Early black holes in cosmological simulations: luminosity functions and clustering behaviour

Colin DeGraf,1⋆ Tiziana Di Matteo,1 Nishikanta Khandai,1 Rupert Croft,1 Julio Lopez2 and Volker Springel3,4

1McWilliams Center for Cosmology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
2Computer Science Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
3Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 68118 Heidelberg, Germany
4Zentrum für Astronomie der Universität Heidelberg, Astronomisches Recheninstitut, Mönchhofstr. 12-14, 69120 Heidelberg, Germany

Accepted 2012 May 10. Received 2012 April 23; in original form 2011 August 24

ABSTRACT
We examine predictions for the quasar luminosity functions (QLFs) and quasar clustering at high redshift (z ≥ 4.75) using MassiveBlack, our new hydrodynamic cosmological simulation which includes a self-consistent model for black hole (BH) growth and feedback. We show that the model reproduces the Sloan QLF within observational constraints at z ≥ 5. We find that the high-z QLF is consistent with a redshift-independent occupation distribution of BHs among dark matter haloes (which we provide) such that the evolution of the QLF follows that of the halo mass function. The sole exception is the bright end at z = 6 and 7, where BHs in high-mass haloes tend to be unusually bright due to extended periods of Eddington growth caused by high-density cold flows into the halo centre. We further use these luminosity functions to make predictions for the number density of quasars in upcoming surveys, predicting that there should be ~119 ± 28 (~87 ± 28) quasars detectable in the F125W band of the WIDE (DEEP) fields of the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) from z = 5 to 6, ~19 ± 7 (~18 ± 9) from z = 6 to 7 and ~1.7 ± 1.5 (~1.5 ± 1.5) from z = 7 to 8. We also investigate quasar clustering, finding that the correlation length is fully consistent with current constraints for Sloan quasars (r₀ ~ 17h⁻¹ Mpc at z = 4 for quasars above m_i = 20.2) and grows slowly with redshift up to z = 6 (r₀ ~ 22h⁻¹ Mpc). Finally, we note that the quasar clustering strength depends weakly on luminosity for low L_{BH}, but gets stronger at higher L_{BH} as the BHs are found in higher mass haloes.

Key words: black hole physics – methods: numerical – galaxies: active – galaxies: haloes – quasars: general.

1 INTRODUCTION
Supermassive black holes are believed to be present at the centre of most galaxies (Kormendy & Richstone 1995) and are found to be correlated with the properties of their host galaxies (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Graham & Driver 2007). These correlations provide strong evidence for a link between the growth of black holes (BHs) and the evolution of their host galaxies, generally attributed to some form of quasar feedback (Burkert & Silk 2001; Granato et al. 2004; Sazonov, Ostriker & Sunyaev 2004; Churazov et al. 2005; Di Matteo, Springel & Hernquist 2005; Kawata & Gibson 2005; Springel et al. 2005b; Begelman, Volonteri & Rees 2006; Bower et al. 2006; Croton et al. 2006; Ciotti & Ostriker 2007; Hopkins, Richards & Hernquist 2007; Malbon et al. 2007; Sijacki et al. 2007).

Perhaps the most fundamental statistical quantity in the study of quasars is the number density, often characterized as a quasar luminosity function (QLF). The QLF has been studied observationally (La Franca et al. 2002; Barger et al. 2003b; Fiore et al. 2003; Ueda et al. 2003; Croton et al. 2004; Cirasuolo, Magliocchetti & Celotti 2005; La Franca et al. 2005; Brown et al. 2006; Richards et al. 2006; Silverman et al. 2008; Ehbrero et al. 2009; Yencho et al. 2009), as well as through simulations, with BHs modelled both semi-analytically (Kauffmann & Haehnelt 2000; Volonteri, Haardt & Madau 2003; Wyithe & Loeb 2003; Granato et al. 2004; Marulli et al. 2008; Bonoli et al. 2009) and directly incorporated in the simulation (Hopkins et al. 2005c, 2006b; Hopkins et al. 2007; Marulli et al. 2009; DeGraf, Di Matteo & Springel 2010).

