The formation of galaxies hosting $z \sim 6$ quasars

Nishikanta Khandai,∗ Yu Feng, Colin DeGraf, Tiziana Di Matteo and Rupert A. C. Croft

McWilliams Center for Cosmology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Accepted 2012 April 2. Received 2012 February 19; in original form 2011 November 2

ABSTRACT
We investigate the formation and properties of galaxies hosting $z \sim 6$ quasars in the gigaparsec scale cosmological hydrodynamical simulation MassiveBlack, which includes a self-consistent model for star formation, black hole accretion and associated feedback. We select the host galaxies based on the sample of quasars in MassiveBlack brighter than the Sloan Digital Sky Survey magnitude limit in the redshift range $5.5 \lesssim z \lesssim 6.5$. We find that quasar hosts in the simulation are compact gas-rich systems with high star formation rates (SFRs) of SFR $\sim 10^2 - 10^3$ $M_{\odot}$ yr$^{-1}$ consistent with observed properties. We show that the star-forming gas in these galaxies predominantly originates from high-density cold streams which efficiently penetrate the halo and grow the galaxy at the centre. The ratio of molecular to total cold gas mass in these galaxies is $M_{\text{mol}}/M_{\text{cold}} \sim 0.1$, much larger than local galaxies of similar masses, indicating that star formation in high-redshift quasar host galaxies is more efficient than their local counterparts. We show that MassiveBlack predicts a deviation from the local $M_{\text{BH}}-\sigma$ and $M_{\text{BH}}-M_*$ relations, implying that black holes are relatively more massive for a given stellar host at these redshifts.

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

1 INTRODUCTION
Supermassive black holes (SMBH) are now ubiquitously found in the nuclei of local galaxies. Tight correlations have been observed between the central black hole (BH) and its host galaxy (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Graham et al. 2001, 2011; Tremaine et al. 2002; Marconi & Hunt 2003; Haring & Rix 2004), implying that the growth of the quasar, powered by the SMBH, is intimately linked to the formation of its host galaxy. At $z \sim 6$ and above, the most direct constraint on the evolution of SMBHs comes from observations of luminous quasars in the Sloan Digital Sky Survey (SDSS; Fan et al. 2006; Jiang et al. 2009). Recently, a detection has been confirmed for a quasar at $z = 7$ (Mortlock et al. 2011). These quasars are mainly optically selected and represent the bright end of the quasar population at this epoch. They are rare, with number densities of $n \sim 1$ Gpc$^{-3}$ and extremely luminous with inferred masses of $M_{\text{BH}} \sim 10^9 M_{\odot}$, suggesting that they are harboured in rare haloes of mass $M_{\text{halo}} \sim 10^{11} M_{\odot}$ at these redshifts. Observations in other bands [far-infrared (FIR) to radio] suggest that the spectral energy distributions (SEDs) of these quasars are similar to those of their low-redshift counterparts (Wang et al. 2008). This means that $10^9$ $M_{\odot}$ BHs are fully developed and already in place even when the Universe was relatively young ($\lesssim 10^7$ yr at $z \gtrsim 6$).

Reprocessed thermal dust continuum emission in the FIR provides a clean method for deriving the total star formation rates (SFRs) in galaxies (Dale & Helou 2002). Bright quasars also add to the FIR luminosity, $L_{\text{FIR}}$, and one needs to correct for it to estimate the SFR of the host galaxy. The excess $L_{\text{FIR}}$ (corrected) for a sample of $z \sim 6$ quasar hosts suggests a massive starburst origin (SFR $\sim 10^{2.7} - 10^{1.4} M_{\odot}$ yr$^{-1}$) for them (Wang et al. 2010, 2011). These authors also detected large reservoirs of molecular gas, with mass $M_{\text{mol}} \sim 10^{10} M_{\odot}$, in these galaxies through the emission of redshifted carbon monoxide (CO) line. These results further corroborate that the observed large star formation activity in these galaxies is sustained through an abundant supply of molecular gas, the fuel for star formation. The cold gas in these galaxies is localized within a radius of a few kpc (Walter et al. 2004, 2009; Wang et al. 2010). The relation between CO luminosity, $L_{\text{CO}(1–0)}$, and $L_{\text{FIR}}$ for these galaxies are similar to those of typical star-forming systems at lower redshift.

However, the relation between the BH mass, $M_{\text{BH}}$, and the bulge velocity dispersion of its host, $\sigma$, is seen to be above the local relation (Wang et al. 2010), indicating that the $M_{\text{BH}}-\sigma$ relation is evolving. This is true even when some of the assumptions, like degeneracy of $\sigma$ with inclination angle, are considered. There is still significant debate whether there may be observational biases influencing these results (Lauer et al. 2007).

*E-mail: nkhandai@andrew.cmu.edu

© 2012 The Authors Monthly Notices of the Royal Astronomical Society: 2012 RAS

doi:10.1111/j.1365-2966.2012.21047.x
The requirement of large volumes and significant resolution to follow galaxy formation (gas inflows into relatively small scales) has made numerical studies of the growth of the first quasars extremely challenging. A number of approaches have resorted to ‘constrained’ simulations. In a pioneering study (Li et al. 2007), high-resolution merger simulations with subgrid models for star formation and growth of BHs have been used to study the formation of $z \sim 6$ quasars. This work extracted merger trees from large, coarse dark matter only simulations and identified the most massive halo candidate at $z = 6$. To simulate at high resolution, the merger trees were populated with isolated galaxies which undergo corresponding mergers, and are endowed with gas and models for star formation and BH growth.

This approach qualitatively reproduces the properties of the SDSS quasar J1148+5251 and its host at $z = 6.42$. Fully cosmological resimulations of selected haloes from the Millennium run were also carried out in a more recent study by Sijacki, Springel & Haehnelt (2009). This approach also finds that the most massive haloes are consistent with being the first quasar hosts.

