how these supermassive black holes (SMBHs) accumulated so much mass in such a short time (see reviews by Volonteri 2010 and Haiman 2013). One possibility is that they grew from BHs left behind by the first generation of stars (e.g. Madau & Rees 2001, Haiman & Loeb 2001, Bromley et al. 2004, Shapiro 2005, Pelupessy et al. 2007, Tanaka & Haiman 2009, hereafter TH09). These 'Population III' (PopIII) stars are thought to have formed from molecular-cooling gas collapsing inside dark matter (DM) minihaloes at redshifts $z \geq 20$ (e.g. Bromm & Larson 2004). PopIII stars are thought to be massive ($\sim 30 - 300 M_\odot$, Heger et al. 2000; but see Turk et al. 2009, Stacy et al. 2010, Hosokawa et al. 2011, Greif et al. 2011b), metal-free, and short-lived; after their deaths, they would have left behind BHs with $\sim 40$ per cent of their original mass (e.g. Zhang et al. 2008b). These 'seed' BHs can then grow by accreting gas and merging with their peers through hierarchical structure formation. Provided that they can accrete efficiently (but see Milosavljević et al. 2004, Alvarez et al. 2007), PopIII seeds can grow to $\sim 10^9 M_\odot$ by $z \approx 7$ without a super-Eddington phase (e.g. Tanaka et al. 2012, hereafter TPH12).

One of the main uncertainties of the above scenario of SMBH formation is when and where the first stars form. Recent studies of the effect of baryonic streaming motions (BSMs) have put new insights into this question (e.g. Tseliakhovich & Hirata 2010). Before cosmic recombination, the excitation of acoustic oscillations in photon-baryon fluids generates relative bulk velocities between the
gas and the DM. During recombination, the sound speed of baryons drops abruptly, so that the mean relative motion of $\sim 30$ km/s becomes supersonic. This bulk motion makes it easier for gas to stream out of DM haloes and thus affects the distribution and evolution of baryons in the early Universe. [Dalal et al. 2010; Mao et al. 2011; Tseliakhovich et al. 2011; Falckov et al. 2012]. In particular, the streaming velocities suppress the formation of PopIII stars [Greif et al. 2011a; Stacy et al. 2011]. The effect is larger at higher redshifts, when the characteristic DM halo potentials are shallower and the bulk streaming velocities are larger compared to the sound speed of the intergalactic medium (IGM).

The earliest-forming PopIII seeds, which are the most vulnerable to BSMs, are also believed to be the most important ‘building blocks’ in the assembly of the first SMBHs in the Pop-III scenario. This is because of the gravitational recoil effect, in which the asymmetric emission of gravitational waves imparts velocities as large as thousands of km s$^{-1}$ to the merged object, relative to the rest frame of the binary’s centre of mass [Peres 1962; Kidder 1995; Favata et al. 2004; Baker et al. 2006]. Mergers of BHs of similar masses result in higher recoil velocities, and are thus more likely to result in ejection from the host halo [Haiman 2004; Yoo & Miralda-Escudé 2004]. Inversely, a seed that has formed with a greater mass than the rest of the population will suffer smaller recoils following a merger and is less likely to be ejected; a seed that forms earlier than the rest may be doubly protected, because it will have had more time to grow before its first merger and also benefit from residing in a deeper host potential well. The survival bias for the most massive BHs (‘survival of the fittest’ [Volonteri & Rees 2006; cf. TH08]) increases in a runaway fashion as the objects that survive mergers become increasingly more massive with respect to their contemporaries. BSMs could thus affect the assembly of SMBHs via mergers by preferentially suppressing the number density of the earliest seed BHs.

In this work, we investigate how this reduction of the number of seed BHs affects SMBH formation in the early Universe. In particular, we are interested in the following two questions: (1) whether the formation of $10^9$ $M_\odot$ SMBHs at $z \geq 7$ is still possible without super-Eddington accretion; and (2) how the BSMs influence the BH mass function in the redshift range $z = 6 - 11$, which could be probed by the James Webb Space Telescope (JWST). We simulate the assembly history of 70 haloes whose masses $M_{\text{halo}}$ at $z = 6$ exceed $10^{12} M_\odot$, the mean Monte Carlo progenitor mass function agrees with the theoretical expectation to within a few percent out to $z \sim 30$. The discrepancy becomes somewhat larger for higher redshift when there are fewer progenitors, e.g. within $\sim 30$ per cent for $z = 40$. We note that compared to the spherical-collapse model, the ellipsoidal model predicts at $z \sim 40$ a factor $\sim 70$ more haloes with $M > 6 \times 10^8 M_\odot$. Thus, employing the latter model is essential in accurately characterizing the population of the very first seed BHs.

We simulate the assembly history of 70 haloes whose masses $M_{\text{halo}}$ at $z = 6$ exceed $10^{12.85} M_\odot$; these haloes represent a comoving volume of $\sim 150$ Gpc$^3$. Because it is intractable to directly simulate all of the haloes below $10^{12.85} M_\odot$ in such a large volume, we instead construct a statistical representation of these lower-mass haloes by simulating narrow mass bins—with logarithmic widths of $\Delta \log M_{\text{halo}} \lesssim 0.5$ ranging from $\log(M_{\text{halo}}/M_\odot) = 8.12.85$—each represented by a sample of 100 haloes, and then scaling up the counting results for each bin to the expected halo abundances (see section 2.6 and table 1 in TH08 for details). In other words, our simulations combine a statistical approximation of the assembly of haloes with $10^{9} M_\odot < M_{\text{halo}}(z = 6)$.

2 METHODS

2.1 The Merger Tree

Following previous works [Volonteri et al. 2003; Yoo & Miralda-Escudé 2004; Bromley et al. 2004; TH09], we employ semianalytic Monte Carlo merger trees to simulate the hierarchical assembly of DM haloes and the growth of their nuclear BHs. An important aspect of the present work is the adoption of the ellipsoidal collapse model [Sheth & Tormen 2002] of DM in the extended Press-Schechter (EPS) formalism. Previous studies that utilized merger trees have used the spherical collapse model [Lacey & Cole 1993], which is known to underpredict the number of the most massive haloes when compared to N-body simulations, with the discrepancy increasing with look-back time.