Another fundamental quantity for quasar populations is the strength of quasar clustering, and its evolution with redshift.
Numerous observational studies have been done (La Franca, Andreani & Cristiani 1998; Porciani, Magliocchetti & Norberg 2004; Croom et al. 2005; Myers et al. 2007; Shen et al. 2007; da Ângela et al. 2008; Ross et al. 2009; Shen et al. 2009), generally finding evidence for increasing clustering amplitude with redshift, in agreement with findings from simulations (Bonoli et al. 2009; Croton 2009; DeGraf et al. 2010). Clustering is especially significant because one can use it to estimate the typical dark matter haloes in which quasars are found simply by comparing the clustering strength of quasars to that of dark matter haloes. In particular, the luminosity dependence of the clustering (if any) can help determine if the bright and faint quasars populate the same haloes, suggesting that they may be the same population of quasars at different phases of their lives (see e.g. Hopkins et al. 2005c,a,d,b, 2006a) or different halo masses, suggesting that they are fundamentally different quasar populations (see e.g. Lidz et al. 2006).

Thus both QLF and quasar clustering can be used to investigate the populations of quasars being observed, how they populate dark matter haloes and how the typical luminosities correlate with hosts. However, investigations at high redshift have been extremely difficult. Simulations have proved difficult due to the volumes needed to produce high-redshift quasars in sufficient numbers, limiting the statistical investigations. Similarly, the difficulty observing such few objects (Willott et al. 2010a), and high-redshift quasar clustering is limited to lower redshift (e.g. $z \sim 4$ by Shen et al. 2009). However, upcoming surveys such as the James Webb Space Telescope (JWST) and the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) have the potential to drastically improve the high-$z$ observations, making now an ideal time to make predictions as to the statistics of high-redshift quasars.

In this paper we use a new large-scale hydrodynamic cosmological simulation to study the earliest supermassive BHs focusing on their luminosity and clustering properties. We take advantage of this simulation’s large volume (it is the largest cosmological simulation which directly models the growth and evolution of BHs) to investigate the statistical properties of the earliest supermassive BHs, focusing on the luminosity function and correlation length. We use these quantities to compare with current observational measurements, to make predictions for upcoming surveys and to investigate the implications of potential features which may be detected.

In Section 2 we describe the numerical simulation used for our analysis. In Section 3 we investigate the QLF (Section 3.1) and quasar clustering (Section 3.2) and compare with observations, and our results are summarized in Section 4.

## 2 Method

In this paper we use a new cosmological hydrodynamic simulation of a $533 h^{-1}$ Mpc box specifically intended for high-redshift investigations (see Table 1). The simulation uses the massively parallel cosmological TreePM-SPH code GADGET-3 (an updated version of GADGET-2, see Springel 2005) incorporating a multiphase interstellar medium (ISM) model with star formation (Springel & Hernquist 2003) and BH accretion and feedback (Di Matteo et al. 2005; Springel, Di Matteo & Hernquist 2005a).

In this paper we use a new large-scale hydrodynamic cosmological simulation to study the earliest supermassive BHs focusing on their luminosity and clustering properties. We take advantage of this simulation’s large volume (it is the largest cosmological simulation which directly models the growth and evolution of BHs) to investigate the statistical properties of the earliest supermassive BHs, focusing on the luminosity function and correlation length. We use these quantities to compare with current observational measurements, to make predictions for upcoming surveys and to investigate the implications of potential features which may be detected.

In Section 2 we describe the numerical simulation used for our analysis. In Section 3 we investigate the QLF (Section 3.1) and quasar clustering (Section 3.2) and compare with observations, and our results are summarized in Section 4.