In order to study directly how and where the first quasars assemble without a pre-imposed choice of halo, large cosmological volumes ($\sim$Gpc$^3$) for capturing rare high-sigma peaks and sufficiently high resolution in order to resolve kpc scales are required. High resolution is also necessary when including subgrid models for star formation, BH accretion and related feedback processes. Keeping these constraints in mind, we have run the largest (currently feasible) hydrodynamic simulation of its kind, MassiveBlack, which includes gravity, hydrodynamics and subgrid models for star formation, BH growth and associated feedback processes. It was run in a cosmological volume of $L_{\text{box}} = 533 h^{-1}$ Mpc with $2 \times 3200^3 \approx 65$ billion particles (dark matter + gas) and a uniform gravitational softening of $\epsilon = 5 h^{-1}$ kpc, with the code P-GADGET which has been extensively modified from the public code GADGET2 (Springel 2005) to run optimally on a large number of multicore processors. The simulations were carried out on $10^5$ cores of Kraken at the National Institute for Computational Sciences (NICS).1 MassiveBlack has the same mass and force resolution of the resimulated dark matter halo of Li et al. (2007); however, it does not rely on an imposed merger scenario to produce luminous quasars at $z \sim 6$ and tracks the assembly of galaxies in a more self-consistent manner. The subgrid models for star formation, BH growth and associated feedback processes in MassiveBlack are similar to the galaxy merger simulations constructed from the merger tree of the resimulated halo in Li et al. (2007); however, the mass resolution for gas and stars in MassiveBlack is much coarser.

The high-redshift observations of galaxies, quasars and their hosts can be used to test standard models of galaxy formation in previously unexplored regimes. Indeed, MassiveBlack has been instrumental in reproducing a number of observational properties of high-$z$ quasars, e.g. the formation and luminosity function of SDSS quasars powered by BHs of mass $M_{\text{BH}} \sim 10^8 M_\odot$ at $z = 6$ (DeGraf et al. 2011a; Di Matteo et al. 2011) and also statistical properties, such as their high-redshift clustering (DeGraf et al. 2011a). The aim of this paper is to take the best current observations of the highest redshift Sloan quasar hosts and compare them to simulation predictions for these objects. Given that MassiveBlack has reproduced properties of the highest redshift quasars, it is very important to check that the brightest, most massive BHs that power them also live in appropriate environments.

---

1 http://www.nics.tennessee.edu

### Table 1. Basic simulation parameters for the MassiveBlack simulation.

| $L_{\text{box}}$ (h$^{-1}$Mpc) | $N_{\text{part}}$ | $m_{\text{DM}}$ (h$^{-1}$M$_\odot$) | $m_{\text{gas}}$ (h$^{-1}$M$_\odot$) | $\epsilon$ (h$^{-1}$kpc) |
|-----------------------------|------------------|---------------------------------|---------------------------------|-----------------|
| 533.333                    | $2 \times 3200^3$ | $2.78 \times 10^8$              | $5.65 \times 10^7$              | 5               |

Our paper is organized as follows. We start by describing the MassiveBlack simulation in Section 2. We next identify a sample of potential quasar hosts at $z \sim 6$ and look at their formation and growth in Sections 3.1 and 3.2. In Section 3.3 we compare properties of these galaxies in MassiveBlack with recent observations. We present our conclusions in Section 4.

## 2 METHODS: THE MASSIVEBLACK SIMULATION

In this section we describe a large hydrodynamic simulation, MassiveBlack, which we have run to study the high-redshift Universe. We have used P-GADGET, a significantly upgraded version of GADGET3 (see Springel 2005, for an earlier version), which we are developing for use at upcoming Petascale supercomputer facilities. MassiveBlack is a cosmological simulation of a Λ cold dark matter ($\Lambda$CDM) cosmology.

It is worth pointing out that the number density of $z \sim 6$ quasars is extremely small, $n \sim$ a few (Gpc$^{-3}$), and that they are hosted by rare massive haloes. Therefore, in order to simulate and resolve these objects, one needs a large cosmological volume as well as high mass and force resolution. MassiveBlack was run with $N_{\text{part}} = 2 \times 3200^3 = 65.5$ billion particles in a comoving volume of side $L_{\text{box}} = 533 h^{-1}$ Mpc and a comoving gravitational softening length of $\epsilon = 5 h^{-1}$ kpc, (see Table 1, for more details).

The initial conditions were generated with the Eisenstein and Hu power spectrum at $z = 159$ and the simulation was evolved to $z = 4.75$. The cosmological parameters used were the amplitude of mass fluctuations, $\sigma_8 = 0.8$, spectral index, $n_s = 0.96$, cosmological constant parameter $\Omega_k = 0.74$, mass density parameter $\Omega_m = 0.26$, baryon density parameter $\Omega_b = 0.044$ and $h = 0.72$ (Hubble constant in units of $100 \text{km s}^{-1} \text{Mpc}^{-1}$).

Along with gravity and smoothed particle hydrodynamics, P-GADGET incorporates a multiphase interstellar medium (ISM) model with star formation (Springel & Hernquist 2003) and BH accretion and feedback (Di Matteo, Springel & Hernquist 2005, 2006; Di Matteo & Hernquist 2005).

### 2.1 Subgrid model for black hole accretion and feedback

In our simulation, BHs are modelled as collisionless sink particles within newly collapsing haloes, which are identified by a friends-of-friends (Davis et al. 1985) halo finder called on the fly at regular time intervals. A seed BH of mass $M_{\text{seed}} = 5 \times 10^7 h^{-1} M_\odot$ is inserted into a halo with mass $M_{\text{halo}} \geq 5 \times 10^{10} h^{-1} M_\odot$ if it does not already contain a BH. The seeding recipe is chosen to match the expected formation of SMBHs by gas directly collapsing to BHs with $M_{\text{BH}} \sim M_{\text{seed}}$ (Bromm & Loeb 2003; Begelman, Volonteri & Rees 2006) or by massive primordial stars collapsing into $\sim 10^5 M_\odot$ mass BHs (Bromm & Larson 2004; Yoshida et al. 2006) at $z \sim 30$ followed by sufficient exponential growth to reach $M_{\text{seed}}$. 

