Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **454,** 1038–1057 (2015) doi:10.1093/mnras/stv2056

The impact of angular momentum on black hole accretion rates in simulations of galaxy formation

Y. M. Rosas-Guevara,[1⋆] R. G. Bower,[1⋆] J. Schaye,[2⋆] M. Furlong,[1] C. S. Frenk,[1] C. M. Booth,[3] R. A. Crain,[2,4] C. Dalla Vecchia,[5] M. Schaller[1] and T. Theuns[1,6]

[1]*Institute for Computational Cosmology (ICC), Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[3]*Department of Astronomy and Astrophysics, The University of Chicago, Chicago, IL 60637, USA*
[4]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[5]*Max Planck Institute for Extraterrestrial Physics, Gissenbachstraße 1, D-85748 Garching, Germany*
[6]*Department of Physics, University of Antwerp, Campus Groenenborger, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*

Accepted 2015 September 2. Received 2015 July 15; in original form 2013 November 29

ABSTRACT

Feedback from energy liberated by gas accretion on to black holes (BHs) is an attractive mechanism to explain the exponential cut-off at the massive end of the galaxy stellar mass function. Most previous implementations of BH accretion in hydrodynamical simulations of galaxy formation have assumed that BHs grow at an accretion rate that is proportion to the Bondi rate. A major concern is that the Bondi accretion rate is inappropriate when the accreting material has significant angular momentum. We present an improved accretion model that takes into account the circularization and subsequent viscous transport of infalling material, and implemented as a 'subgrid' model in hydrodynamic simulations. The resulting accretion rates are generally low in low mass ($\lesssim 10^{11.5}\,\mathrm{M}_\odot$) haloes, but show outbursts of Eddington-limited accretion during galaxy mergers. During outbursts these objects strongly resemble quasars. In higher mass haloes, gas accretion peaks at ∼10 per cent of the Eddington rate, which is thought to be conducive to the formation of radio jets. The resulting accretion rate depends strongly on the effective pressure of the gas surrounding the BH, which in turn depends strongly on halo mass. This induces a sharp transition in the importance of BH feedback. In small haloes, the growth of galaxies is regulated by star formation and supernova feedback, but above a halo mass of $10^{11.5}\,\mathrm{M}_\odot$, rapid BH growth leads to the suppression of star formation and reduced growth of stellar mass with increasing halo mass.

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

1 INTRODUCTION

A fundamental open question in galaxy formation is the role that black holes (BHs) play in shaping the galaxy around them. The observed scaling relations between the mass of the central supermassive BH and the properties of the bulge (Magorrian et al. 1998; Tremaine et al. 2002; Mullaney et al. 2012) suggest that there is an intimate connection between the growth of the central BH and the evolution of its host galaxy. From these observations, however, it is not clear whether the formation of the BH plays an integral part in the galaxy formation process, or whether it is simply a biproduct of the process of galaxy's evolution. There are two lines of argument that suggest that the first option is correct. First, the energies that are available from the formation of a $10^9\,\mathrm{M}_\odot$ BH are enormous, greatly exceeding the binding energy of a galaxy's baryonic halo. Unless the coupling of the accretion energy to the surrounding baryons is extremely weak, it would be surprising if the formation of the BH has little impact on its surroundings. Secondly, there is a strong observed correlation between the mechanical power of radio galaxy lobes in galaxy clusters and the cluster gas cooling rate. Many authors have argued that the energy deposited by the radio galaxy is sufficient to replenish the cooling radiation of the system (McNamara & Nulsen 2007). Careful observations of galaxy groups have revealed evidence of similar levels of energy input to radio galaxies in groups (Antognini, Bird & Martini 2012; Birzan et al. 2012; Ma, McNamara & Nulsen 2013); this regime is more relevant to the connection between BH and galaxy growth.

In phenomenological or semi-analytic models, feedback from active galactic nuclei (AGNs) is an indispensable element that enables the models to reproduce the stellar mass function (SMF) of the local

⋆E-mail: y.m.rosas-guevara@dur.ac.uk (YMR-G); r.g.bower@dur.ac.uk (RGB); schaye@strw.leidenuniv.nl (JS)

universe (Bower et al. 2006; Croton et al. 2006; Bower, McCarthy & Benson 2008). AGN feedback is assumed to be ineffective in low-mass haloes, where the gas cooling time is short compared to the sound-crossing time (White & Frenk 1991), and only to couple effectively in quasi-hydrostatic haloes ($M \gtrsim 10^{12}\,\mathrm{M}_{\odot}$). This dichotomy has some observational support, since the bulk of energy output from quasars appears to be radiated, while the mechanical energy of radio galaxies is trapped in the overall potential. As a result, a characteristic mass scale is introduced where the SMF presents a break: accretion in low-mass haloes is dominated by cold and rapidly cooling gas (since the cooling time is less than the free-fall or sound-crossing time), while accretion in high-mass halo occurs through quasi-hydrostatic cooling gas flows (where the gas is approximately in pressure balance, and the sound-crossing time is less than the cooling time; White & Frenk 1991). The importance of this distinction can be understood if the primary driver of the star formation rate in galaxies is the balance between outflows and inflows (i.e. the star formation rate of galaxies adjusts itself so that the inflow and outflow are in equilibrium). In the case of low-mass haloes, the AGN feedback loop is (assumed to be) ineffective and the balance between gas supply and outflow is set by the supernova (SN)-driven outflow rate. In higher mass haloes, the AGNs regulate the galaxy growth either by offsetting the cooling rate (Bower et al. 2006), or by puffing up the hot gas halo (Bower et al. 2008; McCarthy et al. 2011; Bower, Benson & Crain 2012) so that the cooling rate is reduced. In either case, the result is to suppress the mass of the cold gas and reduce the star formation rate in massive haloes, creating a break in the SMF.

In this scenario, the distinction between rapid cooling and hydrostatic haloes is critical. In the absence of a clear physical scale at which BH feedback becomes effective, the SMF behaves as a power law (Benson et al. 2003; Bower et al. 2012) because the impact of gas ejection builds up over a wide range of stellar mass. However, semi-analytic models make a variety of simplifying assumptions, and it is possible that AGN-driven and star-formation-driven outflows might not combine as simply as is envisaged.

Hydrodynamic simulations have the great advantage that there is no need to make an explicit distinction between hydrostatic and rapidly cooling haloes. Any dependence on the ratio of cooling and dynamical time-scales should emerge from the solution of the hydrodynamic equations. There is a long history of papers that include AGN feedback in numerical simulations. Di Matteo, Springel & Hernquist (2005) and Springel, Di Matteo & Hernquist (2005) introduced BH fuelling in order to study the impact of BH accretion on galaxy mergers (Hopkins et al. 2006, 2007; Sijacki et al. 2007; Barai et al. 2013). More recently, some papers have focused on modelling both feedback modes, *quasar* and *radio*, either making use of two accretion modes (Sijacki et al. 2007; Vogelsberger et al. 2014) or different ways of implementing jet feedback (Dubois et al. 2010; Debuhr, Quataert & Ma 2011).

The approach we follow here is extensively based on the development of the Springel et al. (2005) accretion model presented in (Booth & Schaye 2009, hereafter BS09). Booth & Schaye (2010) emphasize the importance of self-regulation: BHs grow until their energy output is comparable to the binding energy of the halo, resulting in a very tight correlation between halo mass and BH mass. They demonstrate that the slope of the correlation departs from unity because of the variation of halo concentration, and that changes in the feedback efficiency result in an offset in BH mass such that the rate by which the BH releases energy remains fixed. These simulations focused on the physics of BH accretion at high masses, and therefore used a relatively high particle mass. The OverWhelmingly Large Simulations (OWLS) project which consists of a large suite of cosmological, smoothed particle hydrodynamics (SPH) simulations with varying boxes and resolutions, includes AGN feedback in some of its simulations. The highest resolution simulations (512^3 particles for the 25 Mpc box and gas particle mass $1.4 \times 10^6\,h^{-1}\,\mathrm{M}_{\odot}$), run only up to $z = 2$ making it impossible to compare with the observational data on the local galaxy mass function (Schaye et al. 2010) at this resolution. Nevertheless, the analysis of the properties of galaxies in the lower resolution OWLS simulations ($8.7 \times 10^7 h^{-1}\,\mathrm{M}_{\odot}$) at $z = 0$ by Muldrew, Pearce & Power (2013) hints at the problem of a power-law SMF and the absence of a physical scale on which the BH becomes effective. This problem is also seen in the simulations of Puchwein & Springel (2013). Motivated by the success of semi-analytic models, Sijacki et al. (2007) explicitly introduce different feedback schemes for high- and low-mass accretion rates. This approach along with a *radiative mode*, that includes the effects of a strong ionizing radiation emerging from active BHs in the net gas cooling rates, is used in the recent large volume cosmological simulations of the ILLUSTRIS project (Vogelsberger et al. 2014; Sijacki et al. 2015). For a cluster-scale approach, Dubois et al. (2013) follow the evolution of a 10^{11}-$\mathrm{M}_{\odot}$ mass halo progenitor of a cluster at $z = 0$, including SN and AGN feedback. Dubois et al. (2013) found that the feedback from star formation plays a marginal role and that AGN feedback is able to quench star formation when BHs are self-regulating.

In this paper we will take another look at the BH accretion models used in these numerical simulations. In the models above, the BH accretion rate is based on the Bondi estimate. This assumes that material falling through the Bondi radius free-falls into the BH. As recently discussed by several authors (e.g. Krumholz, McKee & Klein 2005; Debuhr, Quataert & Ma 2011; Power, Nayakshin & King 2011; Anglés-Alcázar, ÃÜzel & Davé 2013), neglecting the angular momentum (AM) of the flow may lead to significant errors. Power et al. (2011) propose a new model in which a BH particle represents a self-regulating torus. Orbits are used to estimate whether particles are captured within a given accretion region of the BH. Then gas particles are accreted on to the BH after a viscous time. Both the accretion region and viscous time are given by free parameters in their model. Their method is suitable for ultra high resolution calculations (with particle mass $10^2\,\mathrm{M}_{\odot}$) where the accretion region is $\sim$0.003 kpc, but it is not suitable for the multi-megaparsec scale simulations that are required to study the galaxy population (Muldrew et al. 2013). Debuhr et al. (2011) also propose a BH accretion model that depends on the AM. Specifically, the accretion rate is proportional to the mean gas surface density, the local sound speed squared and the inverse of the rotational angular frequency. They apply their model in isolated major mergers, finding self-regulation of the resulting BH but no clear evidence of suppression of the star formation in the resulting galaxy.

We will present a model that is similar in spirit to that of Debuhr et al. (2011). We propose a simple scheme that takes into account the AM of the gas flow, but does not require such high resolution so that it can be used in simulations of the cosmological population. The subgrid model that we arrive at is suitable for inclusion in simulations of galaxy formation, and thus allows us to revisit the interaction of BH feedback and galaxy formation. We show that including the AM has a profound impact on the behaviour of BH accretion, reproducing the behaviour postulated in semi-analytic models. In contrast to the common interpretation, however, the BH accretion shapes the mass function not because the efficiency of BH *feedback* varies with halo mass, but because BH accretion is strongly suppressed in cold star forming discs (relative to the rate

estimated when AM is not accounted for). We show that the new model not only matches the observed galaxy stellar mass fractions, but also generates accretion patterns that strongly resemble quasars in lower mass haloes and radio galaxies in the highest mass haloes that we probe.

The strategy of the paper is as follows. Most cosmological simulations estimate BH accretion rates on the basis of the Bondi–Hoyle–Lyttleton accretion model (Bondi & Hoyle 1944). We begin Section 2 by summarizing this model and discuss the critical length and critical time-scales in Bondi accretion. In the following subsection, Section 2.1.2, we motivate a simple extension of the Bondi–Hoyle–Lyttleton that allows the model to account for the AM of material flowing through the Bondi radius. This model contains an uncertain parameter that accounts for the effective viscosity of the disc/torus that forms as the flow circularizes: for plausible values of this parameter, the accretion rate may be suppressed by many orders of magnitude relative to the BS09 rate.