### Table 1. Numerical parameters.

| Box size ($h^{-1}$ Mpc) | $N_p$ | $m_{DM}$ ($h^{-1}$ M$_\odot$) | $m_{gas}$ ($h^{-1}$ M$_\odot$) | $\epsilon$ ($h^{-1}$ kpc) |
|-------------------------|-------|-------------------------------|-------------------------------|-----------------------------|
| $533.33$                | $2 \times 3200^3$ | $2.8 \times 10^8$ M$_\odot$ | $5.7 \times 10^7$ M$_\odot$ | $5.0$                        |

© 2012 The Authors, MNRAS 424, 1892–1898
Monthly Notices of the Royal Astronomical Society © 2012 RAS
Degraf, Di Matteo & Springel (2011), finding that it does a good job reproducing the $M_{\text{BH}} - \sigma$ relation, the total BH mass density (Di Matteo et al. 2008), the QLF (Degraf et al. 2010) and the expected BH clustering behaviour (Degraf et al. 2011). This simple model thus appears to model the growth, activity and evolution of supermassive black holes in a cosmological context surprisingly well. In particular, we note that, having run several (smaller) simulations with varying resolution, the average BH properties converge among runs with MassiveBlack resolution and higher, and only at lower resolutions do changes begin to manifest. Thus our results remain in the converged regime with regard to resolution effects. We also note that Booth & Schaye (2009) and Johansson, Naab & Burkert (2008) have adopted a very similar model, and have independently investigated the parameter space of the reference model of Di Matteo et al. (2008), as well as varying some of the underlying prescriptions. For further details on the simulation methods and convergence studies done for similar simulations, see Di Matteo et al. (2008).

Because the simulation saves the complete set of BH properties (mass, accretion rate, position, local gas density, sound speed, velocity and BH velocity relative to local gas) for each BH at every time step, the BH output for such a large simulation is prohibitively difficult to analyse using previous techniques. For this reason, Lopez et al. (2011) developed a relational data base management system specifically for this simulation. A similar strategy has also been followed in the analysis of the Millennium Simulation (Lemson & Virgo Consortium 2006). In addition to providing a substantially more efficient query system for extracting information, this data base is significantly more flexible than traditional approaches. For a complete summary of the data base format and its efficiency, please see Lopez et al. (2011).

3 RESULTS

3.1 Luminosity function

In the left-hand panel of Fig. 1, we show the bolometric QLF for $L_{\text{BH}} > 10^{10} L_\odot$ at $z = 11–5$ (solid lines). We also show the observational data compiled by Willott et al. (2010b) from Sloan Digital Sky Survey (SDSS) main (filled diamonds; Fan et al. 2006), SDSS deep stripe (filled triangles; Jiang et al. 2009) and the Canada–France High-z Quasar Survey (asterisks; Willott et al. 2010a), and that compiled by Hopkins et al. (2007) (filled circles and open diamonds; Fan et al. 2001a,b; Barger et al. 2003b,a; Fan et al. 2003, 2004; Cristiani et al. 2004; Barger et al. 2005; Richards et al. 2005; Silverman et al. 2005). Note that Willott et al. (2010b) use the bolometric corrections of Richards et al. (2006) to convert to bolometric QLF, while Hopkins et al. (2007) use their own corrections, but these corrections are consistent with one another (Hopkins et al. 2007). We find that our simulation is generally consistent with observations, though we tend to predict a steeper slope than observations. Thus this large simulation volume shows that our model is capable of producing Sloan-type BHs at high redshift ($z = 5$ and 6), and in the correct abundances (and in fact of sufficiently high mass; see Di Matteo et al. 2012). We also note that although the luminosity of individual quasars can vary significantly over short time periods, the large number of simulated BHs means that the calculation of the QLF is insensitive to these variations. In particular, we computed the QLF at several subsequent time steps, and found minimal temporal variations (which remains within the error bars shown in Fig. 1).