However, until recently it was not practically feasible to generate merger trees based on the ellipsoidal EPS model, because the fitting formula for the conditional mass function was inaccurate for small time steps $\Delta z$. Zhang et al. 2008a derived an accurate form of the conditional mass function for $\Delta z \lesssim 0.1$—thus enabling the construction of ellipsoidal merger trees—and developed several algorithms that faithfully reproduce the progenitor mass function. We adopt their ‘method B’, which in a follow-up paper (Zhang et al. in preparation) is found to agree best with $N$-body simulations. We do not detail the mathematical formulae and algorithms here, and instead refer the reader to the above papers for a comprehensive description.

We have confirmed the fidelity of our merger trees by comparing the progenitor mass function with the semianalytic predictions of the ellipsoidal EPS model. For example, in 100 realizations of the assembly history of a $10^{12} M_\odot$ parent halo at $z = 6$, the mean Monte Carlo progenitor mass function agrees with the theoretical expectation to within a few percent out to $z \sim 30$. The discrepancy becomes somewhat larger for higher redshift when there are fewer progenitors, e.g. within $\sim 30$ per cent for $z = 40$. We note that compared to the spherical-collapse model, the ellipsoidal model predicts at $z \sim 40$ a factor $\sim 70$ more haloes when $M > 6 \times 10^8 M_\odot$. Thus, employing the latter model is essential in accurately characterizing the population of the very first seed BHs.

We simulate the assembly history of 70 haloes whose masses $M_{\text{halo}}$ at $z = 6$ exceed $10^{12.85} M_\odot$; these haloes represent a comoving volume of $\sim 150$ Gpc$^3$. Because it is intractable to directly simulate all of the haloes below $10^{12.85} M_\odot$ in such a large volume, we instead construct a statistical representation of these lower-mass haloes by simulating narrow mass bins—with logarithmic widths of $\Delta \log M_{\text{halo}} \lesssim 0.5$ ranging from $\log(M_{\text{halo}}/M_\odot) = 8.12.85$—each represented by a sample of 100 haloes, and then scaling up the counting results for each bin to the expected halo abundances (see section 2.6 and table 1 in TH08 for details). In other words, our simulations combine a statistical approximation of the assembly of haloes with $10^{9} M_\odot < M_{\text{halo}}(z = 6)$.
6) < 10^{12.85} \, M_\odot$, alongside a direct realization of those with $M_{\text{halo}}(z = 0) > 10^{12.85} \, M_\odot$.

### 2.2 Implementation of Baryonic Streaming Motions

BSMs suppress the formation of the first stars by inhibiting the collapse of gas into early DM minihaloes. Several numerical studies have shown that in the absence of BSMs, PopIII stars typically form in haloes with virial temperatures $T_{\text{vir}} \sim 1000$ K (e.g. Bromm et al. 2002 and references therein) or, equivalently, circular velocities $v_{\text{circ}} \sim 4$ km s$^{-1}$. Streaming velocities effectively raise this threshold for PopIII formation; Fialkov et al. (2012) fit the following expression for the ‘cooling’ circular velocity threshold to the simulation results of Stacy et al. (2011) and Greif et al. (2011a):

$$v_{\text{cool}} = \sqrt{\frac{\lambda}{v_0^2} + [\alpha v_{\text{BSM}}(z)]^2},$$

with $v_0 = 3.7$ km/s and $\alpha = 4.0$. The case $v_{\text{BSM}} = 0$ corresponds to $T_{\text{vir}} \approx 960$ K.

Because BSMs are coherent on scales of a few comoving Mpc (Tseliakhovich & Hirata 2010) and the DM haloes in our simulation satisfy ($M_{\text{halo}}/\rho_0)^{1/3} \lesssim$ Mpc, we associate each $z = 6$ parent halo and all of its progenitors with a single streaming velocity value at recombination, $v_{\text{BSM}}$, drawn from a Maxwell-Boltzmann distribution with an rms value $\sigma_{\text{BSM}} \approx 30$ km/s. The streaming velocity subsequently decays with time as $v_{\text{BSM}}(z) \propto (1 + z)$.

We convert the circular velocity threshold in equation (1) to a mass threshold $M_{\text{cool}}(z, v_{\text{BSM}})$ through the relation

$$M_{\text{cool}} \approx 4.5 \times 10^5 \left(\frac{v_{\text{cool}}}{4 \text{ km s}^{-1}}\right)^3 \left(\frac{1 + z}{21}\right)^{-3/2} \, M_\odot.$$

We seed a halo with the BH remnant of a PopIII star if its mass is higher than both $M_{\text{cool}}$ and the cosmological Jeans mass (e.g. Barkana & Loeb 2001):

$$M_j = \frac{4\pi}{3} \rho_0 \left(\frac{\lambda_j}{2}\right)^3,$$

where $\lambda_j = c_s \sqrt{\pi/(G\rho_0)}$ is the Jeans length, $\rho_0(z) = 3H(z)/(8\pi G) \times \Omega(z) (1 + z)$ is the mean matter density and $c_s = \sqrt{5kB_{\text{IGM}}/(3\mu m_p)}$ is the gas sound speed in the IGM. (Here, G, k_B and m_p are the usual physical constants, $H = 100h$ km s$^{-1}$ Mpc$^{-1}$ is the Hubble parameter, $T_{\text{IGM}}$ is the IGM temperature, and $\mu$ its mean molecular weight.)