In Section 3, we take on the challenge of the implementing the extended model as a subgrid physics module suitable for cosmological scale simulations. To illustrate the impact of including AM in the calculation, we perform a suite of cosmological simulations. In Section 4, we describe the simulation code, the initial conditions and the simulations used in this study. In Section 5, we compare the impact of BH accretion in the Bondi-like accretion model used by BS09 and the AM-dependent extension of this model. We investigate the accretion history of the most massive BHs in the simulations in Section 6. Finally, we summarize our results and discuss the fundamental implications of the BH accretion model in Section 7. The conclusions of the paper are presented in Section 8. Convergence tests and parameter variations can be found in Appendix A and B, respectively. Throughout we adopt the Λ-cold-dark-matter cosmology with $h = 0.704$, $\Omega_\Lambda = 0.728$, $\Omega_\mathrm{m} = 0.272$, $\Omega_\mathrm{b} = 0.0455$, $\sigma_8 = 0.810$ and $n_\mathrm{s} = 0.967$. These values were derived from the *Wilkinson Microwave Anisotropy Probe* 7 year (*WMAP*7) data (Komatsu et al. 2011).

2 A BH ACCRETION MODEL THAT ACCOUNTS FOR ANGULAR MOMENTUM

BH accretion models are almost universally based on the Bondi accretion rate or its extension, usually referred to as the Bondi–Hoyle–Lyttleton rate, that extends the model to the case of a BH moving relative to the background particles (Bondi & Hoyle 1944). The net circulation of gas around the BH is not taken into account in either formulation. In this section, we describe a simple, physically motivated extension that accounts for the net AM of the surrounding gas.

2.1 The physical scales of BH accretion

In the absence of AM, the accretion rate of the BH is determined by the gas density and effective sound speed at the Bondi radius, r_B. This radius, defined as the radius at which the BH gravity dominates over the thermal and turbulent pressure of the surrounding gas, is a key physical length scale in the problem. It is given by

$$r_\mathrm{B} = \frac{GM_\mathrm{BH}}{c_\mathrm{s}^2} \approx 430 \left(\frac{M_\mathrm{BH}}{10^7\,\mathrm{M}_\odot}\right)\left(\frac{c_\mathrm{s}}{10\ \mathrm{km\,s^{-1}}}\right)^{-2} \mathrm{pc}, \quad (1)$$

where M_BH is the mass of the BH and c_s is the sound speed of the surrounding medium. In the Bondi estimate of the accretion rate, gas within r_B falls directly into the BH. However, when the infalling gas has a net AM, the flow will instead form an accretion disc or torus with characteristic radius, r_circ. For any significant AM, this radius will be larger than the last stable orbit of the BH. The radius of the disc is determined by the circularization radius of the flow

$$r_\mathrm{circ} = \frac{j^2(r_\mathrm{circ})}{GM_\mathrm{BH}}, \quad (2)$$

where j is the specific AM of material in a circular orbit at the circularization radius. The estimate assumes that the mass of the BH dominates within the circularization radius. In order to be accreted by the BH, material within this radius must transfer its AM outwards faster than it is converted into stars or expelled from the accretion region by feedback processes.

To produce a simple model, we assume that the accretion processes can be treated in two parts: first an almost radial infall from the Bondi radius to the circularization radius and then a slower flow of material through the disc to the last stable orbit. We express the specific AM of the flow passing through r_B in terms of the Bondi radius and the net tangential velocity (or circulation speed, V_ϕ) of gas at the Bondi radius, $\boldsymbol{j} = \boldsymbol{r_\mathrm{B}} \times \boldsymbol{V_\phi}$. Assuming angular momentum is conserved in the infall phase, the circularization radius is

$$r_\mathrm{circ} = \frac{r_\mathrm{B}^2 V_\phi(r_\mathrm{B})^2}{GM_\mathrm{BH}} = GM_\mathrm{BH}\frac{V_\phi(r_\mathrm{B})^2}{c_\mathrm{s}^4}$$
$$\approx 430 \left(\frac{M_\mathrm{BH}}{10^7\,\mathrm{M}_\odot}\right)\left(\frac{V_\phi(r_\mathrm{B})}{10\ \mathrm{km\,s^{-1}}}\right)^2\left(\frac{c_\mathrm{s}}{10\ \mathrm{km\,s^{-1}}}\right)^{-4} \mathrm{pc}. \quad (3)$$

This formulation implies that $(r_\mathrm{circ}/r_\mathrm{B}) \leq 1$, since material at the Bondi radius is assumed to be falling into the BH. This implies $(V_\phi/c_\mathrm{s}) \leq 1$ *at the Bondi radius*. On scales larger than the Bondi radius, it is entirely possible that the circulation speed will be greater than the sound speed. This will be the case, for example, in a rotationally supported cold gas disc.

To get a sense of the values of the Bondi and circularization radii relative to the Jeans scale of the interstellar medium (ISM) we include Fig. 1. The left-hand panel illustrates the variation of r_circ (red lines) and r_B (blue lines) with the circular speed and the effective sound speed (or equivalently, the *effective temperature*) of the surrounding gas.[1] We limit coloured lines to the region $(V_\phi/c_\mathrm{s}) \leq 1$, where the circularization radius is smaller than the Bondi radius as required for infalling material.

The left-hand panel in Fig. 1 shows r_B and r_circ for two values of the BH mass M_BH at 4×10^8 (solid lines) and $4 \times 10^7\,\mathrm{M}_\odot$ (dotted lines) and two values for the circular speed V_ϕ, 10 and 100 $\mathrm{km\,s^{-1}}$. Both the Bondi and circularization radii vary from thousands of parsecs when the sound speed is 10 $\mathrm{km\,s^{-1}}$ (gas temperature $\sim 4 \times 10^3$ K) to less than 1 parsec when the sound speed is 10^3 $\mathrm{km\,s^{-1}}$ (gas temperature $\sim 4 \times 10^7$ K). Although both radii decline as the sound speed increases, the circularization radius shows a much more dramatic variation so that the importance of the accretion disc phase is reduced relative to the free-fall phase of the flow.

It is useful to compare the Jeans length λ_Jeans ($\sim$2.2 kpc, black star) for the gas at the star formation density threshold ($n_\mathrm{H}^* = 0.1\,\mathrm{cm^{-3}}$) with an effective temperature 10^4 K.[2] For a BH with

[1] We define the effective sound speed as $c_\mathrm{s} \equiv (\frac{\gamma k_\mathrm{B} T}{\mu m_\mathrm{H}} \frac{1}{f_\mathrm{th}})^{1/2}$ where f_th is the fraction of pressure which is thermal. In our simulations, we assume $f_\mathrm{th} = 1$ but adopt an effective gas temperature that is given by the ISM equation of state (see BS09). We set $\mu = 0.59$.

[2] There is no unique definition of the Jeans length. This usually differs by a factor which depends on the geometry of the object. We

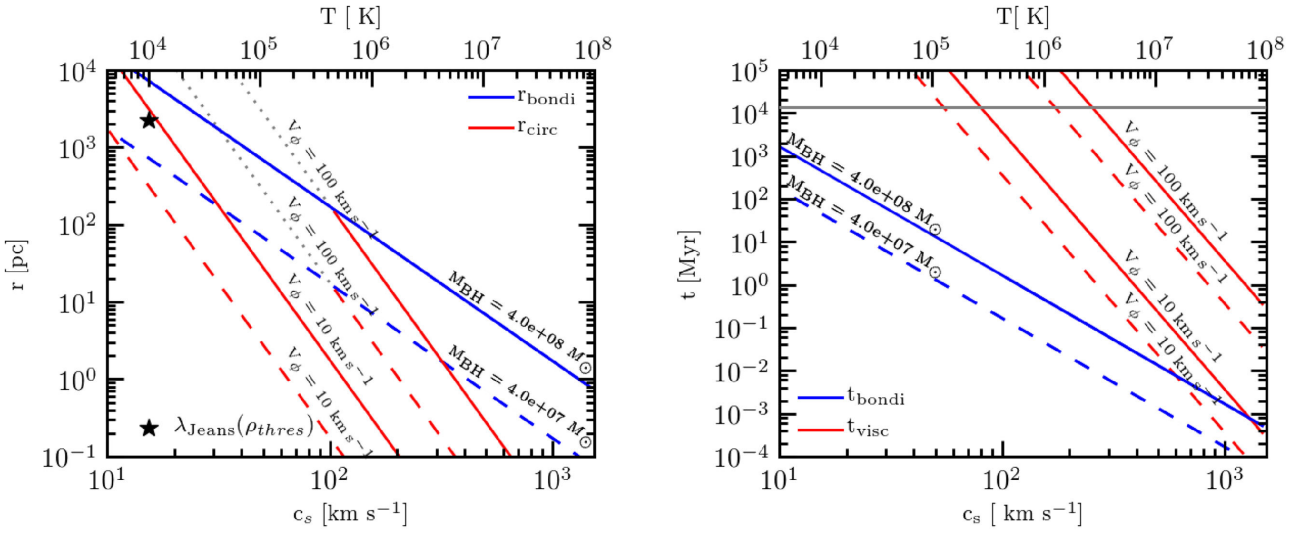

Figure 1. The left-hand panel shows the Bondi (blue lines) and viscous (red lines) radii as a function of the sound speed for $M_{\rm BH} = 4 \times 10^7\,{\rm M}_\odot$ (dashed lines) and $M_{\rm BH} = 4 \times 10^8\,{\rm M}_\odot$ (solid lines) and for $V_\phi = 10$ and $100\,{\rm km\,s^{-1}}$. The grey dotted lines show the unstable regime in which our assumption that $(V_\phi/c_{\rm s}) \leq 1$ breaks down. Although both radii decrease with the sound speed and increase with $M_{\rm BH}$, the circularization radius is a stronger function of the sound speed. The star marker represents $\lambda_{\rm Jeans}$ in a gas at $T = 10^4$ K and at the SF threshold density ($n_{\rm H}^* = 0.1\,{\rm cm^{-3}}$). The simulations are designed so that this is well resolved. The right-hand panel compares the Bondi time (blue lines) and viscous time (red lines) for our fiducial model with $C_{\rm visc} = 2 \times 10^6$ (equation 8). At low values of $c_{\rm s}$, the viscous time can be as long as the age of the Universe, but it reaches much shorter values when the sound speed is high. The Bondi time also decreases with high sound speed but the effect is weaker: even for low circular speed ($10\,{\rm km\,s^{-1}}$), the accretion rate is limited by the viscous time-scale unless the sound speed is as high as $c_{\rm s} \sim 10^3\,{\rm km\,s^{-1}}$.

mass comparable to the Jeans mass ($2 \times 10^7\,{\rm M}_\odot$), the Bondi radius is just less than a kpc, similar to the Jeans length. In practice, however, gas surrounding the BH will usually be far denser than the star formation threshold. Turbulence, induced by star formation (through the winds of hot stars and SNe), will greatly increase the effective sound speed and consequently reduce the Bondi radius. Assuming the net circulation speed is independent of radius, the smaller Bondi radius implies a lower AM for the inflowing gas and consequently a smaller circularization radius.

2.1.1 *The Bondi time-scale and the viscous time-scale*

In the previous section, we showed that the presence of AM introduces an additional spatial scale into the accretion model, corresponding to the radius at which infalling material circularizes. This creates an additional time-scale that is dependent on the AM of the infalling material.