To help better understand the quasar populations which we are simulating and the dark matter haloes in which they are found, we have performed a simple fit to characterize the relation between BH luminosity and host halo mass. To do this, we use a simple halo occupation distribution (HOD) model. Because we are only interested in the high-luminosity sources (and it is exceedingly rare to find multiple high-luminosity objects within a single halo, especially at such high redshift), we model the probability that a halo of a given mass ($M_{\text{Host}}$) will host a quasar above a specified luminosity ($L_{\text{BH, cut}}$) as a cumulative log-normal distribution (which is often used in galaxy HOD models for $\langle N_{\text{cen}} \rangle$; see e.g. Zheng et al. 2005)

$$f(L_{\text{BH, cut}} | M_{\text{Host}}) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\ln(M_{\text{Host}}) - \mu(L_{\text{BH, cut}})}{\sqrt{2} \sigma(L_{\text{BH, cut}})} \right) \right]. \quad (1)$$

Figure 1. Left-hand panel: quasar luminosity function for $z = 11–5$ (solid lines with Poisson error bars), along with the $z = 5$ and 6 observational data points compiled by Willott et al. (2010b) and Hopkins et al. (2007) (see text for complete references to observational data). We also show the predicted QLF from our HOD model using the $z = 5$ fit (dashed lines) and using redshift-specific fits (dotted lines). (See Table 2 for fitted parameters.) Right-hand panel: quasar magnitude function for $z = 11–5$, based on apparent bolometric magnitude. The regions probed by CANDELS and VIDEO surveys are shown as shaded regions (solid red – CANDELS DEEP; hatched blue – CANDELS WIDE; hatched green – VIDEO). The lower limits are based on the survey volume for $z \leq 0.5$. We also show the magnitude limits for JWST NIRCam (black dashed), LSST (pink dot-dashed), Dark Energy Survey (red long-dashed), and VIKING (yellow double dot-dashed).
Table 2. Luminosity dependence of $\mu$ and $\sigma$ (see equation 1). $\mu, \sigma(L_{\text{BH, cut}}) = A_{\mu, \sigma} \times \left(\frac{L_{\text{BH, cut}}}{10^{11} \text{L}_\odot}\right)^{B_{\mu, \sigma}}.$

| Redshift | $A_\mu$ | $B_\mu$ | $A_\sigma$ | $B_\sigma$ |
|----------|---------|---------|------------|---------|
| 5        | 27.05   | 0.0208  | 0.209      | 0.128   |
| 6        | 26.905  | 0.0191  | 0.239      | 0.133   |
| 7        | 27.07   | 0.0173  | 0.298      | 0.0     |

where $f(L_{\text{BH, cut}} | M_{\text{host}}$) is the fraction of haloes with mass $M_{\text{host}}$ which contain a quasar above $L_{\text{BH, cut}},$ and erf is the error function $\left(\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-y^2} dy\right)$ (see Table 2 for fitted parameters). We combine this occupation fraction (using parameters at $z = 5$) with the halo mass function of Reed et al. (2007) to predict the QLF, shown in Fig. 1 (blue dashed line). This predicted QLF matches the simulation well at $z = 5,$ confirming an accurate fit. We note that our HOD approach may slightly underpredict the low luminosities since the HOD parameters are fitted for bright ($>10^{11} \text{L}_\odot$) quasars, and because at low luminosities, satellite BHs (which are not included in our simple HOD) may become significant. For bright objects, however, this approach is very accurate. The analytic function of Reed et al. (2007) is used rather than the mass function obtained directly from our simulation so the approach can be used to estimate out to higher mass haloes and lower redshifts than our simulation would otherwise allow, thereby extending our predictive power. [We also note that the mass function of Reed et al. (2007) is consistent with our simulation.]

We also plot the QLF for higher redshifts using the $z = 5$ HOD fit (dashed lines), and find that it does a surprisingly good job of predicting the higher redshift QLF, with the exception of the luminous quasars at $z = 6$ and 7. This suggests that the manner in which quasars populate haloes is approximately redshift independent, save for the luminous quasars at redshifts 6 and 7. At these redshifts, we find the rarest (i.e. brightest) objects to be brighter than our $z = 5$ fit would expect. In other words, at $z = 6$ and 7, the quasars in the most massive haloes are more luminous than those in comparable haloes at other redshifts, a result of unusually high-accretion rates. To quantify this discrepancy, we performed HOD fits for $z = 6$ and 7 (Table 2) which are shown as dotted lines in Fig. 1. We note that $A_\mu$ remains approximately constant, suggesting that quasars of $\sim10^{11} \text{L}_\odot$ tend to populate similar-size haloes regardless of redshift (at least for $z \geq 5$). However, at $z = 6$ and 7 the smaller $B_\mu$ implies that BHs found in massive haloes at $z = 6$ and 7 tend to be brighter than those in comparable haloes at other redshifts, and the effect gets stronger for higher mass haloes/higher luminosity quasars.