In the absence of baryonic streaming, $M_j \propto c_s^3(1 + z)^{-3/2}$. However, with BSMs the effective Jeans length scales as $\lambda_j \propto c_s^2 + v_{\text{BSM}}$ (Stacy et al. 2011), so that

$$M_j \approx 5.6 \times 10^5 \left(\frac{c_s^2 + v_{\text{BSM}}}{1 \text{ km s}^{-1}}\right)^3 \left(\frac{1 + z}{21}\right)^{-3/2} \, M_\odot.$$

We compute the IGM temperature and the corresponding Jeans mass self-consistently by calculating the photoheating due to the X-rays produced by accreting BHs (see below). As a control, we have also run simulations with $v_{\text{BSM}}$ set to zero, in which $M_{\text{cool}}$ and $M_j$ have uniform values at any given values of $z$ and $T_{\text{IGM}}$.

### 2.3 Other Baryonic Processes

We adopt several semi-analytic prescriptions to treat relevant baryonic processes that contribute to SMBH growth. Many of the model components listed below are described in greater detail in TPH12, to which we refer the reader for specifics.

**BH seeding and IMF.** The initial mass function (IMF) of the stellar population is highly uncertain; recent simulations suggest that the typical mass maybe few tens of $M_\odot$ (Turk et al. 2009; Stacy et al. 2011; Greif et al. 2011b), but some stars may form with much higher masses (Omukai & Palla 2001; Ohkubo et al. 2009). We adopt a Salpeter IMF with a floor of $20 \, M_\odot$ and slope $d\nu/d\log M_\odot \propto M_\odot^{-1.35}$. The masses of the remnant seed BHs are prescribed using a semi-analytic fit (TPH12) to the simulation results of Zhang et al. (2008a). For the high-mass stars ($M_* > 45 \, M_\odot$) of the greatest interest, the remnant BH mass is $\sim 20 - 40$ per cent of the stellar mass; we assume that stars in the pair-instability mass window $140 \, M_\odot < M_* < 260 \, M_\odot$ leave behind no BH remnant (Heger et al. 2003). We place a single seed BH in each DM halo that satisfies a minimum-mass criterion, as described above. A halo is not seeded if it or any of its progenitors have previously formed a seed BH.

**BH mergers and recoil.** We compute the merger time of a merging DM halo using the fitting function of Boylan-Kolchin et al. (2009), assuming a circularity parameter $\eta = 0.5$. If the merger time is longer than the Hubble time, we assume that the secondary halo ends up ‘stuck’ as a satellite and its BH is for all practical purposes removed from the simulation. If the merger time is shorter than the Hubble time, the two nuclear BHs are assumed to merge on the same time-scale as the host haloes (for the purposes of the merger-tree simulation, instantaneously). A random gravitational recoil velocity is assigned to the merged BH according to the fitting formula of Lousto et al. (2010), under the assumption that components of BH spin are uniformly distributed: $0 < a < 0.93$ for the dimensionless spin magnitude, and $0 < \theta < \pi/6$ for spin vector angles with respect to the inspiral plane (see Bogdanovic et al. 2007; Dotti et al. 2009). The recoiling BH is discarded if the recoil velocity is so great that the BH cannot resettle at the halo centre via dynamical friction within a Hubble time (see TH04 for details).

**BH growth and IGM heating.** We assume a simple growth model in which BHs are assumed to grow exponentially at a mean rate of 2/3 of the Eddington rate, with a time-averaged radiative efficiency $\epsilon = 0.07$—provided that a rich supply of cold gas is available (see paragraph below). We assume that growing BHs emit 90 per cent of their radiation as a multicolour disc (Shakura & Sunyaev 1973) with a greybody spectrum (Blaes 2004; Tanaka & Menou 2010), and the other 10 per cent as a hard X-ray corona above 1 keV with a photon index $\Gamma = 2$. The resulting spectral energy distribution peaks above $\sim 1$ keV for BHs below $10^5 \, M_\odot$. The X-rays photoheat the IGM as they are absorbed. Because they have long ($\gtrsim 1$ Gpc) mean free paths, they do so nearly isotropically—the feedback is global.$^2$

---

$^2$ Lyman-Werner background radiation from star formation can also heat the IGM, but Tanaka et al. (2012) found this contri-
for the history of this ‘global warming’ of the IGM by mini-quasars, accounting for the atomic ionization states of H and He. We refer the reader to TPH12 (cf. Haardt & Madau 1996) for the relevant details of cosmological radiative transfer.

As the IGM is heated, its Jeans mass scale increases, so that gas in low-mass haloes can no longer collapse. We implement this negative feedback by assuming that the BHs can only form and grow if their host halo exceeds the Jeans mass, and that they can only grow for $3 \times 10^7$ yr [comparable to the typical active galactic nucleus (AGN) lifetime; e.g. Martinelli 2004] following the most recent merger of their host with another halo exceeding $M_J$. In this model, the X-ray heating of the IGM by mini-quasars is a self-regulatory feedback, in that the emission of the accreting BHs act to suppress the subsequent formation and growth of BHs in low-mass haloes. In models without this self-regulatory feedback, mini-quasars heat the IGM but the IGM temperature has no bearing on BH formation and growth; seed BHs form in pristine haloes above the cooling mass and are allowed to grow continuously at $2/3$ of the Eddington rate.

Because the gas accreted by the BHs following each merger episode is much less than the fraction of cool gas in the merged halo, we assume that BSMs do not affect BH growth in our simulations beyond raising the halo mass threshold above which mergers can trigger accretion. We discuss caveats to this assumption in §4.

3 RESULTS
3.1 Evolution of the cosmic BH population

BSMs and the negative feedback from IGM heating have the same qualitative effect of reducing the number of BH seeds, but they operate preferentially at different epochs. The former has a greater impact at early times when the relative speeds between baryons and DM are greater, and the latter operates at later times after the X-rays from the earliest growing BHs have sufficiently raised the cosmological Jeans mass. To isolate these two effects, we run four sets of simulations, toggling each effect on and off. We use the same DM haloes and merger histories for each set of simulations. The models without regulation by IGM heating vastly overproduce massive BHs compared to the observational constraints; these are shown merely to illustrate, in the simplest model of a steadily exponentially growing BHs, how the late-time BH population is affected when seed BH production is delayed by BSMs.