In the case that the gas around the BH does not have any net circulation, the time-scale for gas accretion is determined by the sound-crossing time at the Bondi radius:

$$t_{\rm Bondi} = \frac{r_{\rm B}}{c_{\rm s}} = \frac{GM_{\rm BH}}{c_{\rm s}^3} \approx 42 \left(\frac{M_{\rm BH}}{10^7\,{\rm M}_\odot}\right)\left(\frac{c_{\rm s}}{10\,{\rm km\,s^{-1}}}\right)^{-3}{\rm Myr}. \quad (4)$$

In the presence of a net circulation, an accretion disc or torus will be formed at the circularization radius. The disc viscosity will cause gas to spiral inwards to the last stable orbit where it will finally be accreted. To estimate this time-scale, we will assume that the disc formed is thin (i.e. that the gas cools efficiently) and that the radial motions are small in comparison to the rotation. In this case, gas follows roughly Keplerian orbits and the disc scaleheight H is given by

$$\frac{H}{R} \sim \frac{c_{\rm s}'(r_{\rm circ})}{v_{\rm circ}'(r_{\rm circ})} = \frac{V_\phi}{c_{\rm s}}\frac{c_{\rm s}'}{c_{\rm s}} \ll 1, \quad (5)$$

where $v_{\rm circ}'$ is the circular velocity within the disc, and $c_{\rm s}'$ is its sound speed. The second equality follows from conservation of specific AM between the circularization radius and the Bondi radius [i.e. $r_{\rm circ} v_{\rm circ}'(r_{\rm circ}) = r_{\rm B} V_\phi(r_{\rm B})$] using equations (1) and (3). To be consistent with these assumptions and that of a thin disc, $c_{\rm s}' \ll c_{\rm s}^2/V_\phi$, which follows if there is strong turbulence in the ISM. Transport through the disc can be described by a diffusion equation where the kinematic viscosity ν is usually parametrized as

$$\nu = \alpha_{\rm visc}\, c_{\rm s}' H, \quad (6)$$

with $\alpha_{\rm visc}$ as a dimensionless number (Shakura & Sunyaev 1973). The viscous time can then be expressed as

$$t_{\rm visc} = [\alpha_{\rm visc}(H/R)^2]^{-1} t_{\rm dyn} \sim \frac{r_{\rm circ}^2}{\nu}, \quad (7)$$

where $t_{\rm dyn}$ is the dynamical time-scale of the disc. The values for H/R lie in the range 0.1–0.001 for a thin accretion disc in the α-disc parametrization (Shakura & Sunyaev 1973) and values of $\alpha_{\rm visc}$ lie in the range ∼0.1–0.3 from observational evidence (Buat-Menard, Hameury & Lasota 2001; Cannizzo 2001a,b; Schreiber et al. 2004; King, Pringle & Livio 2007). Using equations (2), (3), (6) and (7), the viscous time becomes

define $\lambda_{\rm Jeans} \equiv (\frac{c_{\rm s}^2}{G\rho})^{1/2} = c_{\rm s} n_{\rm H}^{*\,-1/2} (\frac{X f_{\rm gas}}{G\mu m_{\rm H}})^{1/2}$, where $f_{\rm gas} = 0.3$ is the gas mass fraction and $X = 0.752$ is the hydrogen mass fraction. Setting $c_{\rm s} \sim 15.3\,{\rm km\,s^{-1}}, \lambda_{\rm Jeans} \sim 2.2\,{\rm kpc}$. We define the corresponding Jeans mass as $M_{\rm Jeans} = 4/3\pi\rho\left(\frac{\lambda_{\rm Jeans}}{2}\right)^3$ that takes the value $2 \times 10^7\,{\rm M}_\odot$.

$$t_{\rm visc} = C_{\rm visc}\frac{r_{\rm circ}}{v_{\rm circ}} = C_{\rm visc}\frac{j^3}{G^2M_{\rm BH}^2}$$
$$= C_{\rm visc}\frac{r_{\rm B}^3V_\phi^3}{G^2M_{\rm BH}^2} = C_{\rm visc}GM_{\rm BH}\frac{V_\phi^3}{c_{\rm s}^6}$$
$$\approx 42\,C_{\rm visc}\left(\frac{M_{\rm BH}}{10^7\,{\rm M}_\odot}\right)\left(\frac{V_\phi}{10\ {\rm km\,s^{-1}}}\right)^3\left(\frac{c_{\rm s}}{10\,{\rm km\,s^{-1}}}\right)^{-6}{\rm Myr}. \quad (8)$$

For a Shakura–Sunyaev disc, $C_{\rm visc} = 2\pi[\alpha_{\rm visc}(H/R)^2]^{-1}$. If we fully understood the transport processes in the rotating disc or torus that feeds the BH, we would simply insert the appropriate values for $\alpha_{\rm visc}$ and (H/R). For a thin disc model, plausible values of $C_{\rm visc}$ lie in the range 10^8–10^3. However, the structure of the circularization disc we are considering is completely unclear, as is the dominant viscous mechanism on relevant scales. These issues are discussed extensively in the literature (e.g. Shlosman, Begelman & Frank 1990, Hopkins & Quataert 2010; King, Zubovas & Power 2011; Power et al. 2011) because the long time-scales implied by the Shakura–Sunyaev formulation make it hard to understand the high efficiency of accretion required to create supermassive BHs at very high redshift. On the scales relevant to our simulations, the appropriate transport mechanism is likely to be gravitational instability (Hopkins & Quataert 2010, 2011) rather than the magnetorotational instability. But these simulations do not accurately treat the multiphase structure of the ISM and stellar feedback, and the value of $C_{\rm visc}$ is therefore extremely uncertain. We adopt an empirical approach by varying $C_{\rm visc}$ to obtain the best match to observed galaxy properties, as discussed in Appendix B. The analysis prompts us to choose $C_{\rm visc} = 2 \times 10^6$ as a fiducial value, although a wide range of values give qualitatively similar results.

The right-hand panel of Fig. 1 shows $t_{\rm Bondi}$ (blue lines) and $t_{\rm visc}$ (red lines) as a function of the effective sound speed (or equivalent gas temperature) for BH masses $M_{\rm BH} = 4 \times 10^7$ M$_\odot$ and $M_{\rm BH} = 4 \times 10^8$ M$_\odot$ and $C_{\rm visc} = 2.1 \times 10^6$ under the same region of our assumption($V_\phi/c_{\rm s} \leq 1$). The Bondi time and viscous time are both decreasing functions of the sound speed, but the viscous time has a stronger dependence on the sound speed. For instance, for a BH with $M_{\rm BH} = 4 \times 10^8$ M$_\odot$ surrounded by gas with $c_{\rm s} \sim 15.3$ km s^{-1}, the Bondi time-scale is $\sim$1 Gyr, but this drops by six orders of magnitude (to $\sim 10^3$ yr) when $c_{\rm s} \sim 10^3$ km s^{-1}. It is important to note, however, that although the accretion time-scale is much shorter when $c_{\rm s}$ is large, the overall accretion rate depends on the local density: at fixed pressure, the accretion rate declines with increasing $c_{\rm s}$. For $t_{\rm visc}$, the sound speed dependence is even stronger: For a circular speed of 10 km s^{-1}, the viscous time is $\sim$ 10 Gyr ($\sim$ the age of the Universe) for $c_{\rm s} \sim 90$ km s^{-1}, and $\sim 10^3$ yr at $c_{\rm s} \sim 10^4$ km s^{-1}. In contrast, both time-scales are relatively weak functions of the BH mass. The figure particularly highlights that the mass accretion rate will be mostly predominantly limited by the viscous time because of the stronger dependence of $t_{\rm visc}$ on the local sound speed.

2.1.2 *BH accretion rate accounting for angular momentum*

Material with sufficiently low specific AM falls through the Bondi radius forming an AM supported disc or torus. This leads to a dependency of the viscous time and circularization radius on the effective sound speed and the circular velocity of gas flowing through the Bondi radius. In this section we propose a revised estimate of the mass accretion rates of BHs that takes gas circulation into account.

The long viscous time-scale of the accretion disc creates a bottleneck for accretion on to the BH. As matter flows through the Bondi radius it will pile up at the circulation radius. If the viscous time-scale is long, this matter will form a nuclear starburst with only a small fraction being accreted into the BH. The challenge is to estimate the mass of the accretion torus/disc, and thus the accretion rate. Our subgrid implementation requires an estimate of the instantaneous accretion rate: gas that has stalled will be represented by macroscopic simulation particles (star formation and winds are handled by other components of the simulation code). The Bondi time sets the characteristic time-scale of the problem. In the absence of circulation, the mass within the Bondi region is $\sim \dot{m}_{\rm Bondi}t_{\rm Bondi}$. We assume that the viscous bottleneck results in this mass building up in the disc/torus and then draining into the BH at a lower rate:

$$\dot{m}_{\rm BH} \sim \frac{\dot{m}_{\rm Bondi}t_{\rm Bondi}}{t_{\rm visc}}. \quad (9)$$

The constant of proportionality is degenerate with the proportionality constant $C_{\rm visc}$. Following this argument, the critical factor is, then, the ratio of the Bondi and viscous times:

$$\frac{t_{\rm Bondi}}{t_{\rm visc}} = \frac{r_{\rm B}c_{\rm s}^{-1}}{C_{\rm visc}[r_{\rm B}V_\phi]^3[GM_{\rm BH}]^{-2}}$$
$$= \frac{1}{C_{\rm visc}}\frac{c_{\rm s}^3}{V_\phi^3}. \quad (10)$$

Thus $t_{\rm Bondi}/t_{\rm visc}$ depends only on $[c_{\rm s}/V_\phi]^3$. Note that if $t_{\rm visc}$ is larger than $t_{\rm Bondi}$ (i.e. $C_{\rm visc}^{1/3}V_\phi > c_{\rm s}$) the accretion rate is limited by the Bondi rate and we can ignore the time spent in the accretion disc phase. We therefore write the BH accretion rate as

$$\dot{m}_{\rm BH} = \begin{cases} \dot{m}_{\rm Bondi}\left[\frac{1}{C_{\rm visc}}\left(\frac{c_{\rm s}}{V_\phi}\right)^3\right] & \text{if } C_{\rm visc}^{1/3}V_\phi > c_{\rm s}, \\ \dot{m}_{\rm Bondi} & \text{otherwise}, \end{cases} \quad (11)$$

where $\dot{m}_{\rm Bondi}$ is the mass accretion rate given in the equation (12) corresponding to the Bondi–Hoyle–Lyttleton accretion model (following Bondi & Hoyle 1944 to allow for the bulk gas motion as well as gas circulation) and $C_{\rm visc}$ is a free parameter which parametrizes the efficiency of AM transport and mass-loss from the disc. We take $C_{\rm visc} = 2.1 \times 10^6$ as our fiducial value.

The circular speed of material at the Bondi radius needs to exceed a critical value before the accretion time-scale is suppressed. In the case of our fiducial, $C_{\rm visc} = 2.1 \times 10^6$, this critical value is given by $V_\phi \approx c_{\rm s}/128$. For sound speeds of 10^3 km s^{-1}, the critical value of V_ϕ is 10 km s^{-1} and for low sound speeds of 10 km s^{-1} it is 0.1 km s^{-1}. Below this critical value, even though a disc/torus forms, transport through the disc is more rapid than the rate at which material flows through the Bondi radius.

3 BH ACCRETION MODELLING IN COSMOLOGICAL SIMULATIONS

Most cosmological simulations include accretion on to central BHs via a subgrid model based on the Bondi–Hoyle–Lyttleton accretion model (Bondi & Hoyle 1944), possibly adding a coefficient that attempts to account for the multiphase nature of gas around the BH that is not resolved at the finite resolution of the simulation (e.g. Di Matteo et al. 2005; Springel et al. 2005; BS09). These models assume that the net circulation of gas in the neighbourhood of the BH can be neglected. In this section we will discuss how these models can be extended to account for the circulation of gas. We focus our description on the model of BS09, but the extension can equally be applied to other implementations of the Bondi–Hoyle–Lyttleton accretion model. It is inherent in the nature of submodels

that the density, effective sound speed and circulation of gas at the Bondi radius will need to be estimated from coarse scale quantities that are resolved in the simulation.