This can be explained by considering the growth history of the massive BHs. Di Matteo et al. (2012) find that the most massive BHs typically undergo a period of rapid (i.e. Eddington) growth in this general redshift range (z ~ 6–8) as a result of high-density streams of cold gas which not only help assemble the first massive haloes, but appear to penetrate to the halo centres. These cold flows facilitate extended periods of Eddington growth of BHs in the most massive haloes, which manifests itself in the QLF by flattening the bright end. This continues until the energy released by the quasar heats the surrounding gas to such a point that it is blown away from the halo centre, and the BHs become self-regulated (see Di Matteo et al. 2011, for an investigation into this behaviour), at which point the bright end will steepen once again. In this picture, the QLF should roughly follow the evolution of the halo mass function except for those BHs undergoing their Eddington growth phase, where the luminosity should be unusually high. We find exactly these results in Fig. 1, where the $z = 5$ fit does a good job except in the bright end at $z = 6$–7, where Eddington growth is common.

In addition to the luminosity function shown in the left-hand panel of Fig. 1, we show the magnitude function in terms of apparent bolometric magnitude in the right-hand panel which we use as a basis for predictions for several upcoming surveys. In particular, we show the regimes probed by the CANDELS WIDE and DEEP surveys\(^1\) (solid red and hatched blue, respectively). These regions are bounded by the magnitude limit of the F125W filter (converted to bolometric magnitude using the spectral energy distribution (SED) of Hopkins et al. 2007), and the volume enclosed by the survey areas ($m < 26.4$ over 0.2 deg$^2$ for WIDE and $m < 27.4$ over 0.04 deg$^2$ for DEEP, assuming a redshift range of ±0.5). We predict that both the WIDE and DEEP programmes will find sources out to $z \sim 7$–8, with each probing a slightly different region of the luminosity function. We expect there to be $\sim119 \pm 28$ ($\sim87 \pm 28$) quasars bright enough in the F125W band of the CANDELS WIDE (DEEP) fields from $z = 5$ to 6, $\sim19 \pm 7$ ($\sim18 \pm 9$) from $z = 6$ to 7 and $\sim1.7 \pm 1.5$ ($\sim1.5 \pm 1.5$) from $z = 7$ to 8. This will drastically increase the number of high-redshift quasars, which will substantially improve the measurements of the faint end of the $z = 6$ QLF, and the different (though overlapping) ranges probed by the WIDE and DEEP programmes will help constrain the faint-end slope. We have also provided a similar bounded volume for the VISTA Deep Extragalactic Surveys Observations Survey (VIDEO), and magnitude limits for JWST’s Near-infrared Camera (NIRCam; black dashed line) (Gardner et al. 2006), Large Synoptic Survey Telescope (LSST; pink dot–dashed), Dark Energy Survey (red long dashed) and VISTA Kilo-degree Infrared Galaxy Survey (VIKING; yellow double dot–dashed), each converted to bolometric magnitudes (using the SED of Hopkins et al. 2007) to provide the means for predicting the number of high-$z$ quasars these surveys should find. Overall, these surveys should drastically improve the measurements of the faint-end QLF by improving the number of observed quasars, and pushing to luminosities an order-of-magnitude lower than current observations.