© 2012 The Authors, MNRAS 423, 2397-2406

Monthly Notices of the Royal Astronomical Society © 2012 RAS
Quasar hosts at $z \sim 6$

by the time the host halo reaches $\sim 10^{10} M_\odot$. Given our resolution, we cannot test the latter prescription but it nevertheless represents a plausible scenario. Once seeded, BHs grow by accreting surrounding gas or by merging with other BHs. Gas is accreted with an accretion rate $\dot{M}_{BH} = 4\pi G^2 M_{BH}^2 \rho / (c_s^2 + v_{BH}^2)^{3/2}$ (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Bondi 1952), where $v_{BH}$ is the velocity of the BH relative to the surrounding gas, $\rho$ and $c_s$ are the density and sound speed of the hot and cold phase of the ISM gas [which when taken into account appropriately as in Pelupessy, Di Matteo & Ciardi (2007) eliminates the need for a correction factor $\alpha$ previously introduced]. We allow the accretion rate to be mildly super-Eddington but limit it to a maximum allowed value equal to three times Eddington rate ($\dot{M}_{Edd}$) to prevent artificially high values, consistent with Begelman et al. (2006) and Volonteri & Rees (2006). The BH radiates with a bolometric luminosity which is proportional to the accretion rate, $L_{bol} = \eta \dot{M}_{BH} c^2$ (Shakura & Sunyaev 1973), where $\eta$ is the radiative efficiency and its standard value of 0.1 is kept throughout, and $c$ is the speed of light. Some of the liberated energy is expected to couple thermally to the surrounding gas. In the simulation, 5 per cent of the radiated energy does this. This energy is deposited isotropically on gas particles that are within the BH kernel (32 nearest neighbours) and acts as a form of feedback (Di Matteo et al. 2005). The value of 5 per cent is the only free parameter in the model and was set using galaxy merger simulations (Di Matteo et al. 2005) to match the normalization in the observed $M_{BH}-\sigma$ relation. BHs also grow by merging once one BH comes within the kernel of another with a relative velocity below the local gas sound speed.

This model for the growth of BHs has been developed by Di Matteo et al. (2005) and Springel et al. (2005). It has been implemented and studied extensively in cosmological simulations (Li et al. 2007; Sijacki et al. 2007, 2009; Colberg & Di Matteo 2008; Di Matteo et al. 2008; Booth & Schaye 2008; Croft et al. 2009; DeGraf et al. 2010, 2011b,c; Chatterjee et al. 2012), successfully reproducing basic properties of BH growth, the observed $M_{BH}-\sigma$ relation and the BH mass function (Di Matteo et al. 2008), the quasar luminosity function (DeGraf et al. 2010) and the clustering of quasars (DeGraf et al. 2011c).

We use a relational data base management system developed by Lopez et al. (2011) specifically for this simulation to track the history of BH properties (e.g. mass, accretion rate, position, local gas density, sound speed, velocity and BH velocity relative to local gas) which are saved for each BH at every time step. For a complete summary of the data base format and its efficiency, the reader is referred to Lopez et al. (2011).

3 PROPERTIES OF HOST GALAXIES OF $z \sim 6$ QUASARS

We now focus on individual properties of host galaxies of $z \sim 6$ quasars in this section. We look at how these galaxies and the BHs at their centres were assembled and compare our results with their observed properties in the latter part of this section.

3.1 Star formation and black hole growth

Our simulation allows us to follow the growth of SMBHs and their host galaxies up to $z \sim 5$. In Fig. 1, we plot the SFR and BH accretion rate in $M_\odot$ yr$^{-1}$. The SDSS flux limit of $m_i < 20.2$ for $z > 3$ (Shen et al. 2009) and converted to a bolometric luminosity (and then an accretion rate) using the SED of Hopkins, Richards & Hernquist (2007). This SDSS flux cut roughly corresponds to BH accretion rates of about $10 M_\odot$ yr$^{-1}$. We have chosen a conservative criterion

Figure 1. SFR (red) and BH accretion rate (black) in $M_\odot$ yr$^{-1}$. The SDSS flux limit of $m_i < 20.2$ for $z > 3$ (Shen et al. 2009) for the quasar sample is shown by a blue line.
for selecting this sample by assuming that luminosities of the quasar sample cross the SDSS detection limit for at least 10 time steps, which translates to them being bright for approximately 14 Myr (of the 206 Myr between $z = 6.5$ and 5.5). Such a condition ensures that quasars which have an artificially high luminosity (or accretion rate) in a single time step (e.g. when a gas particle comes extremely close to the BH for just a single time step) are not selected. In Fig. 1 we find that the SFR in the galaxy is strongly correlated with the growth of the central BH. The central BHs grow rapidly through a period of sustained Eddington accretion (typically between $8 < z < 6$) and are continuously fed by streams of high-density gas (Di Matteo et al. 2011). The typical masses of the host haloes in this sample grow from $M_{\text{halo}} = 10^{11.5} M_\odot$ at $z = 8$ to $M_{\text{halo}} = 10^{12.4} M_\odot$ at $z = 6$. Prior to its peak accretion phase, the BH grows more rapidly than the stellar mass of its host. Star formation is regulated by feedback from the BH, and typically is suppressed just prior to the peak accretion phase of the BH. This is a typical feature of this model and has been seen in many previous works (Di Matteo et al. 2005; Li et al. 2007; Sijacki et al. 2009). Once the BH accretion becomes feedback dominated it deposits enough energy in its vicinity and shuts off further growth by expelling gas in its surrounding star-forming region. At its peak, the SFR of the host galaxy is extremely high, $\mathcal{O}(10^4) M_\odot \text{yr}^{-1}$, and is consistent with observations for the SFR of quasar hosts at these redshifts (Wang et al. 2010, 2011). We make a direct comparison with observations in Section 3.3. Most of the SMBHs fall within the Sloan detection limit when they grow through Eddington limited accretion and attain $M_{\text{BH}} = 10^9 M_\odot$ during this phase. The one exception is the first object, which peaks early on at $z = 7.5$ then undergoes a merger (as will become evident in Fig. 2) at $z = 5.5$ to become a $M_{\text{BH}} = 10^9 M_\odot$ BH.