3.1 The OWLS BH accretion model

Following Springel et al. (2005), a BH seed with a mass $m_{\rm seed}$ is injected into each new Friends-Of-Friends (FOF) halo that exceeds the mass threshold, $m_{\rm halo,\,min}$. In order to prevent the BH from being ejected by two-body encounters, the BH seed is assumed to track the position of the minimum potential in the halo. Subsequently, BHs grow by two processes: mergers with other BHs or accretion of nearby gas particles. The gas accretion obeys a modified version of the Bondi–Hoyle–Lyttleton formula:

$$\dot{m}_{\rm Bondi} = \alpha \frac{4\pi G^2 M_{\rm BH}^2 \rho}{(c_{\rm s}^2 + v^2)^{3/2}}, \tag{12}$$

where $M_{\rm BH}$ is the mass of the BH, $c_{\rm s}$ is the sound speed, ρ density of the local medium and v is the velocity of the BH relative to the ambient medium. The coefficient α is a dimensionless efficiency parameter that is introduced to account for the multiphase nature of the gas around the BH. The complex density structure of the ISM cannot be explicitly realized unless the simulations have a resolution significantly better than a few pc (Creasey, Theuns & Bower 2012). In the original Springel et al. (2005) model, α is a fixed number that ensures that the BHs grow rapidly during galaxy mergers. The authors set $\alpha = 100$. In BS09, the structure of the dense star-forming ISM is treated using an imposed polytropic equation of state that matched on to the ideal gas law at a threshold density, $n_{\rm H}^*$ (Schaye & Dalla Vecchia 2008). Gas at lower densities is assumed to be single phase, and the macroscopic density to therefore be a good estimate of the gas density at the Bondi radius of the BH. BS09 therefore extend the model, so that α becomes a function of local density:

$$\alpha = \begin{cases} 1 & \text{if } n_{\rm H} < n_{\rm H}^* \\ \left(\frac{n_{\rm H}}{n_{\rm H}^*}\right)^{\beta} & \text{otherwise} \end{cases}. \tag{13}$$

BS09 show that the parameter β does not play a crucial role in the growth of BHs which is usually limited by the AGN feedback. We set the parameter β to 2 which is the fiducial value in the OWLS project. We will refer to equations (12) and (13) as the Bondi-like model in what follows.

The subgrid accretion model also assumes that the BH accretion rate cannot exceed the Eddington accretion rate limit,

$$\dot{M}_{\rm Edd} = \frac{4\pi G m_{\rm p} M_{\rm BH}}{\sigma_{\rm T}\, c\, \epsilon_{\rm r}}, \tag{14}$$

where $m_{\rm p}$ is the proton mass, $\sigma_{\rm T}$ is the Thomson cross-section for the scattering of free electrons and $\epsilon_{\rm r}$ is the fraction of the energy liberated per accreted rest mass energy and is set at 0.1. The final accretion rate is

$$\dot{m}_{\rm BH} = \min(\dot{m}_{\rm Bondi}, \dot{M}_{\rm Edd}). \tag{15}$$

3.2 Accounting for AM-dependent accretion on the subgrid model

In order to implement the AM-dependent accretion rate as a subgrid model, we need to estimate the circulation speed of the gas surrounding a BH using the SPH smoothing kernel. We determine the weighted AM of the surrounding gas and then divide by the smoothing length h:

$$V_\phi = \left| \sum_{i=0}^{N_{\rm SPH}} \boldsymbol{r}_i \times \boldsymbol{v}_i\, m_i W(\boldsymbol{r}_i, h) \frac{1}{\rho h} \right|, \tag{16}$$

where $W(r, h)$ is the SPH smoothing kernel used in an SPH code, h is the SPH smoothing length of the BH and ρ is the smoothed density given by $\rho = \sum_{i=0}^{N_{\rm SPH}} m_i W(r_i, h)$.

The cosmological simulations almost never resolve the Bondi radius, so the subgrid scheme must extrapolate the circularization speed that is measured to smaller scales, just as it is needed to extrapolate the resolved gas density in the neighbourhood of the BH to the density at the Bondi radius. Following a similar approach, $V_\phi(r_{\rm B})$ should be extrapolated from the resolved circulation as follows:

$$V_\phi(r_{\rm B}) = (\frac{r_{\rm B}}{h})^{\gamma} V_\phi(h), \tag{17}$$

where γ can, in principle, be positive or negative depending on the relative mass of the BH and the structure of the surrounding gas. For simplicity, we will set $\gamma = 0$ in this paper. This assumes a flat rotation curve in the region around the BH and those processes that disrupt the rotation of the accretion disc on resolved scales would also disrupt the circulation on the scale of the Bondi radius. Note that the choice of γ is largely degenerate with $C_{\rm visc}$, so that if a different value of γ was chosen, $C_{\rm visc}$ would have to be recalibrated.

In a cold, rotationally supported gas disc, $(V_\phi/c_{\rm s}) \geq 1$ on the smallest scales that we are able to resolve. In this situation, the circularization lies outside the Bondi radius and we continue to apply equation (8) as the viscous time-scale. Clearly in future work it would be possible to develop a more complex model in which the viscous time-scale had a more complex dependence on the spatial scale of the circularization radius. Our present approach is intended to be as simple as possible.

The subgrid model provides an estimate of the instantaneous accretion rate. Following Springel et al. (2005), BH particles are described by two masses, a subgrid mass used in the accretion rate calculation and a particle mass used in gravitational calculations. This allows the subgrid BH to be less massive than the particle mass used in the simulation, but ensures that gravitating mass is strictly conserved. If the BH subgrid mass exceeds the particle mass, a particle within the BH neighbourhood is stochastically selected and its mass added to the particle mass of the BH particle. This allows the two BH masses to track each other. The surrounding gas particles continue to form star and potentially generate outflows.

Calculating the circular velocity using equation (16) gives stable values of V_ϕ even though the smoothing length and AM may fluctuate wildly. In order to set an upper limit on the numerical fluctuations in V_ϕ, we calculate the relative change in V_ϕ, $\mathrm{d}V_\phi/V_\phi$, between consecutive time steps. Fig. 2 shows the combined distribution for each BH with final mass $>10^6\,{\rm M}_\odot$. The distribution of $\mathrm{d}V_\phi/V_\phi$ is well characterized by a Cauchy distribution; we use half the 16–84 percentile range to determine a Gaussian-equivalent σ of 0.20. The distribution includes physically driven changes in the accretion rate, but the width of the core allows us to set an upper limit on the numerical noise arising from the SPH averaging of less than ∼20percent. This variation is likely to be dominated by physical gas flows (as opposed to numerical noise) driven, for example, by secular evolution and feedback from stars and the BH. We also find that the estimated value of the circular speed is largely independent of the number of neighbours used in the smoothing kernel. Furthermore, in Appendix A1, we show that the values of V_ϕ and $c_{\rm s}$ converge well with increasing particle number in idealized galaxy

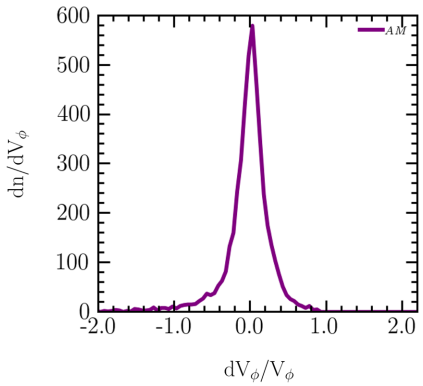

Figure 2. The normalized distribution of fractional fluctuations of V_ϕ, dV_ϕ/V_ϕ for the BHs in the simulation AM (see Section 4). In order to place an upper limit on the numerical noise in the measurement of V_ϕ, we compute the fractional change between consecutive time steps. This variation is an upper limit on numerical nose since the fluctuations may be dominated physical variations in the gas flows near to the BH (driven, for example, by BH and stellar feedback). The form of the histogram is described well by a Cauchy distribution with a Gaussian equivalent σ of 0.20.

simulations. Note that our approach is very different from Debuhr et al. (2011), who estimate the circular speed using the gravitational mass only. This difference is important: variations in the circular speed (driven by local gas conditions) result in distinctive accretion patterns for stable discs and merger induced outbursts. The circular speed is then used to compute the viscous time-scale using equation (8), and thus the modified accretion rate using equation (11). The accretion rate is also constrained to be smaller than the Eddington accretion rate (see equation 14). Other aspects of the implementation are identical to BS09 which we will describe briefly in Section 4.

4 THE NUMERICAL CODE AND HYDRODYNAMIC SIMULATIONS

The simulations we present are based on the GADGET-3 SPH code (Springel 2005), adding enhancements to reduce the simulation viscosity when the time derivative of the flow divergence is small (Cullen & Dehnen 2010) and to ensure that time steps of particles receiving feedback energy are limited (Durier & Dalla Vecchia 2012). These enhancements will be described in more detail in Dalla Vecchia (in preparation). The code uses an extensive network of subgrid modules to account for the turbulent pressure of the ISM, and implements star formation following the empirical Kennicutt–Schmidt law (Schaye & Dalla Vecchia 2008), chemical enrichment (Wiersma et al. 2009b) and cooling tracking 11 elements (Wiersma et al. 2009a). These modules were originally developed for the OWLS (Schaye et al. 2010) and GIMIC (Crain et al. 2009) simulation projects. Here we use the metallicity-dependent gas density threshold of (Schaye 2004, as in OWLS model SFTHRESZ) and a revised treatment of the equation of state (Dalla Vecchia & Schaye 2012). Feedback from stars is implemented as stochastic thermal energy injection events, using a fixed heating temperature of $10^{7.5}$ K, in order to avoid spurious radiative losses (Dalla Vecchia & Schaye 2012, see also Creasey et al. 2011). We moderate the SN feedback efficiency, $f_{\rm E}$ (the fraction of SN energy available to perform feedback) as a function of the local dark matter velocity dispersion (Oppenheimer & Dave 2006; Okamoto, Gao & Theuns 2008) in order to obtain a good match to the faint end slope of the galaxy SMF.

Table 1. A list of the simulations used in this paper. Each simulation has the same SN feedback parameters, and a comoving volume of (25 Mpc)^3. The simulation use 2×360^3 particles, with initial baryonic particle mass $1.4 \times 10^6\,h^{-1}\,{\rm M}_\odot$ and dark matter mass particle $6.3 \times 10^6\,h^{-1}\,{\rm M}_\odot$; the mass of seed BHs is set $m_{\rm seed} = 10^4\,h^{-1}\,{\rm M}_\odot$ and the minimum halo mass in which BH seeds are injected is $10^{10}\,h^{-1}\,{\rm M}_\odot$. The columns show (1) name of the simulation; (2) efficiency with which energy emitted by a BH is coupled into the ambient gas; (3) radiative efficiency of BH accretion disc.

Name (1)	$\epsilon_{\rm f}$ (2)	$\epsilon_{\rm r}$ (3)	Accretion model (4)
NO-AGN	–	–	–
BS09	0.15	0.1	Bondi-like accretion model equations (12) and (13).
BS09-LE	0.015	0.1	Bondi-like accretion model equations (12) and (13).
AM	0.15	0.1	AM accretion model equation (11), $C_{\rm visc} = 2 \times 10^6$.
Appendix B	0.15	0.1	AM accretion model, $C_{\rm visc} = 6 \times 10^4$–$6 \times 10^6$.

A similar variation is commonly used in semi-analytic models (e.g. Guo et al. 2011; Bower et al. 2012) and is supported by small-scale simulations of galaxy winds such as Creasey et al. (2012). The feedback efficiency $f_{\rm E}$ is set as a function of the equivalent halo virial temperature, varying between 1.0 (in low-temperature haloes) and 0.1 (in high-temperature systems), following the equation

$$f_{\rm E} = 1.0 - 0.9 \left(\frac{1}{1 + \exp\left[-2(\log T - \log T_0)\right]} \right). \quad (18)$$

After some experimentation, we set the transition temperature to $T_0 = 10^{5.5}$ K.

In this paper, our focus is the interaction between feedback from BH accretion and the formation of galaxies. We fix the SN feedback efficiency and its scaling and vary the parameters of BH feedback. Throughout the paper, for AGN feedback we use the stochastic heating model of BS09 with a heating temperature of 10^8 K, and a density power $\beta = 2$ in equation (13). By default we assume that the energy liberated per accreted rest mass energy is $\epsilon_{\rm r} = 0.1$, and that the heating efficiency (this is the fraction of liberated energy coupled to the surrounding gas medium) is $\epsilon_{\rm f} = 0.15$; we show the effect of reducing the heating efficiency in Section 5. Accretion is always limited to be less than the Eddington accretion rate. BH seeds are inserted into haloes that exceed an FOF halo mass of $10^{10}\,h^{-1}\,{\rm M}_\odot$, corresponding to ∼1500 dark matter particles. Such haloes are well defined and there is no danger of injecting BHs into spurious systems (BS09). BHs are injected with an initial mass of $10^4\,h^{-1}\,{\rm M}_\odot$. They are allowed to merge within their SPH smoothing kernel and have relative velocity less than 0.5 of the sound speed of the surrounding gas. In Section 3.2, we describe the enhancement of the standard accretion model to account for the circular speed of gas within the smoothing kernel of the BH.