3.2 Quasar clustering

In addition to the number density of quasars, we look at the clustering properties characterized by the correlation length $r_0,$ defined as the scale at which the correlation function $\xi(r_0) = 1.$ In Fig. 2 we show the predicted correlation length as a function of lower luminosity cut for $z = 1$–6. These predicted correlation lengths are obtained using a halo mass function (Reed et al. 2007) and the halo bias factor of Sheth, Mo & Tormen (2001), weighted by the fraction of haloes hosting sufficiently luminous quasars (equation 1, using the $z = 5$ parameters for $z \leq 5$ and the $z = 6$ fit at $z = 6$). We note that the halo bias factor is not well constrained for the very high mass end (see e.g. Pillepich, Porciani & Hahn 2010; Tinker et al. 2010), and has not been investigated significantly at these high redshifts. Nonetheless, we find that the Sheth et al. (2001) bias factor does a good job of modelling our halo bias, though at $z \sim 6$ it underestimates by ~5 per cent relative to our simulation (note that we adjustment the bias factor by 5 per cent to account for this when calculating $r_{0,\text{BH}}$). This technique is important as it allows us

\(^1\)http://candels.ucolick.org

© 2012 The Authors, MNRAS 424, 1892–1898

Monthly Notices of the Royal Astronomical Society © 2012 RAS
to extend our predictions to higher luminosities where we do not have sufficient BHs to obtain clustering statistics.

To confirm the validity of this approach, we calculate the correlation length directly from the simulation using a maximum likelihood estimator (see e.g. Croft et al. 1997), and show the comparison in the inset plot (dashed lines – HOD model; solid lines – direct from simulation), finding good agreement in the luminosity range probed directly by our simulation. Thus we confirm that the large-scale clustering of luminous quasars can be well modelled by the clustering of dark matter haloes and a relatively simple quasar HOD (equation 1). In the main plot we also show curves of constant apparent magnitude (solid black curves), and a curve for the magnitude cut used for quasar selection in SDSS \(m_B < 20.2\) for \(z > 3\) and \(m_B < 19.1\) for \(z < 3\) (solid red curve). These curves show that using a magnitude cut rather than a luminosity cut will increase the observed redshift evolution. We also see that the evolution of the clustering strength with luminosity is relatively minor for low \(L_{BH}\) (where BHs occupy the slowly evolving end of the halo mass function), but becomes more significant at high luminosities where BHs occupy the high-mass tail of the mass function.

In addition, we note the effect of the change in the HOD model from \(z = 5\) (red) to 6 (orange). For low luminosities (below \(\sim 10^{13} L_\odot\)) there is minimal effect, due to the similarity in \(A_J\). However, the smaller \(B_m\) at \(z = 6\) (Table 2) means that higher luminosity quasars will typically be found in smaller haloes at \(z = 6\) than at \(z = 5\), consistent with findings of increasing \(M_{BH}/M_{\text{halo}}\) ratio with increasing redshift (e.g. Wang et al. 2010). [Note that this conclusion is for luminosity-selected quasars, which are biased against lower mass BHs. Therefore, we do not conflict with findings of a decreasing \(M_{BH}/M_{\text{halo}}\) ratio among volume-selected hosts (e.g. Alexander et al. 2008; Willott et al. 2010b).] In fact, we do find a slight decrease in the \(M_{BH}/M_{\text{halo}}\) relation from \(z = 6\) to 5.1.] This decrease in the typical host halo mass results in a correspondingly lower correlation length, and because the effect grows with luminosity, the difference manifests itself in the shallower slope of \(r_0\) versus \(L_{BH}\), even resulting in a crossover at \(\sim 4 \times 10^{13} L_\odot\). We note that this crossover is highly dependent on the HOD at high luminosities, and we are forced to extrapolate our HOD to reach these values. Thus this crossover is not a strong prediction (see further discussion below) but the general suppression is a clear trend in our simulation. In this way we see that \(r_0\) could be effected by a change in the HOD, but the effect should be fairly minor and will only occur at extremely high luminosities \((\gtrsim 10^{13.5} L_\odot)\).