We look at the evolution of the environment around three typical host galaxies (the first three objects of Fig. 1) from $z = 7.5$ to 5.0 in Figs 2–4. In the top row of these figures, we plot the gas distribution, colour coded by temperature. The middle row is the gas distribution colour coded by the SFR and the bottom panel is the distribution of stars. The large blue circle denotes the virial radius of the halo and BHs are shown as smaller circles with radii proportional to their mass.

The SMBH in these examples have different growth histories, e.g. the first SMBH becomes feedback dominated at $z = 7.5$ then grows again due to a merger at $z = 5.5$ (seen in Fig. 2). Apart from this SMBH, the others grow to a mass of $10^9 M_\odot$ through Eddington-limited accretion by $z = 6$. As seen in Figs 3–4, these haloes are continuously fed by cold streams down to $z \sim 6$, prior to the feedback-dominated phase. At this redshift, feedback from the BH injects energy but is not able to disrupt the stream. Star formation is still sustained at SFR $\sim 10^5 M_\odot \text{yr}^{-1}$ until $z \sim 5.5$. By $z = 5.5$ feedback from the BH starts to destroy the inner structure of the cold streams, inhibiting further growth; the SFR drops down by an order of magnitude below its peak value. There are still pockets of cold star-forming gas around the BH; however, the depleted gas due to star formation is no longer replenished by the cold streams. By $z = 5$ the inner part of the smooth cold stream is mostly destroyed and the gas in the central region is heated to $T \sim 10^4 K$ and pushed out. Dense subhaloes which have not been destroyed by feedback continue to contribute to the growth of the host galaxy and the central BH, though at a much reduced rate.

In the middle row of Figs 2–4, we again look at the distribution of gas around the same object, but now colour coded by the SFR. The redshift of each panel is the same as in the top row. We find that the central region of the galaxy is forming most of the stars. Star formation also occurs in dense clumps of gas located on cold filamentary streams, though at a much reduced rate. The central region has a sustained period of star formation down to $z = 5.5$ for most objects. During this period, the BH has grown at the Eddington rate through smooth accretion of cold gas. Most of the star formation occurs in the most centrally located gas that also feeds the BH which Di Matteo et al. (2011) demonstrated is cold-flow fed. We wish to point out that the subgrid model and its predictions should be viewed in the context of cosmological growth of BHs rather than detailed physics of the accretion disc, which we do not resolve.

Figure 2. An example of the growth of a typical $z \sim 6$ quasar host galaxy. This host corresponds to the first object of Fig. 1. The top row visualizes the gas distribution colour coded by temperature across six redshifts. The projected density ranges from $\sim 10^5$ to $\sim 10^8 M_\odot (\text{kpc h}^{-1})^{-2}$. In the middle row, we visualize the gas distribution colour coded by the SFR. The bottom row shows the distribution of stars. The large blue circle is the virial radius of the halo. The smaller circles are BHs with the radius proportional to their mass.
in the simulation. This is true for all subgrid models which are employed in cosmological simulations including the star formation model which is used here.

In the bottom row of Figs 2–4, we look at the distribution of stars. As expected, the locations of stars are preferentially found near star-forming gas particles (middle row). We are unable to accurately predict the morphologies of the galaxies due to the limited resolution of the simulation.

3.2 Growth of host galaxies through cold streams

Here we investigate the origin of the star-forming gas in the quasar host galaxies. In particular, we will test whether most of this star-forming gas is indeed entering the halo via cold streams rather than cooling from a shock-heated phase. We wish to point out that we do not characterize the redshift of accretion of these star-forming gas particles, but rather look at the origin of all the star-forming gas during the host galaxy’s peak star-forming activity.

As an example, we look at the object in Fig. 4. We examine how the star-forming gas in the halo at its peak star-forming activity, i.e. at $z = 5.75$, was accreted. We trace the histories of these star-forming gas particles back to $z = 7.5$ and plot their temperature as a function of (physical) separation from the SMBH in Fig. 5. To make comparison with observations easier, in the discussion that follows all length scales are quoted in physical units. For star-forming gas particles, the effective temperatures for the two-phase medium (Springel & Hernquist 2003) are shown. For non-star-forming gas particles, the temperature for the single-phase medium is shown.

In Fig. 5, the filled black circles denote star-forming gas particles and the open blue squares denote the gas particles which have zero SFR. The horizontal dot–dashed line is the virial temperature, the

Figure 3. Same as in Fig. 2 but for the second object of Fig. 1.