We plot the mass evolution of the nuclear BH population in Fig. 1. In all simulations, X-rays from mini-quasars ‘globally warm’ the IGM. The panels on the left-hand side show results from simulations where the rising IGM temperature provides no feedback on the BHs, whereas the panels on the right-hand side show results where the warming provides a negative feedback on the formation and growth of nuclear BHs, as described in the previous section. The results from simulations implementing BSMs are shown in thick lines, and those from simulations without BSMs are shown in thin lines with lighter colours. The solid curves in the upper panels show the global BH density $\rho_{BH}$ for all nuclear BHs as a function of redshift; the dashed curves show the density for only the BHs exceeding $10^7 \, M_\odot$. The lower panels show the rate of change in the universal BH density $\dot{\rho}_{BH}$ (solid black or grey curves), along with contributions from new seed formation (dotted green) and gas accretion (dashed blue), as well as losses due to gravitational recoil (long-dashed red). The losses due to BHs being ‘stuck’ in unmerged satellite haloes are not shown. For the simulations with the ‘global warming’ feedback, the $\dot{\rho}_{BH}$ curves become very crowded near where the IGM heating begins to suppress seed BH formation and accretion. We have included a zoom-in view of this region of the plot (magenta boxes, lower right-hand panel) for ease of viewing.

As anticipated, the primary effect of BSMs is to suppress the formation of seed BHs (PopIII stars) at early times. The increase in the effective Jeans mass delays seed BH formation by $\Delta z \sim 3 - 4$ at $z > 20$ (Greif et al. 2011a) in simulations both with and without self-regulation. A somewhat counterintuitive result is that the suppression of $\rho_{BH}$ by BSMs (relative to the control simulations with BSMs turned off) decreases with time, even in the simple ‘no self-regulation’ models in which all BHs grow steadily at $2/3$ of the Eddington rate. By $z \approx 8$, the total SMBH mass density and its growth rate in the cases with and without BSMs are nearly indistinguishable. This is because the ‘extra’ seeds that form in the absence of BSMs do not contribute efficiently to $\rho_{BH}$. Many become satellites as their host haloes become tidally stripped during a minor merger, and those that do merge with other BHs are often ejected from the host halo by the gravitational recoil effect. The vulnerability of the earliest seed BHs to the recoil effect was pointed out by TPH12, who had also shown that the number and masses of massive BHs in merger-tree simulations depended weakly on the number of seed BHs formed. In other words, simply doubling (or halving) the number of seed BHs does not double (or halve) the total mass of nuclear BHs at later times.

The BH mass densities in simulations with and without BSMs converge earlier and more strongly in the presence of the global warming, self-regulating feedback by mini-quasars. To illustrate why this occurs, we have plotted in Fig. 2 the temperature history of the IGM, $T_{IGM}(z)$ (top panels) in our simulations, along with the mass scales that govern seed formation and accretion, $M_{cool}$ and $M_J$ (bottom panels). As with the previous figure, thick black curves show results from the simulations with BSMs, and thin grey curves show those without BSMs. We show the distribution of the cooling mass $M_{cool}$ in blue. The case with no BSMs is highlighted in light blue, the region between zero streaming velocity and the $1\sigma$ upper bound is shaded in blue, and the mean value of $M_{cool}$ is shown a solid blue curve. As with Fig. 1, we have zoomed in on the region of the figure where the IGM heating feedback begins to affect BH formation and accretion (bottom panels).

When $M_J$ exceeds $M_{cool}$, the formation of new BHs is rapidly suppressed. In addition, because the host haloes must exceed the Jeans mass and have recently merged with another such halo for BHs to continue growing, the Jeans mass threshold also acts as a thermostat for BH accretion. In the simulations without BSMs, the IGM heats slightly faster.
than in the case with BSMs, because there are more BHs to photoheat the IGM. However, this also means that the Jeans mass rises faster, and the negative feedback becomes effective earlier. Once the cosmological Jeans mass exceeds the Jeans cooling mass threshold, BH growth becomes strongly coupled to the IGM temperature, and BSMs effectively become irrelevant for BH growth and formation, at least in the context of the models considered here. This occurs at \( z \approx 16 \) without BSMs, and \( z \approx 14 \) with BSMs, but at \( z \leq 14 \) the IGM temperature (and reionization) histories are virtually indistinguishable.

3.2 Impact of BSMs on the most massive \( z \approx 6 \) quasar SMBHs

Let us now focus on how BSMs affect the population of the most massive SMBHs at \( z \approx 6 \). We consider the total mass \( \sum M_{\text{BH}} \) contained in the progenitor haloes that eventually assemble a single massive halo at \( z = 6 \) (i.e. the total mass of BHs inside a given merger tree). In Fig. 3 we plot the distribution of this quantity for the 70 most massive haloes (\( M_{\text{halo}} > 10^{12.85} M_\odot \)) at \( z = 6 \). That is, we take the 70 sets of progenitors (with each set associated with its own streaming velocity value) that eventually assemble our 70 most massive haloes at \( z \approx 6 \), compute \( \sum M_{\text{BH}}(z) \) for each set, and summarize the distribution. Here, \( \sum M_{\text{BH}}(z) \) denotes the total BH mass per parent halo, not the total BH mass in the 150 Gpc\(^3\) comoving volume of our simulations. As with the earlier figures, the darker (lighter) shades show the cases with (without) BSMs, while the left-hand (right-hand) panels show the cases without (with) self-regulation due to IGM heating. The purple lines in each panel show the mean values per progenitor set, and these are enveloped by shaded regions that denote the \( \pm 1\sigma \) distribution bounds. The top panels show the total accumulated \( \sum M_{\text{BH}} \) per parent halo, and the panels below show the distribution of cumulative BH masses created as new seeds (\( \pm 1\sigma \) bounds shaded in green), accreted (blue) and ejected via recoil (red). Because the masses of the parent haloes at \( z = 6 \) differ by less than a factor of 3, we have opted not to weigh \( \sum M_{\text{BH}} \) by the parent halo mass.