In Fig. 3 we show the evolution in our predicted correlation length for SDSS-type quasars based on our \(z = 5\) HOD fit (for luminosity cuts corresponding to \(m_i < 19.1\) – black; \(m_i < 20.2\) – pink; and \(m_i < 23.0\) – blue), together with the observed measurements of Shen et al. (2009) (obtained with and without inclusion of negative data points, shown with dashed and solid error bars, respectively). The curves assume constant occupation behaviour (i.e. the \(z = 5\) HOD parameters are used at all redshifts). Based on earlier work, we would expect relatively minor evolution in the HOD between \(z = 5\) and 3 (Chatterjee et al. 2012), so the extrapolation should remain fairly accurate for \(z > 3\). This extrapolation shows that our simulation is able to reproduce the large correlation lengths \((r_0 \sim 17 h^{-1}\text{Mpc at } z = 4)\) from observational constraints. In addition, we match the rough evolution with redshift for \(z > 3\), and we expect \(r_0\) to continue to increase with redshift, reaching \(\sim 22 h^{-1}\text{Mpc by } z = 6\). Because we use a redshift-independent HOD model, this agreement with the observed evolution of \(r_0\) with redshift suggests that high-\(z\) quasar clustering evolution can be completely explained by the evolution in the clustering of dark matter haloes, i.e. without evolution in the mass of the typical host halo. Below \(z = 3\) we expect the active galactic nucleus (AGN) HOD to evolve more quickly (Chatterjee et al. 2012), so the extrapolation should not be considered an accurate prediction at lower redshifts (the dotted curves of Fig. 3), but we nonetheless appear to remain broadly consistent with observations.

In addition to the curves for the \(z = 5\) HOD, we also show the effect of the change in occupation distribution from \(z = 5\) to 6 (dashed lines). At \(z = 6\) the shallower slope in the \(L_{BH}-M_{\text{halo}}\) relation means that a given luminosity cut will correspond to a smaller typical host mass (and thereby produce a smaller \(r_0\) at \(z = 6\) than \(z = 5\). Because the difference in the HOD is in the slope \(B_m\)
of the $L_{\text{BH}}-M_{\text{Halo}}$ relation, the suppression gets stronger at higher luminosities and has essentially no effect on lower luminosities, which we show with the $m_1 = 23.0$ curve of Fig. 3. We note that the magnitude of the suppression is highly dependent on the slope of $\mu$ in equation (1) (found in Table 2) and is only significant at luminosities above those well probed by our simulation. As such, we cannot accurately estimate the magnitude of the suppression (or the exact luminosity at which the $z = 5$ and 6 curves of Fig. 2 will cross), but we nonetheless expect a suppressed slope in the $r_0$ evolution from $z = 5$ to 6. Given the approximate magnitude of the suppression found in our simulation (of the order of $\sim 5\%$) and the difficulties in measuring $r_0$ to high precision at such high redshifts, it is unlikely that such suppression will be detected in the near future, but it is nonetheless an effect which should be considered wherever possible.

### 4 CONCLUSIONS

In this paper we have investigated high-redshift BHs found within a new large-scale hydrodynamic simulation, focusing on the luminosity function and clustering behaviour. We model the QLF for $z \geq 5$, probing quasars to luminosities up to $\sim 10^{47} L_\odot$. We find reasonable agreement with observational data at $z = 5$ and 6, generally falling within the variation between surveys, confirming that our model is capable of producing Sloan-type quasars at very early times as well as matching their observed abundances.

Using a HOD fit for the central AGN at $z = 5$ (which we have provided), we find that the evolution in the QLF is well described by the evolution in the halo mass function (at least for the redshifts modelled by our simulation) without significant evolution in the occupation distribution, except for the high-luminosity end at $z = 6$–7 where we find a significantly flatter luminosity function. We postulate this flattening of the luminosity function to be a result of the largest BHs tending to undergo unusually rapid growth (and thereby producing unusually high luminosities) during these redshifts, a conclusion supported both by the difference in the HOD fits at these redshifts and by direct investigation into the light curves of these BHs (Di Matteo et al. 2012). At lower redshift (by $z \sim 5$), self-regulation suppresses the most massive BH growth, resulting in fainter bright-end quasars, and a corresponding steepening of the QLF. In particular, we note that this increase in luminosity at $z \sim 6$–7 should make observations of such high-$z$, luminous ($\sim 10^{47} L_\odot$) quasars easier than would otherwise be expected.