The simulations presented in this paper are summarized in Table 1. All the simulations use the same subgrid parameters to describe star formation and stellar feedback, and only differ in the BH subgrid physics. There are in total four simulations:

(i) NO-AGN: this simulation does not include BH physics.

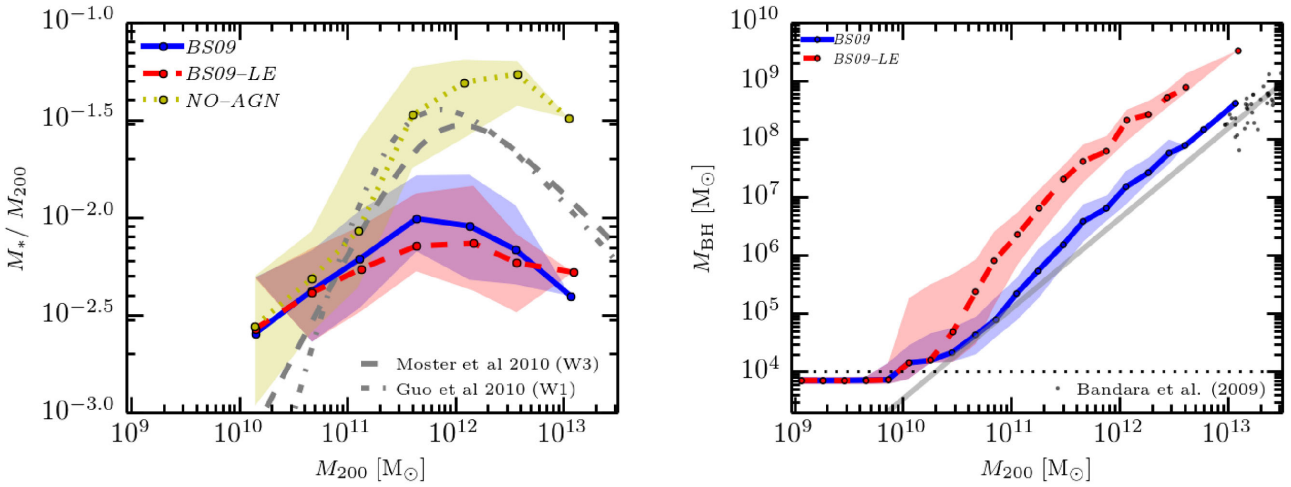

Figure 3. The left-hand panel shows the stellar mass fraction as a function of M_{200}, while the right-hand panel shows the $M_{\rm BH}$–M_{200} relation. The lines show the medians of M_*/M_{200} (or $M_{\rm BH}$) for the simulations NO-AGN (yellow, dotted line), BS09 (blue, solid line) and BS09-LE (red, dashed line); see Table 1. The shaded region shows the 10–90 percentile range of the data. The grey dashed and dash–dotted lines in the left-hand panel represent abundance matching results derived from Moster et al. (2010) and Guo et al. (2010), respectively, and the grey line in the right-hand panel indicates the observational $M_{\rm BH}$–M_{200} relation derived from observations by Bandara, Crampton & Simard (2009). Their observational data is shown as grey points. The figure shows that efficient BH self-regulation creates a strong correlation between BH mass and halo mass, but strongly suppresses the formation of stars across a wide range of halo mass. Although changing the efficiency of BH feedback alters the normalization of the $M_{\rm BH}$–M_{200} relation, it has little effect on the host galaxy properties.

(ii) BS09: we use BS09's model with the feedback efficiency set to $\epsilon_{\rm f} = 0.15$ in the Bondi-like accretion equations (12) and (13). Accretion is also limited by the Eddington rate.

(iii) BS09-LE: we set this simulation as the above except that the feedback efficiency is set to $\epsilon_{\rm f} = 0.015$.

(iv) AM: this simulation uses the AM-dependent accretion model equation (11), with $C_{\rm visc} = 2.1 \times 10^6$.

The results are not particularly sensitive to the moderate choice of $C_{\rm visc}$ (by two orders of magnitude), and we show some example variations in Appendix B. All of the models assume that the accretion rate cannot exceed the Eddington limit (equation 14).

Initial conditions were generated using second-order Lagrangian perturbations (Jenkins 2010) at a starting redshift of 127. We use initial gas and dark matter particle masses of 1.4×10^6 and $6.3 \times 10^6\,h^{-1}\,{\rm M}_\odot$, respectively; the *comoving* gravitational softening lengths are set to $2\,h^{-1}$ kpc with a maximum physical softening of $0.5\,h^{-1}$ kpc. The simulations are carried out in periodic boxes of 25 Mpc on each side. The largest structure that forms in the simulation has a mass of $1.65 \times 10^{13}\,{\rm M}_\odot$ at $z = 0$. Comparing with large simulation volumes we find that a box of this size does not noticeably bias galaxy properties as a function of virialized halo mass (Haas et al. 2013), allowing us to efficiently test the impact of different subgrid models. Because of the limited box size, we compare galaxy properties to the stellar fractions ($f_{\rm star} = M_*/M_{200}$) inferred from observations as a function of halo mass, rather than to the galaxy SMF itself.

5 THE GLOBAL IMPACT OF BH ACCRETION ON GALAXIES

In this section we examine the impact of BH feedback on the galaxies in the simulations described above. We focus on the stellar mass fraction–M_{200} (halo mass) relation and the $M_{\rm BH}$–M_{200} relation, in particular. We first present results based on the Bondi-like models (BS09 and BS09-LE), and then contrast them with the AM-dependent model.

5.1 Simulations with BS09 Bondi-like accretion models

First, we look at the simulations based on the Bondi-like accretion models, / and BS09-LE, and compare these with the NO-AGN simulation. The left-hand panel of Fig. 3 shows the stellar mass fraction, $M_{\rm star}/M_{200}$ (where the stellar mass is measured within a 30 kpc radius), for central galaxies as a function of M_{200}. Given the small box size, this provides a convenient way to compare with observational data. The solid lines represent the median relation in bins of halo mass, with the 10–90 percentile range of the data indicated as coloured regions. Grey lines show observational estimates based on abundance matching from Moster et al. (2010) and Guo et al. (2010).

The NO-AGN simulation (yellow colour) is consistent with the abundance matching results up to a halo mass of $10^{12}\,{\rm M}_\odot$, suggesting that in the physical Universe the feedback from AGNs has little or no effect in low-mass haloes. The lack of AGN feedback leads to an overproduction of stellar mass in above $10^{11.5}\,{\rm M}_\odot$ haloes in disagreement with the abundance matching results. Fig. 3 also shows the BS09 (blue line) and BS09-LE (red line) simulations. In both simulations, the impact of BH feedback has a strong effect on a wide range of halo mass, causing a reduction of the stellar mass fraction by a factor of 2 in $10^{11}\,{\rm M}_\odot$ haloes mass and an order of magnitude in the high haloes mass at $10^{12}\,{\rm M}_\odot$ relative to observations. The models are incompatible with the observational data across the full halo mass range probed.

We note that the effects of AGN feedback become significant at lower masses and are more severe than in the low-redshift OWLS simulations, which agree well with observations (McCarthy et al. 2010), because our particle masses are lower by nearly two orders of magnitude (allowing us to convincingly resolve galaxies with stellar masses greater than $10^9\,{\rm M}_\odot$ with 500 particles).

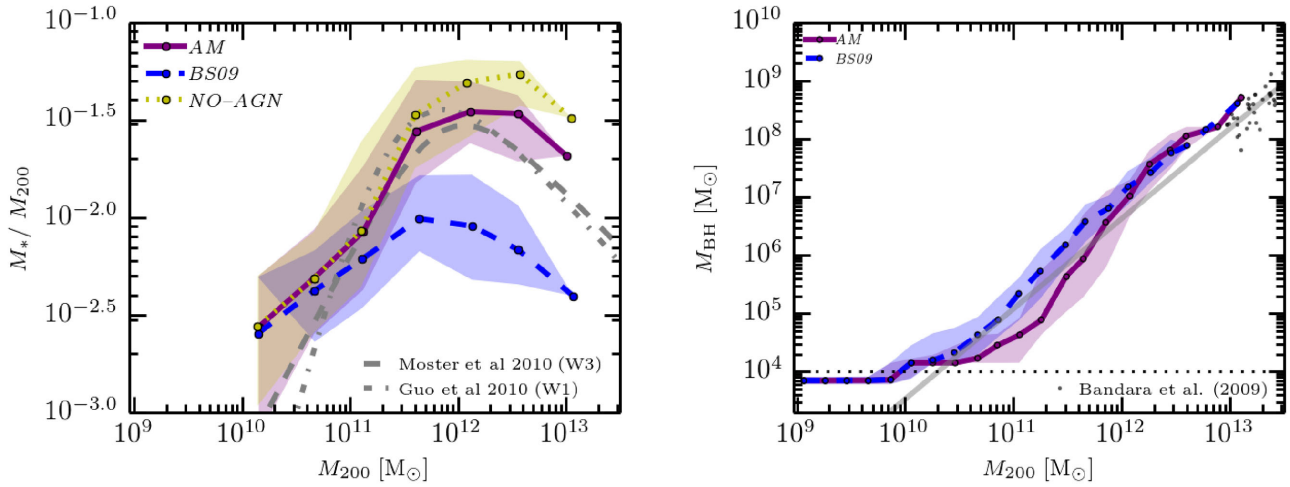

Figure 4. The left-hand panel shows the correlation of stellar mass fraction with M_{200} for the AM-dependent accretion model, contrasted with two of models considered in Section 5.1. The right-hand panel shows the $M_{\rm BH}$–M_{200} relation for the same models. The yellow dotted line and blue (here dashed) coloured lines correspond to those in Fig. 3. The AM simulation is shown in purple. The AM model results in resemble agreement with both the abundance matching relationship of the stellar mass fraction (grey lines) and with the observational $M_{\rm BH}$–M_{200} relation. The key to achieving this match is that BHs residing in haloes with mass below $\sim 10^{12}$ M$_\odot$ grow less than required for self-regulation.

Comparing the simulations BS09 and BS09-LE, where the efficiency of feedback differs by an order of magnitude ($\epsilon_{\rm f} = 0.15$ and 0.015, respectively), it is clear that the AGN feedback produces a similar suppression of star formation and low stellar mass fractions in both runs, which is consistent with the findings of BS09 and Booth & Schaye (2010).

The tendency for the BHs to self-regulate, and thus to correlate strongly with the binding energy of the (inner) halo has been emphasized by Booth & Schaye (2010). As a result, a strong correlation is expected between the BH mass and M_{200}. The right-hand panel of the Fig. 3 shows the BH mass as a function of M_{200} for the BS09 and BS09-LE simulations. In order to provide an observational baseline, black stars show estimates of the BH mass in 48 galaxy-scale strong gravitational lenses from the Sloan lens ACS (Bandara et al. 2009) and grey lines represent the fit of data from Bandara et al. (2009) based on the $M_{\rm BH}$–σ_* relation of Gultekin et al. (2009). The *BS09* simulation gives a correlation between the BH mass and halo mass in good agreement with observations as found by Booth & Schaye (2010). As expected, reducing the feedback efficiency by a factor of 10 produces more massive BHs at fixed halo mass, leading to a $M_{\rm BH}$–M_{200} relation that is offset from the observational data. As has already been shown by Booth & Schaye (2010), the total feedback energy remains the same in both models, consistent with the idea that the BHs grow until they begin to unbind the gas halo of the system. At this point they become self-regulating growing in synchronization with the halo mass.

The results from the two panels suggest that the self-regulated growth of the BH overwhelms the stellar mass growth of the central galaxy if BH accretion is efficient. The gas, which would have been able to form stars in the absence of AGN feedback, is expelled from the system or prevented from accreting, starving the central galaxy of fuel for further star formation. A comparison of the red and blue curves shows that the issue cannot be resolved by altering the efficiency with which energy is deposited. However, following the scenario discussed in Section 2, we have already noted that a vital component of the accretion model is missing: gas that is supported by AM cannot accrete as rapidly as the Bondi-like BS09 formula would suggest.