Note that negative feedback from IGM heating does not strongly affect the total seed mass created in these haloes (green curves). This is because most of the seeds are formed before the rise in \( T_{\text{IGM}} \) and \( M_J \) turns off seed production.

---

**Figure 1.** Evolution of \( \rho_{\text{BH}} \), the universal mass density of nuclear BHs. Simulations without BH self-regulation due to IGM heating are presented on the left-hand panels, and those with self-regulation are presented on the right-hand panels. The thick, darker lines indicate simulations with BSMs included, and the thin, lighter-coloured lines show those without streaming. Upper panels: the global mass density of BHs \( \rho_{\text{BH}} \) as a function of redshift (solid lines), and the mass density of only the BHs with masses above \( 10^5 M_\odot \) (dashed lines). Lower panels: the growth rate of the global nuclear BH mass density \( \dot{\rho}_{\text{BH}} \): net growth (solid black/grey); growth due to new seed formation (dotted green) and gas accretion (dashed blue) and losses due to gravitational recoil (long-dashed red).

**Figure 2.** The thermal history of the IGM in the four simulations. As with the previous figure, left-hand (right-hand) panels and light (dark) curves show simulations without (with) self-regulation and without (with) BSMs, respectively. Top: the IGM temperature history \( T_{\text{IGM}}(z) \). The temperature merely rises due to miniquasar photoheating in the simulations without self-regulation, but in those with self-regulation the temperature in turn affects the mass scale on which BHs can form and accrete. Middle and bottom: the evolution of the cosmological Jeans mass \( M_J \) as the IGM is heated. We have shown the distribution (shaded) and mean (solid thick line) of the cooling mass \( M_{\text{cool}} \) threshold for seed BH formation in blue. The lower bound of the shaded region, highlighted with a lighter shade of blue, shows the case without streaming. The upper bound represents a \( 1\sigma \) deviation from the mean. The bottom panels show a zoom-in of the region marked by magenta boxes in the middle panels, just before and after \( M_J \) overtakes \( M_{\text{cool}} \).
As pointed out by [12], these early-forming BHs cause the ‘global warming’ but are largely unaffected by it. The distribution of $\sum M_{\text{BH}}$ has a much narrower spread in the models without streaming velocities. Note that the thin curves are enveloped by narrow shaded bands denoting the $\pm 1\sigma$ bounds, whereas the shaded bands around the thick curves are wider (in some cases, the $1 - \sigma$ scatter is larger than the mean, resulting in the shaded regions extending to zero). Because each of the 70 haloes have millions of seed-forming progenitors, the total BH formation rate averages out to uniform values in each of these haloes. The subsequent growth of the BHs is uniform in the models without self-regulation, and determined by $T_{\text{IGM}}(z)$ (which is a global quantity) and the halo merger history (which averages out over large numbers) in the models with self-regulating feedback. In the models with BSMs, there is a larger scatter in the seed formation rates at early times due to statistical variations in $M_{\text{cool}}$ from halo to halo, which then propagates to the subsequent accreted and ejected masses. However, at late times, the statistical fluctuations decrease as (i) the spread in $M_{\text{cool}}$ decays with the streaming velocities, and (ii) in the models with self-regulation, the Jeans mass overcomes $M_{\text{cool}}$ as the relevant mass scale once the IGM is heated to $\sim 100 \, \text{K}$.

We see no evidence in our simulations that the BSMs impede the ‘survival of the fittest’ mode of merger-driven growth Volonteri & Rees 2008; [10] discussed in [11]. Naively, one might expect that suppressing the earliest seed BHs (which would get a head start in growth) would increase the number of ejected BHs (because mergers between BHs of similar masses result in larger recoil velocities). The number of BH mergers is slightly reduced in the simulations with BSMs (see also [3.3] andFig. 6), but the fraction of mergers that result in ejection are not appreciably greater.

Fig. 3 reemphasises the qualitative findings discussed earlier: BSMs can reduce the number of BHs at early times, but this relative suppression is smeared out at late times because (i) total BH masses have a sublinear dependence on the total number of seeds formed, since many seeds are lost as satellites or via gravitational recoil, and (ii) negative feedback due to IGM heating further tends to suppress the differences at earlier times by acting as a thermostat for BH formation and growth. The figure also confirms that while BSMs can be a powerful suppressant at early times in places that result in ejection are not appreciably greater.

3.3 Impact of BSMs on the $6 < z < 11$ quasar luminosity function

Our results so far indicate that BSMs are unlikely to have an appreciable effect on the SMBH population at $z \sim 6$, at least in the models we have considered. What about higher redshifts? We show in Fig. 4 the quasar luminosity function predicted by the models with self-regulating global warming feedback, at $z = 6, 9, 10$ and 11. These are shown as shaded regions, to account for the Poisson errors of our simulated sample and the ambiguity of the duty cycle in our model (i.e. whether our Eddington ratio of 2/3 for accreting BHs means that the BHs are shining all the time at $(2/3)L_{\text{Edd}}$, or only $2/3$ of the time at $L_{\text{Edd}}$). As with previous figures, the black, thick curves show the results from simulations with BSMs, and the grey, thin curves show the ones without. For convenience, we have also converted the number densities into the density per square degree per unit redshift, and expressed the results in terms of the flux density $F_\nu$:

$$F_\nu = \frac{L}{4\pi d_L^2} b \nu,$$

where $d_L$ is the luminosity distance and $b$ is the bolometric correction. We have marked the design flux density limit of JWST’s NIRCam instrument (10$^4$ s exposure) at $\approx 3 \, \mu\text{m}$ (from [http://www.stsci.edu/jwst/overview/design/]) by red vertical lines. The bolometric correction $b$ can vary from AGN to AGN, as well as with the intrinsic wavelength; for example, [12] report $b \approx 4.2, 5.2$ and 8.1 at wavelengths of 1450 Å (redshifted to $3 \, \mu\text{m}$ at $z = 19.7$) 3000 Å ($z = 9$) and 5000 Å ($z = 5$), respectively. Previous works (see table 2 in [12]) have arrived at values of $b$ within $\sim 30$ per cent of those quoted above. Given that the typical value of $b$ is unlikely to vary by more than a factor of two in the redshift range of interest here ($z \approx 6$ to 11, or intrinsic wavelengths of 2500–4300 Å), we adopt the constant value $b = 5$ for simplicity.