Figure 4. Same as in Fig. 2 but for the third object of Fig. 1.
Figure 5. Temperature histories of the star-forming gas of the central halo at $z = 5.75$ as a function of separation from the central SMBH. All length scales in this figure are in physical units. This particular halo is the same as in Fig. 4 and has a peak SFR at $z \sim 5.75$. Star-forming gas particles in the central halo are tracked backwards with respect to the reference redshift of $z = 5.75$ when this halo was at its peak star-forming activity. The horizontal dot–dashed line denotes the virial temperature (kelvin) of the halo, the solid vertical line is the virial radius of the halo and the dashed vertical line is the gravitational softening length. The filled black circles denote gas particles with non-zero SFR and the open blue squares are gas particles with zero SFR. The bottom-right panel indicates the distribution of $T_{\text{max}}$ for star-forming gas particles at $z = 5.75$.

vertical solid line is the virial radius and the vertical dashed line is the gravitational softening length. Prior to $z = 5.75$, one can identify four distinct regimes in the $T$–$r$ plots of the gas particles which end up forming stars in the halo: (1) a cold stream of non-star-forming gas with temperatures of $\sim 10^4 \text{ K}$ beyond $r \sim 10 \text{ kpc}$, (2) clumps of star-forming gas outside $r \sim 10 \text{ kpc}$ with temperatures $10^4 \lesssim T \lesssim 10^6 \text{ K}$, (3) hot non-star-forming gas outside $r \sim 10 \text{ kpc}$ with temperatures $10^5 \lesssim T \lesssim 10^7 \text{ K}$ and finally (4) dense star-forming gas within $r \sim 10 \text{ kpc}$ with temperatures $10^3 \lesssim T \lesssim 10^5 \text{ K}$ and finally (4) dense star-forming gas within $r \sim 10 \text{ kpc}$ with temperatures $10^4 \lesssim T \lesssim 10^6 \text{ K}$. Star formation only occurs in dense environments, i.e. at the centre of the halo and clumps located in filaments as seen in Figs 2–4. A further investigation reveals that the cold non-star-forming gas (with $T \sim 10^4 \text{ K}$) is also located in filaments, whereas the hot non-star-forming gas is diffused and spread out across the halo.

As can be seen in Fig. 5, most of the gas that is star forming at $z = 5.75$ does not come from the diffuse hot medium. The major mode of gas accretion for the host galaxy is from gas in filamentary streams. The streams have two components, dense star-forming clumps and cold non-star-forming gas that penetrates deep into the halo well within the 10-kpc region of the central BH. At this point, gas is dense enough to form stars in the central galaxy and also fuel the growth of the BH. The mass of gas in star-forming clumps in the stream is small compared to the smoother cold non-star-forming gas within it. These clumps are subhaloes accreted on to the halo through filaments. Feedback from supernovae and the BH heats up the gas at the centre, so that the temperature systematically increases to $T \sim 10^6 \text{ K}$ with decreasing separation. This mode of accretion continues down to $z \sim 5.75$.

By $z = 5$ the BH has injected enough energy into the surrounding medium to destroy the inner structure of the cold stream. The BH’s growth has become self-regulated. The BH has also regulated the growth of its host. We find little star formation within 3 kpc of the BH, but a residual amount of dense gas is still clumped in the region with $3 \lesssim r \lesssim 10 \text{ kpc}$, so that star formation still persists in this part of the central galaxy.

A sustained period of cold accretion is largely responsible for the high SFR of most objects. The exception is the first object in Fig. 1 where a merger at $z = 5.5$ is responsible for increasing the SFR. As seen in Figs 2–4, star formation occurs only in dense regions, i.e. mostly in the central halo and at a reduced level in dense clumps.
within filaments. At its peak star-forming activity, 95 per cent of star formation occurs within the 3-kpc region of the central galaxy. The numbers quoted for this example are representative of those for the full sample, fluctuating by only a few per cent for the other galaxies.

We track the entire temperature history of the star-forming gas particles of the host galaxy at its peak star-forming phase, i.e. at \( z = 5.75 \). We define \( T_{\text{max}} \) as the maximum temperature that a gas particle attains prior to its first star-forming phase, i.e. prior to it first being in a two-phase medium (Kereš et al. 2005, 2009). A distribution of \( T_{\text{max}} \) will then indicate whether these gas particles were accreted through the hot or cold mode (Kereš et al. 2005, 2009). This is shown in the bottom-right panel of Fig. 5. As can be seen, the majority (>85 per cent) of the star-forming gas is accreted through the cold mode, the remaining gas (<15 per cent) is heated while accreting on to the halo and then cools to form stars. These hot gas particles are the same as the diffuse, hot, non-star-forming gas seen in the first four panels of Fig. 5. The smooth component of the cold stream is still significant, constituting ~70 per cent of its mass, whereas ~30 per cent lies in clumps consistent with Dekel et al. (2009). These results are, however, sensitive to the numerical resolution. It is likely that at higher resolution the structure of the stream will be different. The fraction of mass in the smooth component of the stream will be reduced as smaller clumps within the stream are better resolved. It is also possible that at higher resolution fluid instabilities will be better captured and disrupt some of the inflowing gas. However, we believe that we are in the converged regime for the objects considered here. For example, the results of Kereš et al. (2005) have converged for gas particles with mass \( m_{\text{g}} = 1.06 \times 10^8 M_{\odot} \) which is larger than the mass resolution of MassiveBlack. Additionally, tests have shown that the BH history (Di Matteo et al. 2008) and the bright end of the quasar luminosity function (DeGraf et al. 2010) have converged at the resolution of MassiveBlack.

However, within the context of our simulation, the numbers presented here represent a lower bound for cold mode accretion of star-forming gas since a larger fraction of star-forming gas particles prior to \( z = 5.75 \) should have accreted on to the halo through the cold mode. We find that if we considered all the gas, without distinguishing between star-forming and non-star-forming gases, then ~70 per cent of the gas at \( z = 5.75 \) is accreted cold and ~30 per cent is accreted through the hot mode.

The analysis presented here indicates that quasar host galaxies form stars very efficiently. The massive star formation is sustained through a supply of gas mostly from cold streams that are able to penetrate the halo and reach the centre without being heated to the virial temperature of the halo, consistent with earlier work at lower redshift (Dekel et al. 2009). A smaller fraction of gas <15 per cent does get heated to \( T_{\text{vir}} \) before finally cooling and forming stars. At its peak, ~95 per cent of the star formation occurs within a 3-kpc region of the halo centre, consistent with observations (Walter et al. 2004, 2009; Wang et al. 2010).