We used our luminosity function to provide estimates on the number density of detectable high-redshift quasars for several upcoming surveys. In particular, we expect the CANDELS survey to find quasars up to $z \sim 7$ in both the WIDE and DEEP programmes, with each probing a slightly different region of the QLF. At $z \sim 6$ these programmes should each detect dozens of quasars across a wide range of luminosities, drastically improving the observational constraints on the faint end of the high-redshift QLF, particularly the faint-end slope. In addition, we have provided our estimated number density at each redshift for several additional upcoming surveys, thereby providing the approximate number density of quasars to be found as well as the survey areas necessary to reach the highest redshifts.

We also investigate the clustering behaviour of these high-redshift quasars, finding luminosity-dependent correlation lengths of the order of $\sim 10$–14 h$^{-1}$ Mpc. Using our HOD model and the theoretical clustering of dark matter haloes, we show the correlation length as a function of quasar luminosity, finding relatively weak luminosity dependence at low $L_{\text{BH}}$, but with increasing $L_{\text{BH}}$ dependence at higher luminosities (where BHs are typically found in haloes in the steep end of the halo mass function). We also compare our predicted $r_0$ to high-redshift observations and find excellent agreement, with our simulation predicting $r_0 \sim 15$–20 h$^{-1}$ Mpc for Sloan-type quasars. We also roughly match the evolution of $r_0$ with redshift using a redshift-independent HOD, suggesting that high-redshift quasar clustering evolution is fully explained solely by the evolution in clustering of dark matter haloes without a change in the typical host halo mass.

Finally, we note the effect a change in the quasar occupation distribution from $z = 5$ to 6 has on $r_0$, finding that it should suppress the clustering strength of high-luminosity quasars at $z = 6$. Our limited sample size and evolution in the halo bias factor found in our simulation make it difficult to quantify the magnitude of this suppression, but we nonetheless conclude that some suppression should occur (though likely below the sensitivity of upcoming surveys).

### ACKNOWLEDGMENTS

This work was supported by the National Science Foundation, NSF Petapps, OCI-0749212 and NSF AST-1009781. The simulations were carried out on Kraken at the National Institute for Computational Sciences (http://www.nics.tennessee.edu/).

### REFERENCES

Alexander D. M. et al., 2008, AJ, 135, 1968
Barger A. J. et al., 2003a, AJ, 126, 632
Barger A. J., Cowie L. L., Capak P., Alexander D. M., Bauer F. E., Brandt W. N., Garmire G. P., Hornschemeier A. E., 2003b, ApJ, 584, L61
Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., Wang W.-H., Steffen A. T., Capak P., 2005, AJ, 129, 578
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bondi H., 1952, MNRAS, 112, 195
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, MNRAS, 396, 423
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Bromm V., Loeb A., 2003, ApJ, 596, 34
Brown M. J. I. et al., 2006, ApJ, 638, 88
Burkert A., Silk J., 2001, ApJ, 554, L151
Chatterjee S., DeGraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, MNRAS, 419, 2657
Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91
Cioffi D., Ostriker J. P., 2007, ApJ, 665, 1038
Cirasuolo M., Magliocchetti M., Celotti A., 2005, MNRAS, 357, 1267
Colberg J. M., Di Matteo T., 2008, MNRAS, 387, 1163
Cristiani S. et al., 2004, ApJ, 600, L119
Croft R. A. C., Dalton G. B., Efstathiou G., Sutherland W. J., Maddox S. J., 1997, MNRAS, 291, 305
Croft R. A. C., Di Matteo T., Springel V., Hernquist L., 2009, MNRAS, 400, 43
Crook C. S., Cochrane R., Boyle B. J., Shanks T., Miller L., Outram P. J., 2004, MNRAS, 349, 1397
Crook C. S. et al., 2005, MNRAS, 356, 415
Croton D. J., 2009, MNRAS, 394, 1109
Croston D. J. et al., 2006, MNRAS, 365, 11
da Angela J. et al., 2008, MNRAS, 383, 565
DeGraf C., Di Matteo T., Springel V., 2010, MNRAS, 402, 1927
Di Matteo T., Springel V., 2011, MNRAS, 413, 1383
Di Matteo T., Springel V., Hernquist L., 2005, Nat, 433, 604