Our results at $z \approx 6$ agree well with the model luminosity functions of [17] and Shankar et al. 2009, as well as the observationally inferred luminosity

![Figure 3. The distributions of total BH mass per parent halo, as a function of $z$, for the 70 most massive $z = 6$ parent haloes in our simulations. The 1σ bounds are shaded, again with left-hand (right-hand) panels and light (dark) curves showing simulations without (with) self-regulation and without (with) BSMs, respectively. The mean of the distributions are shown with purple lines. BSMs introduce a wide scatter in BH occupation in different sets of progenitors at high $z$, but at lower redshifts the scatter decreases as the streaming motions decay and seeds are merged inefficiently due to BHs being stuck as satellites or being ejected via gravitational recoil. On average, BSMs do not strongly affect the total BH mass in progenitor haloes at $z < 15$.](image)
function of [Willott et al. (2010)], at \( L > 10^{46} \text{ erg s}^{-1} \), where the data are most robust. We have plotted these luminosity functions (see fig. 7 of [Willott et al. (2010)] in the upper left-hand panel of Fig. 4 as in blue curves [Willott et al. 2010 in solid, Shankar et al. 2009 in dotted, and Hopkins et al. 2007 in dashed curves). Our models appear to overproduce the less luminous quasars, but the number density of these dim objects is less certain [Willott et al. 2010].

For this specific SMBH growth model, BSMs have the largest effect at the high end of the (mini-) quasar luminosity function out to \( z \approx 10 \). Coincidentally, this is comparable to the largest distance at which JWST could detect growing BHs. It is plausible that models that predict the quasar population at such high redshifts can overpredict their abundance if they do not take BSMs into account. However, the suppression affects only the very massive end of the mass function at \( z \gtrsim 10 \), i.e. objects whose number densities are \(< 1 - 100 \text{ deg}^{-2} \) per unit redshift and thus are extremely difficult to detect even if their numbers were not suppressed by BSMs. In our models, BSMs reduce the masses of the most massive miniquasar BHs by a factor of a few at \( z > 10 \). At \( z \approx 6 \), the luminosity functions for the simulations with and without BSMs are almost indistinguishable. Our results suggest that the most massive BHs may actually benefit slightly from BSMs, as the negative effects due to fewer merging progenitors are outweighed by the positive effect of weaker X-ray heating at early times. In any event, the effects of BSMs on the BH mass function at \( z < 10 \) are sufficiently small as to be dwarfed by other theoretical uncertainties.

### 3.4 Impact of BSMs on BH merger rates at high \( z \)

Finally, we address whether BSMs could impact the number of mergers of massive BHs in the early Universe. Such mergers are primary targets of proposed space-based gravitational-wave detectors such as eLISA [Danzmann et al. 2013] and DEIGO [Kawamura et al. 2011]. In Fig. 4 we show the merger rates of BHs in the redshifted binary mass ranges \( 10^2 \text{ M}_\odot < (1+z) M < 10^7 \text{ M}_\odot \) (top/left curves), \( 10^4 \text{ M}_\odot < (1+z) M < 10^7 \text{ M}_\odot \) (middle curves, approximately coincident with the eLISA sensitivity window), and \((1+z) M > 10^7 \text{ M}_\odot \) (bottom/right curves). We have considered binaries with a minimum mass ratio \( M_2/M_1 > 0.01 \), neglecting BHs in satellite haloes whose dynamical merger times are longer than the Hubble time. As with Figures 4 through 8 left-hand (right-hand) panels and thin grey (thick black) curves denote models without (with) self-regulating feedback from IGM heating and streaming velocities, respectively. For the models considered, BSMs do not affect the expected merger rate by more than a factor of 2 at \( z < 15 \). Once again, our results suggest that BSMs will not appreciably affect the observability of nuclear BHs in the early Universe.

---

3 [https://www.elisascience.org](https://www.elisascience.org)
4 The factor \( 1+z \) arises from the degeneracy between the mass and redshift of a gravitational wave source (e.g. Hughes 2002).

© 0000 RAS, MNRAS 000, 000-000

Figure 4. Quasar luminosity functions in the simulations with self-regulating feedback due to IGM heating, with (thick, black curves) and without (thin, grey curves) BSMs. We have selected snapshots of the luminosity function at \( z = 6, 9, 10 \) and 11. The \( z = 6 \) luminosity functions of Willott et al. (2010), Shankar et al. (2009), and Hopkins et al. (2007) are shown in the upper left-hand panel in solid, dotted and dashed blue curves, respectively. BSMs appear to suppress the masses/numbers of SMBHs by a factor of few in the range of \( \sim 10^6 \) to \( \sim 10^9 \text{ M}_\odot \) (\( \sim 10^{44} \) to \( \sim 10^{47} \text{ erg s}^{-1} \)). The most drastic difference in the cases with and without BSMs occur at the most luminous (massive) end of the luminosity function at \( z > 10 \). While this could reduce the number of luminous quasars that could be detected by JWST at these redshifts, these objects are so rare (\( < 10 \text{ deg}^{-2} \) per unit redshift at \( z = 11 \)) that they are difficult to find in a blind search in any case.