One unappealing solution is to only inject BHs into higher mass haloes i.e. suppressing the impact of BHs on lower mass galaxies by hand. For example, injecting BHs at a halo mass of $10^{11.5}\,h^{-1}$ M$_\odot$ provides a good description of galaxy properties. However, this is clearly unsatisfactory: at the resolution we consider, the haloes of $10^{10}\,h^{-1}$ M$_\odot$ contain more than 1500 particles and are well defined. There is no physical justification for not attempting to model BHs in these haloes.

Another simple solution is to adopt a much lower accretion rate. Decreasing β in equation (13) where the multiphase nature of the ISM is ignored, achieves the required goal of suppressing BH growth within galaxies. However, this strongly suppresses the accretion rates of early BHs making it difficult to form a BH population at high redshift as required by observations of massive BHs at redshifts $\sim$7 (Mortlock et al. 2012).

5.2 Simulation with the AM accretion model

In this section, we explore the effect of accounting for AM in the model. Fig. 4 is similar to Fig. 3, except that we show the AM simulation where the accretion rates account for AM as in equation (11), in purple. The left-hand panel shows that the AM simulation leads to stellar mass fractions that are in much better agreement with the abundance matching relation (grey lines) and still able to reproduce the turnover of the relation in high-mass haloes. Note that the stellar mass fraction found in haloes less massive than 10^{12} M$_\odot$ is similar to those in the NO-AGN simulation, supporting the idea that AGN feedback plays a minor role in regulating star formation in small galaxies. Looking at the $M_{\rm BH}$–M_{200} relation in the right-hand panel, we see that the BHs residing in haloes with masses lower than 10^{12} M$_\odot$ in the AM simulation grow more slowly than expected for self-regulation. However, above this halo mass, both models lock on to the same self-regulated relation, in good agreement with observations.

The model we have shown adopts $C_{\rm visc} = 2.1 \times 10^6$. The dependence of the outcome population on the value of this parameter is surprisingly weak – as illustrated in Appendix B. We will argue below that the mass scale at which BH accretion becomes

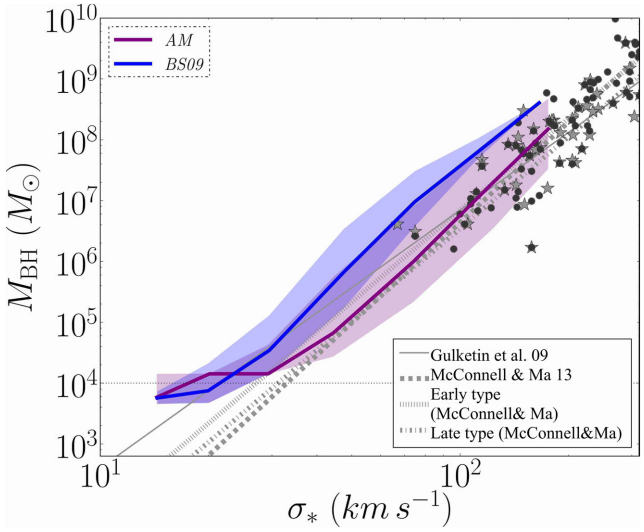

Figure 5. M_{BH}–σ_* relation for the models presented in Fig. 4. Line colours are reproduced from that figure. Observational data are shown in grey, taken from Gultekin et al. (2009) and McConnell & Ma (2013) with the grey solid and dashed lines the authors' respective observational fits. The grey dotted and dash–dotted lines are fits only considering early-type galaxies and late-type galaxies, respectively. The median relation from the AM simulation (purple line) is in good agreement with the fit expected for the early-type galaxies of McConnell & Ma (2013) at the high σ_* end while it encompasses the observational data for late-type galaxies at the lower σ_* values.

self-regulating is set by the properties of the halo rather than by the details of the AM accretion model.

While the BHs in lower mass haloes fall below the (extrapolated) observational relation, the correlation between BH mass and halo mass is a very indirect observational test of the models. A better approach is to compare the BH mass to the stellar velocity dispersion of the model. This is shown in Fig. 5. Since the models have significantly different stellar mass and dynamics, we would expect the models to differ in this plot, and this is indeed seen. Observational data is shown from Gultekin et al. (2009) and McConnell & Ma (2013) with the grey solid and dotted lines, respectively, showing the expected median relations for early-type galaxies (dotted line) and late-type galaxies (dot–dashed line). The difference between observational relations arises because McConnell & Ma (2013) include data for bright galaxy centrals of a cluster resulting in a steeper fit. The AM model is clearly able to pass the test offered by the M_{BH}–σ_* relation.

6 THE ACCRETION HISTORY OF BLACK HOLES

6.1 A case study

In order to understand the transition between the low- and high-mass halo behaviour better, we plot the 'timeline' of a typical massive BH in the simulation in Fig. 6. The curves in the upper panel illustrate the mass of the most massive progenitor (red line, left-hand axis), its accretion rate (coloured dots and right-hand axis) and the Eddington accretion rate (red line, right-hand axis) as a function of redshift. The points plotted below the horizontal line represent BH accretion rates $\dot{m}_{BH} < 10^{-4}$ $M_\odot yr^{-1}$. The colour code corresponds to $\log_{10}(C_{visc}) + 3\log_{10}(V_\phi/c_s)$ [where $\log_{10}(C_{visc}) = 6.3$]. This quantity is the factor by which the mass accretion rate has been suppressed by taking AM into account. For comparison, the FOF mass and the total stellar mass of the host halo are shown in the lower panel as black and green lines, respectively (left-hand axis), along with the total star formation rate as the blue line (right-hand axis).

The BH is injected with a seed mass of 1.4×10^4 $M_\odot$. Initially BH growth is limited by the Eddington accretion rate, but the BH growth soon settles down. Although the BH undergoes episodes of accretion close to the Eddington limit, the absolute mass growth is small at high redshift. Over time, the suppression factor increases in importance. Eventually, promoted by a merger event that begins at $z \sim 5$, the BH undergoes a more sustained period of Eddington-limited accretion. At the end of this event, the accretion rate drops due to the suppression by the AM. There is marginal further BH growth until $z \sim 3.0$ when another period of Eddington-limited accretion begins. The BH undergoes two further periods of Eddington-limited accretion until its host halo reaches a mass of 2×10^{12} $M_\odot$. It would seem reasonable to associate these high accretion rate events with quasar activity. During these high accretion rate periods, the suppression factor is much lower than during the quiescent periods in between. Below $z = 1$, the nature of accretion on to the BH appears to change again. Although the accretion rate fluctuates, it rarely reaches the Eddington limit, typically peaking at 1–10 per cent. The outburst are more frequent, but not sustained. The correspondence between colour and accretion rate suggests that the accretion is limited by the AM. The suppression factor becomes more important as redshift decreases.

The lower panel of Fig. 6 allows us to contrast the growth of the halo with that of the host galaxy of the BH. Since the plot shows the total halo mass in each case, it does not identify galaxy mergers well. However, it does emphasize the steady growth of the halo and its constituent galaxies, which is quite at odds with the episodic growth of the BH. Despite the periods of strong BH activity, the star formation rate of the halo seems little affected until the halo reaches a mass of 2×10^{12} $M_\odot$ at $z \sim 2$. In the following outburst, the star formation rate is clearly set back, with the outburst at $z \sim 1.1$ initiating a sustained decay.

The simplest way to understand what drives the sustained accretion events seen in Fig. 6 is to examine a movie of the evolution.[3] Six frames from the movie are shown in Fig. 7, illustrating some of the important features. Examining the sequence in detail, however, shows that the outburst typically lags behind the obvious morphological disturbance. The first three outbursts follow a clear pattern. Immediately after the outburst, the accretion leads to the expulsion of gas surrounding the BH. Although a strong wind is generated, the void which this leaves around the BH is quickly refilled by cool filaments that efficiently carry cool gas to the centre of the halo. The final outburst is qualitatively different, however: once the central gas is ejected, the surrounding matter does not flow in to replace it. Although there continues to be some cool material at the centre of the halo, this does not settle quickly around the BH. We can understand the change in behaviour from the nature of accretion in haloes of different mass. As Van de Voort et al. (2010) have demonstrated in the OWLS simulations (see also White & Frenk 1991; Kereš et al. 2005; Dekel & Birnboim 2006), the fuelling of high-mass haloes is qualitatively different from their low-mass counterparts; high-mass haloes are dominated by slow accretion from the hot halo, while lower mass haloes are supplied through rapid inflows along cold streams.

[3] Available from the website http://star-www.dur.ac.uk/~rgb/yrg2014/phase_movie.avi.

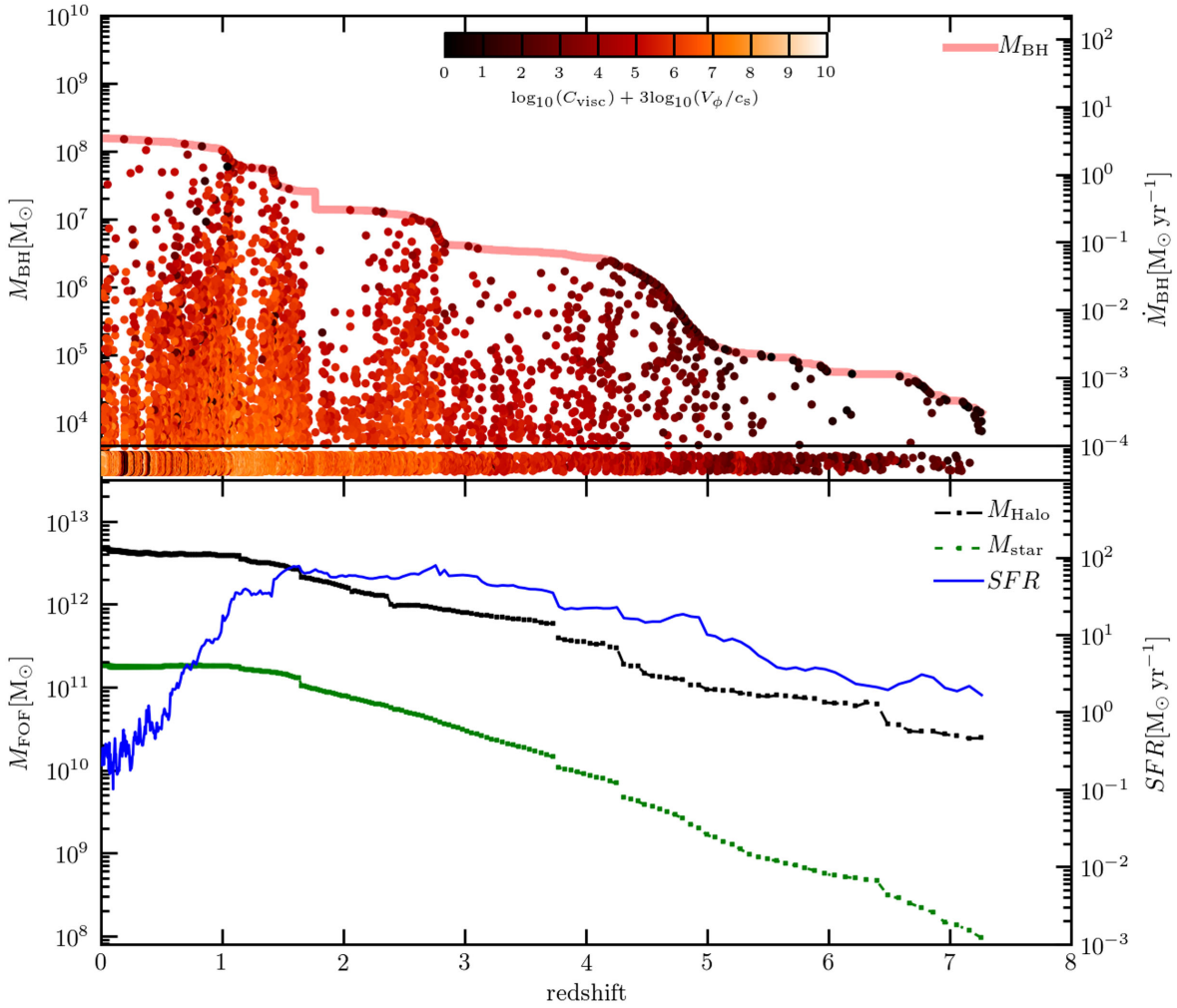

Figure 6. History of a massive BH and its host galaxy in the AM simulation. The top panel shows the growth of the second most massive BH at $z = 0$ as a function of redshift (red line and left-hand y-axis) as well as its instantaneous mass accretion rate (coloured points and right-hand axis). The colour code indicates the log suppression factor by which the mass accretion rate is reduced, $\log_{10}C_{\rm visc} + 3 \times \log_{10}V_\phi/c_{\rm s}$ where $\log_{10}C_{\rm visc} = 6.3$ for the simulation. High accretion rates shows small suppression factor. The red line represents simultaneously BH mass and Eddington accretion. Above $z = 6$, the BH frequently accretes at Eddington limit, but its mass and consequently its energy output are small. At intermediate redshift, the BH undergoes phases of high mass accretion rate($z \sim 4.7$, 2.75 and 1.1), interspersed with periods of quiescence. Below $z = 1$, the BH shows frequent outbursts reaching 1–10 per cent of the Eddington rate. The bottom panel shows the FOF mass (black long dash–dotted line) and total stellar mass (green dash–dotted line) of the host halo (on the left-hand y-axis) as a function of the redshift. The blue solid line represents the total star formation rate (right-hand y-axis).