3.3 Comparison with observations

In this section, we compare properties of quasar hosts in the MassiveBlack simulation with recent observations (Wang et al. 2010). These observations specifically constrain observables for the host galaxies such as SFR, molecular gas \( M_{\text{mol}} = M(\text{H}_2 + \text{He}) \) and the \( M_{\text{BH}}-\alpha \) relation. We also look at the \( M_{\text{BH}}-\alpha \) relation for the host galaxies in the MassiveBlack simulation and compare it with the local relation in observations.

3.3.1 Star formation rates and cold gas

The reprocessed emission in the FIR from star formation heated dust is used to provide an estimate of star formation from observations of \( z \approx 6 \) quasar host galaxies (Wang et al. 2008). The contribution from quasars is removed and the remaining FIR luminosities, \( L_{\text{FIR}} \), are then converted to a SFR (Wang et al. 2010, 2011), using

\[
L_{\text{FIR}} \approx 2.55 \times 10^{-10} \left( \frac{M_{\odot}}{M_{\text{IR}} \text{yr}^{-1}} \right) \left( \frac{L_{\text{FIR}}}{L_{\odot}} \right). \tag{1}
\]

The sample of quasar host galaxies in the observations of Wang et al. (2010, 2011) falls in the redshift range \( 5.78 \lesssim z \lesssim 6.43 \). We compile the SFR of the quasar hosts in our simulation (seen in Fig. 1) and compare it with observations in the left-hand panel of Fig. 6. We find that most of these host galaxies have a peak SFR \( \sim 10^3 M_{\odot} \text{yr}^{-1} \) in the redshift range \( 5.5 \lesssim z \lesssim 6.5 \), comparable to the observed SFR of quasar host galaxies. However, our model seems to lack objects close to the observed SFR of \( 1500 M_{\odot} \text{yr}^{-1} \) for some host galaxies.

It is of course possible that our star formation model is somewhat simplistic, particularly at these redshifts, since we do not model star formation from molecular gas (e.g. Krumholz & Gnedin 2011). Further investigation of the star formation modelling is beyond the scope of this paper, so instead we prefer to compare directly to associated measurements of the cold and molecular gas, which is what the observations can constrain.

In the right-hand panel of Fig. 6, we plot the cold gas mass \( M_{\text{cold}} \) of the host (solid lines). We find that the evolution of \( M_{\text{cold}} \) broadly follows the trend seen in the SFR. On the other hand, observations probe the molecular gas mass \( M_{\text{mol}} \) of these hosts which is not directly modelled in our simulations.

The size of the cold molecular gas reservoir which fuels star formation in observed quasar host galaxies has been estimated through redshifted CO emission (Bertoldi et al. 2003; Walter et al. 2003; Carilli et al. 2007; Wang et al. 2010). These studies indicate molecular gas masses of \( \gtrsim 10^{10} M_{\odot} \) in these objects. Since our simulations do not model molecular gas, we instead estimate the amount of molecular gas by using the SFR as a proxy for \( M_{\text{mol}} \), i.e. by converting the SFR to \( L_{\text{FIR}} \) (equation 1) and using the relation between the CO (1–0) line luminosity, \( L'_{\text{CO}(1-0)} \) and \( L_{\text{FIR}} \) for local star-forming systems such as local starburst spiral galaxies, ultraluminous infrared galaxies and high-z submillimetre galaxies (Solomon & Vanden Bout 2005):

\[
\log (L_{\text{FIR}}) = 1.7 \times \log \left( L'_{\text{CO}(1-0)} \right) - 5. \tag{2}
\]

\( L'_{\text{CO}(1-0)} \) can then be converted to a cold molecular gas mass \( M_{\text{mol}} = aL'_{\text{CO}(1-0)} \) with \( a = 8.0 M_{\odot} (K \text{ km s}^{-1} \text{ pc}^2)^{-1} \) (Wang et al. 2010). Using these relations, we look at the evolution of molecular gas for quasar host galaxies in the MassiveBlack simulation and compare them with observations (Wang et al. 2010, 2011) in the right-hand panel of Fig. 6, shown as dot–dashed lines. Here again we find that we are able to reproduce the amount of molecular gas in these galaxies at redshifts \( 5.5 \lesssim z \lesssim 6.5 \). The estimates are in better agreement than the observed SFR. Wang et al. (2010) also find that \( z \sim 6 \) quasar hosts lie above the local \( L_{\text{FIR}}-L'_{\text{CO}(1-0)} \) relation and attribute it to unknown contributions from quasars to \( L_{\text{FIR}} \) and different dust temperatures and CO line ratios, even though they attempt to correct for them.

The fraction \( M_{\text{mol}}/M_{\text{cold}} \) is typically 10 per cent. This is much higher than the 1 per cent seen in observations of local galaxies (Obreschkow & Rawlings 2009). This indicates that \( z \sim 6 \) quasar host galaxies are forming stars more efficiently than local galaxies.
Figure 6. Left: evolution of SFR for the hosts of the luminous quasars. Right: evolution of $M_{\text{cold}}$ (solid lines) and $M_{\text{mol}}$ (dot-dashed) of the hosts of the most massive BHs. Data points in both panels are from observations (Wang et al. 2010, 2011). Data points on the right-hand panel are estimates for $M_{\text{mol}}$.

Figure 7. Evolution of $M_{\text{BH}}–\sigma$ (left) for quasar–host systems in the MassiveBlack simulation up to $z = 5.5$ (solid line) and further down to $z = 4.75$ (dotted line). The dashed line is the best-fitting local relation of Tremaine et al. (2002). The filled triangles are data from observations of $z \sim 6$ quasars and host galaxies (Wang et al. 2010) assuming an average inclination angle $\theta_{\text{inc}} = 40^\circ$ and open triangles are the same observations without any assumption for the inclination angles. In the right-hand panel, we show the evolution of the $M_{\text{BH}}–M_*$ relation for the same objects as in the left-hand panel. Comparison is made with the best-fitting local relation of Haring & Rix (2004) (dashed line).