### 4 CONCLUSIONS

We have investigated the possible consequences of the early suppression of the formation of seed BHs by the relative bulk motion of baryons against DM haloes in the early Universe. We investigated how BSMs can affect the population of SMBHs at later times (out to \( z = 6 \)) by way of a toy model in which all BHs grow at 2/3 of the Eddington limit, and a more realistic one in which they do so for a short period after their host haloes experienced a merger in which both haloes exceed the cosmological Jeans mass. In both pairs of models, we found that the effects of BSMs on observable quantities—(i) the total mass density of nuclear BHs, (ii) their mass function, (iii) their contributions to photoheating the IGM, and (iv) their merger rates—to be marginal. The only exception is that we found the BSMs could suppress the masses of the most massive BHs at \( z > 10 \); however, the suppression occurs only for the rarest objects that are unlikely to be found by JWST in any case. Our models suggest that the suppression of the nuclear BH population could be much greater at \( z > 20 \), but it is unclear whether this has observational consequences. Although we have only considered PopIII seeds, a simple extrapolation of our findings suggests...
that models with more massive seeds, in which seeds form later (e.g. Begelman et al. 2006, Lodato & Natarajan 2006, Latif et al. 2013, Schleicher et al. 2014), would be even less affected by BSMs.

An important caveat to the above summary is that we have assumed that the suppression by BSMs of the cold gas content in DM minihaloes only affects BH seed formation, and not their subsequent growth. Essentially, we have assumed that the amount of gas in the host halo is not a bottleneck for BH growth. There are extreme examples, such as present-day dwarf galaxies or the z ∼ 10 minihaloes far below the cosmological Jeans mass in our simulations, where the severe dearth of cool gas can almost completely inhibit BH growth; but the suppression by BSMs of the cold gas fraction and central gas densities at z ∼ 20 is thought to typically only be a factor of a few (e.g. Greif et al. 2011b, Naoz et al. 2013). Since nuclear BHs typically consume less than 1 per cent of the total baryonic mass in their host haloes, the factors that limit their growth are unlikely to be directly related to factors of a few in the amount of gas in the halo, but rather to opportunity (e.g. frequency of major mergers of the host halo) or other regulatory mechanisms (local or global feedback). The nuclear BH population may be more sensitive to suppression by BSMs if BH masses tend to scale strongly with the cold gas fraction of the host halo, or if growing BHs is somehow extremely sensitive to the gas content of the host at z ≥ 15 (e.g. through density and temperature thresholds for feedback or secular instabilities that could help fuel BH growth). Additionally, Greif et al. (2011b) found that BSMs can help drive turbulence inside minihaloes, which can impact the gas accretion rates and thus affect both the PopIII star masses as well as the growth of the seed BHs they leave behind (e.g. Krumholz et al. 2006).

Another caveat is the use of the cosmological Jeans mass in regulating BH formation and growth. It has been suggested that the more pertinent mass scale may instead be the so-called filtering mass, which depends on the prior temperature history of the IGM (Gnedin 2000). The Jeans mass may overestimate by a factor of a few the halo mass scale on which negative feedback from IGM heating becomes effective (e.g. Naoz et al. 2013).

In a similar vein, if the merger time-scales of BHs depend sensitively on the gas content and density profiles of their host (e.g. Mayer et al. 2007), then BSMs may act to suppress or flatten the BH merger rate event as a function of redshift.

We have also not included in our merger trees the suppression of DM halo abundances by BSMs. The results of Fialkov et al. (2012) suggest that this effect may be as important at z ∼ 30 as the increase in M_{min} (the threshold halo mass for PopIII star formation). Accounting for this effect would also somewhat increase the impact of the BSMs. However, considering that this suppression is about a factor of 2 at z ∼ 30 and that its relative importance is much lower at z ≥ 30 when BSMs are greatest, it is unlikely to change our basic conclusion.

We conclude that unless SMBH growth and mergers are extremely sensitive to order-unity fluctuations in the cold gas density of their host haloes, BSMs are unlikely to play a dramatic role in suppressing the formation of the first SMBHs. They may reduce the luminosities (masses) of the most luminous quasars at z > 10 by a factor of a few, but otherwise we find the overall average effect on the nuclear BH population and their observables at 6 ≤ z ≤ 10 to be marginal. BSMs could, however, still leave detectable imprints on small (∼ 100 Mpc) scales. If quasars and galaxies are systematically less luminous in regions affected by BSMs, such spatial correlations could be detectable in future large-area, high-redshift surveys that measure clustering properties of quasars and galaxies (see Dalal et al. 2011, Tseliakhovich et al. 2013). Similarly, spatial correlations in quasar luminosity and star formation can cause reionization to take place more inhomogeneously, which could leave detectable imprints in the 21cm power spectrum (Maio et al. 2011, McQuinn & O’Leary 2012, Visbal et al. 2012).

ACKNOWLEDGEMENTS

The authors are grateful to Jun Zhang for helpful discussions regarding the merger-tree algorithm and to Ryan O’Leary for discussions of the literature. We thank the anonymous referee, whose careful comments improved the clarity of this work. This work was partially supported by NASA grant NNX11AE05G (to ZH).
REFERENCES