6.2 BH accretion and its dependence on local gas properties

In order to better understand the interplay between the pressure and AM of the gas around the BH, we show mass accretion rates of the central BHs residing in the nine most massive haloes ($M_{200} > 10^{12.5}$ M$_\odot$ at $z = 0$) in Fig. 8. Each panel divides the plot into different redshift bins showing BH accretion rate as a function of the local effective fluid pressure, with points coloured by the log suppression factor as indicated in Fig. 6.

The points plotted below the horizontal line represent BH accretion rates of $\dot{m}_{\rm BH} < 10^{-4}$ M$_\odot$yr^{-1}. Points which are coloured dark red have a relatively low circular speed and low suppression factor; orange through yellow and white points, with higher circular speeds relative to local sound speed, are offset to much lower accretion rates. The diagram is somewhat complicated by points for which the accretion rate is Eddington limited, but the Eddington limit only significantly distorts the diagram in the highest redshift panel.

Fig. 8 confirms the behaviour seen for a single BH in Section 6.1. The accretion rates tend to increase strongly with the pressure of the surrounding gas. Most gas particles close to the BH lie on the ISM equation of state so that the pressure and the sound speed are closely correlated. In part, the trends in pressure arise because of the strong modulation of the accretion rate by the BS09 α parameter in equations (12) and (13). As the gas pressure increases in the simulation, the subgrid model assumes that the density distribution of the ISM becomes increasingly clumpy and cold. This trend is modulated by the viscous suppression factor (shown by the colouring of the points) since increasing pressures reduce the importance of the circulation speed (the increasing gas pressure reduces the Bondi radius with the result that the circulation radius and the viscous time-scale are also reduced). In addition to the main trend, a few points have a high pressure but a low suppression factor and relatively low accretion rates. These are episodes when the BH accretes from the hot (but low density) intragroup medium, rather than from a central pool of ISM.

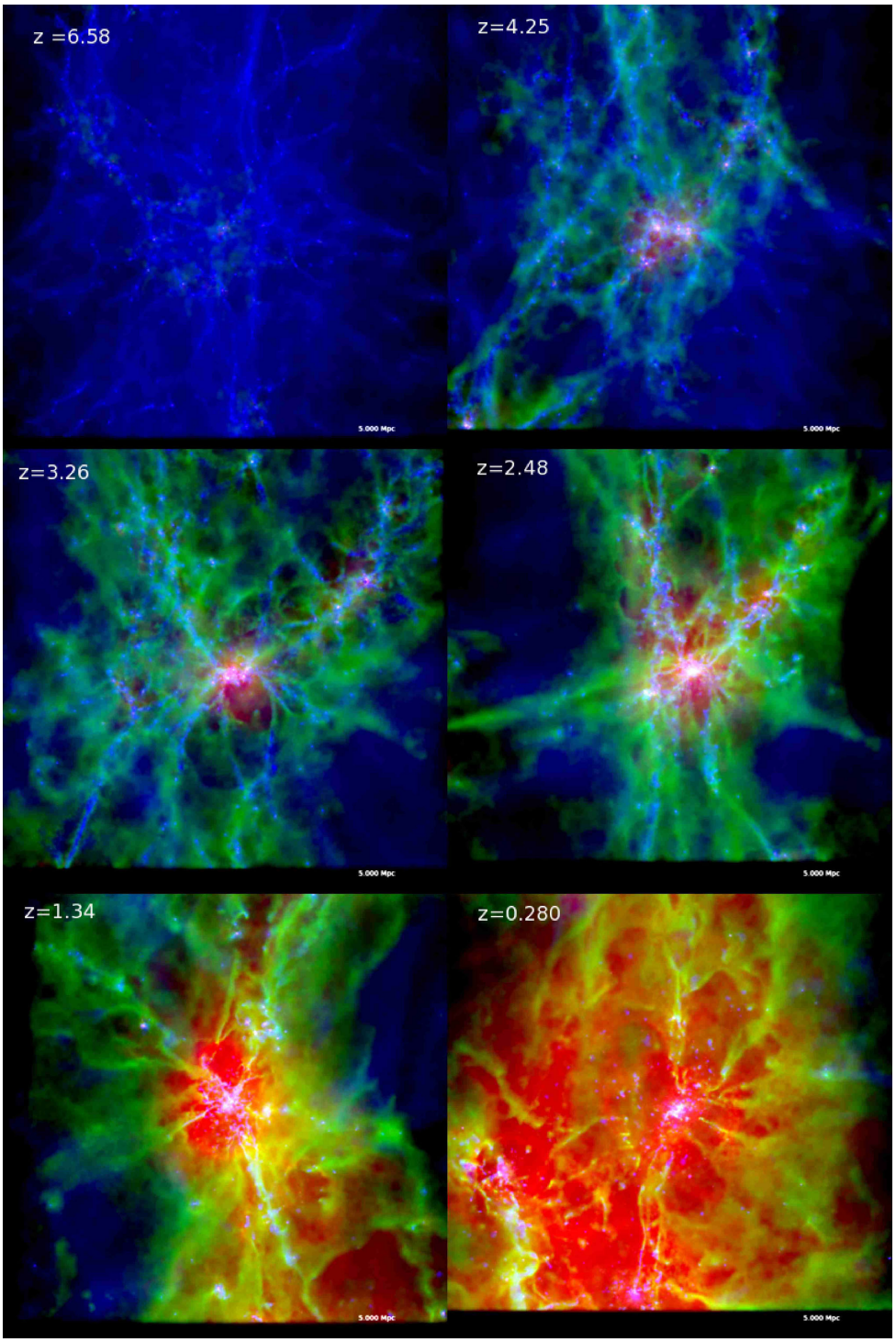

Figure 7. This sequence of images illustrates the evolution of the galaxy and halo of the BH shown in Fig. 6. The colour channels show the projected gas density of gas with temperature $T < 10^{4.5}$ K (blue), $10^{4.5} < T < 10^{5.5}$ K (green) and $T > 10^{5.5}$ K (red). The projected star density is shown in pink. The box above shows 5 Mpc on each side, representing 1/125 of the total simulation volume. The redshifts have been chosen to illustrate the change in the behaviour of AGN feedback. The complete movie can be downloaded from the website shown in footnote 3.

Focusing on the differences between the redshift panels, we see first that the typical accretion rate increases across the three first panels, but then gradually decreases. This confirms the trend seen for an individual BH in Fig. 6. In the highest redshift panels, BHs accrete close to their Eddington limits. In subsequent panels, the accretion rates remain similar (and decline slightly) while the BHs grow significantly in mass. Examination of Fig. 6 shows that two factors contribute. First, the pressure in the surrounding medium reaches a maximum in the intermediate redshift panels and then slowly declines to lower redshift. This possibly arises from the features of the BH's host galaxies: galaxies are less concentrated at lower redshifts as well as the galaxy gas fractions decline below

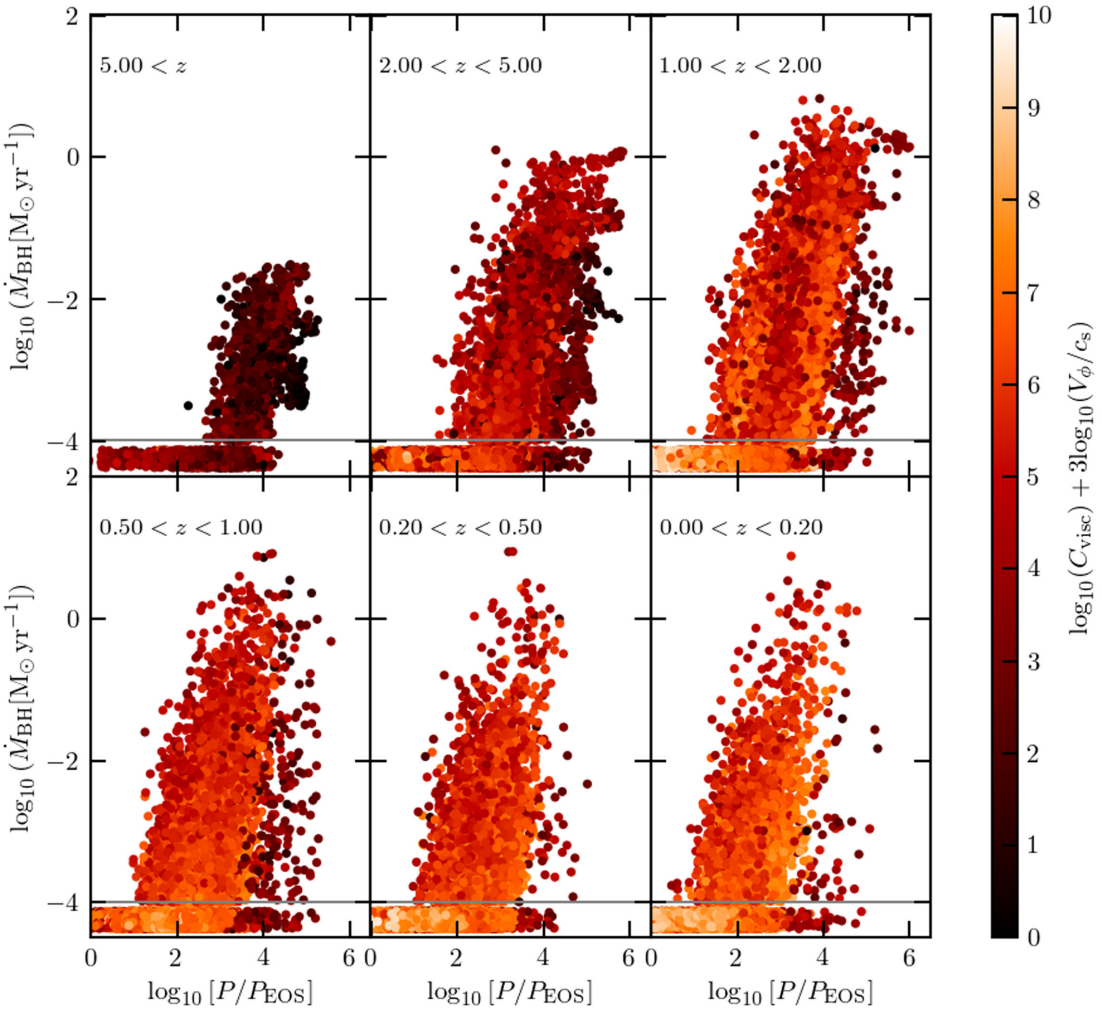

Figure 8. BH accretion histories plotted against the surrounding gas pressure (in units of $P_{\rm EOS,0}/k = 2.3 \times 10^3 {\rm cm}^{-3}\,{\rm K}$) for the most massive progenitor of the central BHs residing in the nine haloes with $M_{200} > 10^{12.5}\,{\rm M}_\odot$ at z=0. Each panel represents a different redshift bin as indicated in all panels. The colour scale indicates the $\log_{10}$ suppression factor as in Fig. 6. Significant accretion events occur when the suppression factor is small and the surrounding gas pressure is high.

$z = 1$. As a result, gas presses around the BH less tightly and tends not to accrete as strongly, despite the growth of BH mass. For a fixed BH mass, the accretion rate depends on pressure as $P^{15/8}$ (for an effective equation of state with polytropic index $\gamma_{\rm EOS} = 4/3$ and subgrid density parameter $\beta = 2$). This is indeed the slope seen in the figure. The suppression factor also becomes more important at lower redshift as clear from Fig. 8.