3.3.2 The $M_{\text{BH}}–\sigma$ and the $M_{\text{BH}}–M_*$ relation

We now look at the evolution of $M_{\text{BH}}–\sigma$ and the $M_{\text{BH}}–M_*$ relation between the central BH and its host galaxy, where $\sigma$ is the velocity dispersion of the bulge, using stars within the half-mass radius as a proxy for the bulge.

The evolution of the $M_{\text{BH}}–\sigma$ relation for the MassiveBlack sample is shown in the left-hand panel of Fig. 7 (solid line). Comparison is made to the best-fitting local relation (dashed line, Tremaine et al. 2002) and observations of $z \sim 6$ quasar–host systems (triangles, Wang et al. 2010). In this panel, we show the evolution up to $z = 5.5$ (solid lines) in order to better compare with observations at these redshifts, and continue to $z = 4.75$ (dotted lines). Since the inclination angle is degenerate with the CO linewidths, an average inclination angle of $\theta_{\text{inc}} = 40^\circ$ was assumed (filled triangles, Wang et al. 2010); the open triangles do not consider any inclination angle. As expected, we see that the BHs grow more rapidly than their host galaxies and eventually end up above the local relation of Tremaine...
et al. (2002). By \( z = 5.5 \) the \( M_{\text{BH}} - \sigma \) relation compares reasonably well with observations that assume an inclination angle. However, observations indicate that there is an object with \( \sigma \sim 120 \text{ km s}^{-1} \) and \( M_{\text{BH}} \sim 10^{9.7} \text{ M}_\odot \) even when the assumption of inclination angle is considered. We do not find such an object within our sample and this discrepancy may have to do with the uncertainties in the inclination angle.

Next we look at the evolution of the \( M_{\text{BH}} - M_* \) relation in the MassiveBlack simulation (solid lines) and compare with the best-fitting local relation of Haring & Rix (2004) (dashed line) in the right-hand panel of Fig. 7. Again consistent with previous properties such as the SFR–\( M_{\text{BH}} \) relation and the \( M_{\text{BH}} - \sigma \) relation, we find that the BH is assembled more rapidly than the central host galaxy so that at high redshift the \( M_{\text{BH}} - M_* \) relation is well above the local relation of Haring & Rix (2004). It would be interesting with future simulations to see what mechanisms eventually drive them to the local relation.

The steep growth of \( M_{\text{BH}} \) which oupaces the stellar component of the host galaxy is common for all these systems. The cyan and red lines in Fig. 7 correspond to the first two objects of Fig. 1. A closer inspection shows that these objects undergo relatively dry mergers with subhaloes in the feedback-dominated phase. The subhaloes have a larger stellar component compared to gas and therefore during this merger both \( M_* \) and \( M_{\text{BH}} \) grow more rapidly than \( M_{\text{BH}} \). This can be seen as a relative flattening of the \( M_{\text{BH}} - \sigma \) and \( M_{\text{BH}} - M_* \) relations in Fig. 7. It is plausible that dry mergers are partially responsible for eventually bringing the high-redshift \( M_{\text{BH}} - \sigma \) and \( M_{\text{BH}} - M_* \) relations on to the well-constrained local relation (Tremaine et al. 2002; Haring & Rix 2004).

Full cosmological simulations with gas physics, BH growth and feedback have been carried out for the most massive dark matter halo at \( z = 6 \) of the Millennium run (Sijacki et al. 2009). These simulations produced an SMBH of mass \( M_{\text{BH}} = 2.01 \times 10^9 \text{ M}_\odot \text{ h}^{-1} \) at \( z = 6.2 \). These authors found that the SMBH has assembled most of its mass during extended periods of Eddington-limited accretion peaking at \( z \sim 6 \). The fate of the \( z = 6 \) SMBH and its host was further tracked down to \( z = 2 \). During this phase, the accretion rate becomes systematically sub-Eddington due to feedback, and decreases even in absolute terms with time. By extrapolating the accretion rates to \( z = 0 \) and taking into account the merger history of the halo from the Millennium catalogue, they conclude that it is possible to build a BH of mass \( M_{\text{BH}} \sim 3 - 4 \times 10^{10} \text{ M}_\odot \text{ h}^{-1} \) in a halo of mass \( M_{\text{halo}} \sim 10^{12} \text{ M}_\odot \text{ h}^{-1} \). These results are consistent with observations of BHs of mass \( M_{\text{BH}} \sim 10^{10} \text{ M}_\odot \) in the centre of the brightest cluster galaxies (BCGs; McConnell et al. 2011).

The evolution of galaxies that end up as BCGs in the local Universe has been studied with semi-analytical models of galaxy formation based on the Millennium simulation (De Lucia & Blaizot 2007). These authors find that stars in BCGs form early (50 per cent by \( z \sim 5 \) and 80 per cent by \( z \sim 3 \)). The satellite galaxies that accrete onto to BCGs at lower redshifts have little gas content, low SFRs, quite red colours and have an early-type morphology. These results therefore suggest that at \( z < 5 \) relatively dry mergers are responsible for building up the BCGs observed today. If we assume that the BH growth is slow in these satellite galaxies (since these are not formed in rare sigma peaks at \( z = 6 \)) and that considerable star formation has already taken place prior to the growth of the BH (or even seeding of the BH in our model), then it is possible that the BH-to-stellar mass ratio is below or at most lies on the local relation. Dry mergers of these satellites would then flatten the high \( M_{\text{BH}} - \sigma \) and \( M_{\text{BH}} - M_* \) to the more local relation. However, to demonstrate this effect, one has to either run MassiveBlack to \( z = 0 \) or study the evolution of a larger sample of luminous quasars and their hosts at \( z \sim 6 \) to \( z = 0 \) as in Sijacki et al. (2009).