Ade, P. A. R. et al. 2013, ArXiv e-prints 1303.5076
Alvarez, M. A., Wise, J. H., & Abel, T. 2009, ApJ, 701, L133
Baker, J. G., Centrella, J., Choi, D., Koppitz M., van Meter J. R., Miller M. C. 2006, ApJ, 653, L93
Barkana, R., & Loeb, A. 2001, Phys. Rep., 349, 125
Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289
Blaes, O. M. 2004, in Beskin V., Henri G., Menard F., Pelletier G., Dalibard J., eds, Accretion Discs, Jets and High Energy Phenomena in Astrophysics. Springer, New York, p. 137-185
Bogdanović, T., Reynolds, C. S., & Miller, M. C. 2007, ApJ, 661, L147
Boylan-Kolchin, M., Ma, C., & Quataert, E. 2008, MNRAS, 383, 93
Bromley, J. M., Somerville, R. S., & Fabian, A. C. 2004, MNRAS, 350, 456
Bromm, V., & Larson, R. B. 2004, ARA&A, 42, 79
Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Nature, 459, 49
Carroll, S. M., Press, W. H., & Turner, E. L. 1992, ARA&A, 30, 499
Dalal, N., Pen, U.-L., & Seljak, U. 2010, J. Cosmol. Astropart. Phys., 11, 7
Danzmann, K. et al. 2013, ArXiv e-prints 1305.5720
Dotti, M., Montuori, C., Decarli, R., Volonteri M., Colpi M., Haardt F., 2009, MNRAS, 398, L73
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Fan, X. et al. 2001, AJ, 122, 2833
Fan, X. et al. 2006, AJ, 132, 1176
Favata, M., Hughes, S. A., & Holz, D. E. 2004, ApJ, 607, L5
Fialkov, A., Barkana, R., Tseliakhovich, D., & Hirata, C. M. 2012, MNRAS, 424, 1335
Gnedin, N. Y. 2000, ApJ, 542, 535
Greif, T. H., White, S. D. M., Klessen, R. S., & Springel, V. 2011a, ApJ, 736, 147
Greif, T. H., Springel, V., White, S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011b, ApJ, 737, 75
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Haiman, Z. 2004, ApJ, 613, 36
Haiman, Z. 2013, in Wilkinson T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library, Vol. 396, The First Galaxies. Springer, Berlin, p. 293
Haiman, Z., & Loeb, A. 2001, ApJ, 552, 459
Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, ApJ, 591, 288
Hinshaw, G. et al. 2012, ArXiv e-prints 1212.5226
Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731
Hosokawa, T., Omukai, K., Yoshida, N., & Yorke, H. W. 2011, Science, 334, 1250
Hughes, S. A. 2002, MNRAS, 331, 805
Jarosik, N. et al. 2011, ApJS, 192, 14
Kawamura, S., Ando, M., Seto, N., et al. 2011, Class. Quantum Gravity, 28, 094011
Kidder, L. E. 1995, PRD, 52, 821
Krumholz, M. R., McKee, C. F., & Klein, R. I. 2006, ApJ, 638, 369
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013, MNRAS, 433, 1067
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Lodato, G., & Narayan, P. 2006, MNRAS, 371, 1813
Lousto, C. O., Campanelli, M., Zlochower, Y., & Nakano, H. 2010, Class. Quantum Gravity, 27, 114006
Madau, P., & Rees, M. J. 2001, ApJ, 551, L27
Maio, U., Koopmans, L. V. E., & Ciardi, B. 2011, MNRAS, 412, L40
Martini, P. 2004, in Coevolution of Black Holes and Galaxies. ed. L. C. Ho (Cambridge Univ. Press, Cambridge), p. 169
Mayer, L., Kazantzidis, S., Madau, P., Colpi M., Quinn T., Wadsley J., 2007, Science, 316, 1874
McQuinn, M., & O’Leary, R. M. 2012, ApJ, 760, 3
Milosavljević, M., Bromm, V., Couch, S. M., & Oh, S. P. 2009, ApJ, 698, 766
Mortlock, D. J. et al. 2011, Nature, 474, 616
Naoz, S., Yoshida, N., & Gnedin, N. Y. 2013, ApJ, 763, 27
Ohkubo, T., Nomoto, K., Umeda, H., Yoshida, N., & Tsuruta, S. 2009, ApJ, 706, 1184
Omukai, K., & Palla, F. 2001, ApJ, 561, L55
Pelupessy, F. I., Di Matteo, T., & Ciardi, B. 2007, ApJ, 665, 107
Peres, A. 1962, Phys. Rev., 128, 2471
Rundle, J. C., Brotherton, M. S., & Shang, Z. 2012, MNRAS, 422, 478
Schleicher, D. R. G., Palla, F., Ferrara, A., Galli, D., & Latif, M. 2013, ArXiv e-prints 1305.5923
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 19, 337
Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, ApJ, 690, 20
Shapiro, S. L. 2005, ApJ, 620, 59
Sheth, R. K., & Tormen, G. 2002, MNRAS, 329, 61
Stacy, A., Bromm, V., & Loeb, A. 2011, ApJ, 730, L1
Stacy, A., Greif, T. H., & Bromm, V. 2010, MNRAS, 403, 45
Tanaka, T., & Haiman, Z. 2009, ApJ, 696, 1798 (TH09)
Tanaka, T., & Menou, K. 2010, ApJ, 714, 404
Tanaka, T., Perna, R., & Haiman, Z. 2012, MNRAS, 425, 2974 (TPH12)
Tseliakhovich, D., Barkana, R., & Hirata, C. M. 2011, MNRAS, 418, 906
Tseliakhovich, D., & Hirata, C. 2010, Phys. Rev. D, 82, 083520
Turk, M. J., Abel, T., & O’Shea, B. 2009, Science, 325, 601
Visbal, E., Barkana, R., Fialkov, A., Tseliakhovich, D., & Hirata, C. M. 2012, Nature, 487, 70
Volonteri, M. 2010, A&A, 18, 279
Volonteri, M., Haardt, F., & Madau, P. 2003, ApJ, 582, 559
Volonteri, M., & Rees, M. J. 2006, ApJ, 650, 669
Willott, C. J. et al. 2009, AJ, 137, 3541
Willott, C. J. et al. 2010, AJ, 140, 546
Yoo, J., & Miralda-Escudé, J. 2004, ApJ, 614, L25
Zhang, J., Ma, C.-P., & Fakhouri, O. 2008a, MNRAS, 387, L13
Zhang, W., Woosley, S. E., & Heger, A. 2008b, ApJ, 679, 639