4 Conclusions

MassiveBlack has previously been successful in reproducing a number of important properties of quasars at high redshift within observational constraints: (i) the formation and abundances of luminous Sloan-type quasars powered by BHs of mass \( M_{\text{BH}} \sim 10^9 \text{ M}_\odot \) at \( z = 6 \) (DeGraf et al. 2011a; Di Matteo et al. 2011), (ii) statistical properties such as the luminosity function of quasars and its evolution as well as high-redshift quasar clustering (DeGraf et al. 2011a). Importantly, MassiveBlack has indicated that cold flows can sustain the growth of BHs at Eddington rates so that they attain masses of \( \sim 10^{9} \text{ M}_\odot \) within the first billion years (Di Matteo et al. 2011).

This picture of cold flows feeding both quasars and their host galaxies at high redshift should be viewed in the context of subgrid modelling of star formation, BH growth and feedback in cosmological simulations. The success of the model eventually lies in its ability to reproduce current observations of these systems (which it does). However, one may question the validity of this picture at smaller scales since our resolution is limited to a few kpc and not e.g., the accretion disc around the BH. However, recent simulations of massive galaxies at \( z \sim 6 \) show that this picture is valid at scales 10–100 pc \( h^{-1} \) (Dubois et al. 2011). They find, consistent with the results presented here, that the growth of the most massive galaxies is due to the extremely efficient funnelling of gas from cold streams. We therefore conclude that results presented here are not strongly affected by resolution.

In this paper, we have looked at the formation of galaxies hosting \( z \sim 6 \) quasars with the MassiveBlack simulation. We are able to reproduce a number of observational properties of these galaxies. We summarize our findings below.

(i) At \( z \sim 6 \), \( M_{\text{BH}} \sim 10^9 \text{ M}_\odot \) BHs are already in place. The growth of the BHs and their host galaxies are strongly correlated. The BHs regulate the growth of their host galaxies.

(ii) BHs grow faster than the host galaxy and this is reflected in the deviations seen in our study from the local \( M_{\text{BH}} - \sigma \) and \( M_{\text{BH}} - M_* \) relations; our results are however consistent with observational findings for these deviations at \( z \sim 6 \).

(iii) The cold streams that fuel the BH are also responsible for feeding the host galaxy to sustain the excessive starburst in them. The accretion of gas, which eventually forms stars in the host, is dominated by cold accretion with the hot mode accretion being subdominant (\( \lesssim 15 \) per cent). We find that 30 per cent of the accreted mass in the stream is in clumps.

(iv) Quasar host galaxies grow in extreme environments. We are able to reproduce the high SFR of \( \sim 10^9 \text{ M}_\odot \text{ yr}^{-1} \) seen in observations in redshifts 5.5 \( \lesssim z \lesssim 6.5 \). Our simulation suggests that the observed galaxies are at the peak of their star formation activity at these epochs.

(v) Our derived estimates of molecular gas in these galaxies are consistent with observations. Such large reservoirs of cold molecular gas (\( M_{\text{mol}} \geq 10^{10} \text{ M}_\odot \)) are responsible for fuelling star formation in these galaxies.

(vi) Using the SFR as a proxy for \( M_{\text{bol}} \) we find that the ratio of molecular to cold gas is larger than seen for local galaxies. This again indicates the host galaxies are forming stars more efficiently than their local counterparts.

(vii) We find that most of the star formation occurs within a compact 3–6 kpc (physical) region around the BH at the peak of its
star-forming activity. These scales are again consistent with observations (Walter et al. 2004, 2009; Wang et al. 2010).

(viii) Given that we are able to reproduce observations between $5.75 \lesssim z \lesssim 6.5$, our simulations suggest that these quasars and their host galaxies are seen at the peak of their assembly.

From the results presented here, we expect that the $M_{\text{BH}}-\sigma$ relation in the high-redshift Universe is very different from the local Universe. This has been indicated in earlier work (Di Matteo et al. 2008), though the sample was too small to put strong constraints on the relation at $z = 5$ and above. Given the large volume and high resolution (hence large sample of galaxies and BHs), Massive-Black is well suited for predicting the evolution of the $M_{\text{BH}}-\sigma$ and $M_{\text{BH}}-M_*$ relations at $z = 5$ and above. It would also be interesting to look at the mechanisms responsible for their evolution. We will address these issues in subsequent work.

ACKNOWLEDGMENTS

This work was supported by NSF award OCI-0749212. This research was supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Kraken at the National Institute for Computational Sciences (http://www.nics.tennessee.edu). NK would like to thank Volker Springel for useful comments on the manuscript.

REFERENCES

Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bertoldi F. et al., 2003, A&A, 409, L47
Bondi H., 1952, MNRAS, 112, 195
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Bromm V., Loeb A., 2003, ApJ, 596, 34
Carilli C. L. et al., 2007, ApJ, 669, L9
Chatterjee S., DeGraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, MNRAS, 419, 2657
Colberg J. M., Di Matteo T., 2008, MNRAS, 387, 1163
Croft R. A. C., Di Matteo T., Springel V., Hernquist L., 2009, MNRAS, 400, 43
Dale D. A., Helou G., 2002, ApJ, 576, 159
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
DeGraf C., Di Matteo T., Springel V., 2010, MNRAS, 402, 1927
DeGraf C., Di Matteo T., Khandai N., Croft R., Lopez J., Springel V., 2011a, preprint (arXiv:1107.1254)
DeGraf C., Oborski M., Di Matteo T., Chatterjee S., Nagai D., Richardson J., Zheng Z., 2011b, MNRAS, 1136
DeGraf C., Di Matteo T., Springel V., 2011c, MNRAS, 413, 1383
Dekel A. et al., 2009, Nat, 457, 451
Di Matteo T., Springel V., Hernquist L., 2005, Nat, 433, 604
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33

This paper has been typeset from a TeX/BiTeX file prepared by the author.