In Fig. 9 we show the dependence of the accretion rate suppression factor on galaxy stellar mass, with points coloured by the specific star formation rate of the host galaxy. The figure, shows that V_ϕ and $c_{\rm s}$ are usually comparable (blue solid line), but that galaxies tend to undergo excursions to much smaller suppression factors, as we have seen in the case study. The black dashed line shows the median suppression factor in bins of log stellar mass with a width of 0.3 dex. It is clear that the suppression factor has only a weak dependence on stellar mass with the accretion flow being only slightly less suppressed in the larger objects. This has important implications: the strong increase of BH mass as a function of halo mass is not the result of a direct dependence of V_ϕ on the halo mass. Instead, this dependence must rise from the reaction of the halo to feedback from the BH, or from a halo mass dependence of the gas pressure around the BH. We consider this further in Section 7.

6.3 Implications for observable AGNs

We will present an in-depth comparison of the model with observational data in a future paper. Here, we limit ourselves to a brief look at the Eddington ratio distribution of accreting sources for the *AM* model. In Fig. 10, we plot a histogram of the Eddington ratio distribution (defined as $\dot{M}_{\rm BH}/\dot{M}_{\rm Edd}$). We weigh each source by its accretion luminosity so that the histogram shows the contribution to the growth of the BH mass density as a function of both BH mass and Eddington ratio. We have chosen to plot the distribution at $z = 1$ (combining outputs in the range $z = 0.78$–1.12); similar plots can be constructed at other redshifts, but the features of the distribution functions remain similar.

The purple solid line shows the Eddington ratio distribution of all BHs: the integral under this plot gives the total mass growth rate of BHs at this epoch. Because, we have weighted the source contribution by their accretion luminosity, the highest Eddington

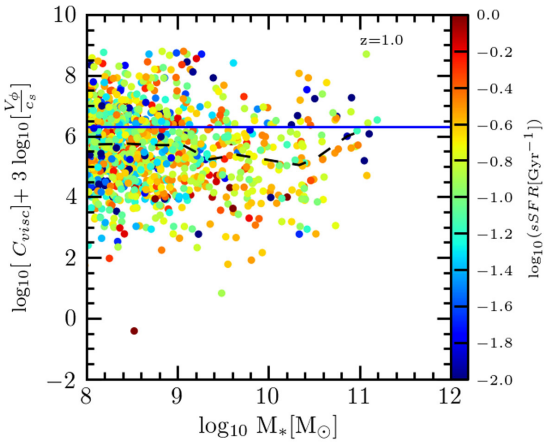

Figure 9. The dependence of the accretion suppression factor, $\log_{10}[C_{\rm visc}] + 3\log_{10}[V_\phi/c_{\rm s}]$, on galaxy stellar mass for the fiducial case, $\log_{10}C_{\rm visc} = 6.3$. The dashed line represents the median of the accretion suppression factor in bins of log galaxy mass with a width of 0.3. The blue line shows the value of $\log_{10}C_{\rm visc}$. The panel shows the relation at $z = 1$, but the plot is similar at other redshifts. In the majority of galaxies, V_ϕ and $c_{\rm s}$ are comparable and the accretion rate is suppressed by a factor $C_{\rm visc}$. In some galaxies, however, the suppression factor is much smaller and the AM of the accreting gas is no longer the limiting time-scale. There is little trend in the median suppression factor with galaxy mass, as illustrated by the black dashed line, or with specific star formation rate, as indicated by the colours of the points.

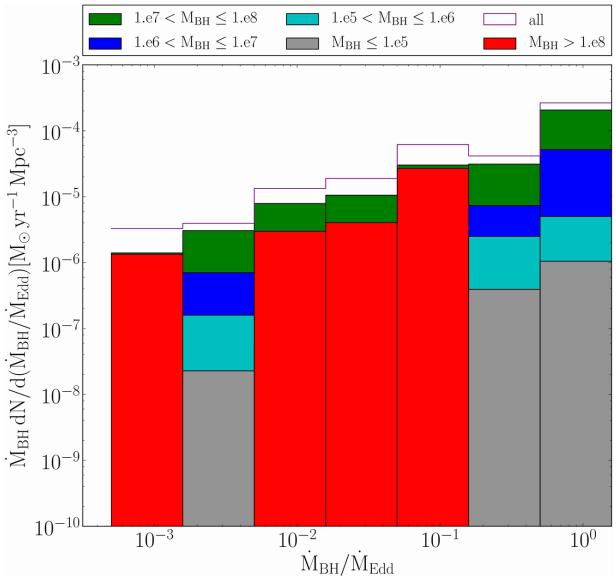

Figure 10. The distribution of Eddington ratio weighted by the mass accretion rate of each accretion source (i.e. the accretion luminosity). This plot shows the distribution at $z = 1$. The total distribution is shown as a purple line; The contributions of BHs in different mass ranges are illustrated by different colours. For BHs less massive than $10^8\,{\rm M}_\odot$, the distributions are similar, with larger BHs making a greater contribution to the BH mass budget. Above $10^8\,{\rm M}_\odot$, however, high Eddington ratio sources become rare.

ratio sources dominate the BH mass growth budget. If we had simply counted sources, the histogram would show that low Eddington ratio sources are far more numerous. It is instructive, however, to examine how this distribution is built up for BHs of different masses. Different colour lines distinguish the contributions from BHs of different masses, starting from the lowest mass BHs, below $10^5\,{\rm M}_\odot$ (grey region), and building up to show the contribution of the most massive BHs, above $10^8\,{\rm M}_\odot$ (red region). Because of the weighting by accretion rate (or equivalent to luminosity), the smallest BHs contribute little to the total. The shape of the Eddington ratio distribution is similar for BH masses up to $10^8\,{\rm M}_\odot$. The relative similarity of the distributions has previously been noted in observational data (Aird et al. 2012).

Above a BH mass of $10^8\,{\rm M}_\odot$, however, high Eddington ratio sources become rare, and the growth of these BHs is dominated by accretion events that are ∼0.1 of Eddington. Although there are only a handful of such high-mass BHs in the simulation, they make a significant contribution to the energy budget because of their mass. Studies of quasar clustering suggest that quasars do not occupy the most massive haloes at low redshift. Instead, massive haloes are occupied by radio galaxies (Romano-Diaz et al. 2010; Angulo et al. 2012; Fanidakis et al. 2013), whose power output tends to be in the form of a kinetic jet, but not in the form of highly visible luminosity. It is quite plausible that this dichotomy arises from a change in the Eddington rate distribution with BH mass (Meier 2001; Nemmen et al. 2006). We will test this aspect of the simulation against observations in more detail in a future paper using a larger simulation volume.

7 DISCUSSION

The simulations that we have presented in the previous sections show the importance of accounting for the AM of accreting gas. Compared to runs that implement a Bondi-like accretion formula, accounting for the AM of gas surrounding the BH results in less BH growth and greater stellar mass growth. This is particularly evident in haloes less massive than $\sim 10^{11.5}\,{\rm M}_\odot$. Compared to runs that do not include any feedback the revised BH accretion model leads to a sharp turnover in the stellar mass fraction of haloes.

In this section, we briefly consider the physical processes that establish this behaviour. In particular, we would like to understand what creates the transition in BH accretion efficiency at $10^{11}-10^{12}\,{\rm M}_\odot$. It is notable that this mass scale is weakly dependent on the accretion viscosity parameter $C_{\rm visc}$ (see Appendix B) which suggests than this mass scale is not caused by the subgrid accretion model. Moreover, this mass scale approximately corresponds to the transition between 'rapid cooling' haloes, in which the cooling time is short compared to the dynamical time, and 'hydrostatic' haloes in which the cooling time is long compared to the halo dynamical time. The key to understanding the nature of this transition is to examine the balance of gas accretion versus star formation and outflows. Since the mass locked up in stars is relatively small, the most important consideration is the outflow of gas. This may be driven by stellar feedback or AGN feedback from the growth of the BH. Recent papers have emphasized that outflows lead to self-regulation, both for feedback from star formation and from BHs (White & Frenk 1991; Bower et al. 2006, 2008; also Booth & Schaye 2010; Schaye et al. 2010; Dubois et al. 2013 and Puchwein & Springel 2013). The observed dependence of stellar mass fraction on halo mass (and hence the shape of the galaxy mass function) requires that star formation dominates the self-regulation at low-halo masses, while it is dominated by BHs at high stellar masses. The challenge is to understand why this switchover occurs.

We begin by considering the physical behaviour of Bondi-like accretion models. In the simulations presented here, we use a lower particle mass than BS09 to better sample star formation in lower mass galaxies, but the same ISM equation of state. Without

accounting for the AM, BH accretion is efficient across a wide range halo mass , (BS09; Booth & Schaye 2010). This results in the balance of gas inflow and outflow being regulated by BH growth at the expense of the growth of stellar mass. This drives the SMF to a power-law shape that is in complete contrast to the observed Schechter function (see Benson et al. 2003). The balance of BH and stellar mass growth cannot be redressed by altering the feedback parameters (as opposed to accretion model): reducing the efficiency of BH *feedback* results in greater BH growth but does not improve the efficiency with which inflowing material is converted into stars.

In contrast, including AM in the calculation reduces BH accretion rates in smaller haloes allowing the inflow and outflow in these haloes to be regulated by star formation. Above a halo mass of $10^{11.5}$ $M_{\odot}$, however, BH accretion becomes more efficient, creating a turnover in the stellar mass–halo mass relation.

We began this paper by stressing the role that AGN feedback plays in shaping the properties of galaxies. In semi-analytic models, the key to reproducing the observed dependence of the stellar mass fraction on halo mass is the link between effective AGN feedback and the halo cooling time. This results in the full impact of AGN accretion only being felt in haloes more massive than $\sim 10^{11.5}$ $M_{\odot}$. We have shown that the AM-dependent BH accretion model used in this paper has a similar overall impact, limiting BH self-regulation to the most massive haloes in the simulation.

But what sets the scale at which this occurs? One possibility is that the suppression factor $V_{\phi}/c_{\rm s}$ may be strongly dependent on system mass. However, we demonstrate that this is *not* the case in Fig. 9. The absence of a strong correlation in this plot shows that the rapid rise of BH mass with halo mass does not directly result from the dependence of the accretion rate on V_{ϕ}. The origin of this halo mass dependence is more subtle, and we consider other possibilities below.

We have already noted that the halo mass scale corresponds to the transition from the rapid cooling regime (in which accretion shocks cool rapidly) to the hydrostatic regime (in which the cooling time of gas is long compared to the dynamical time of the system). We can expect that this creates a difference in the mode by which gas is supplied to the central galaxy (White & Frenk 1991; Kereš et al. 2005; Dekel & Birnboim 2006; Van de Voort et al. 2010): in lower mass haloes, accreted gas flows almost directly on to the galaxy, fuelling a cold gas disc at the centre of the halo, while in more massive systems accreted gas will be shock heated to the virial temperature of the halo before slowly cooling out at the centre. These differences are evident in our simulations, but this cannot be the complete picture. The different modes of halo accretion may well explain why AGN feedback is much more effective at suppressing star formation in the central galaxies of hydrostatic haloes, but it does not explain why BHs accrete slowly in low-mass haloes, while they are able to self-regulate in massive haloes.

There are at least five mechanisms that might couple the accretion efficiency of low-mass objects to the mode of gas accretion and the structure of the surrounding halo. We consider each briefly below.

(i) A transition in the AM of the accreted material. Direct accretion of cool filaments may result in greater AM in the discs formed in lower mass haloes. Such a transition could be driven by the change from rapid cooling (with accretion occurring through misaligned streams) to more spherical accretion through the surrounding hot gas halo. Examining Fig. 9, however, shows that the halo mass dependence of this factor is extremely weak.

(ii) A transition in the merger rate. Efficient accretion is likely to be driven by mergers of stellar clumps that disrupt the AM of the gas disc, while it is not immediately evident why this would be strongly dependent on the halo mass (since the halo merger rate is weakly dependent on halo mass); and indeed this explanation is also ruled out by Fig. 9.

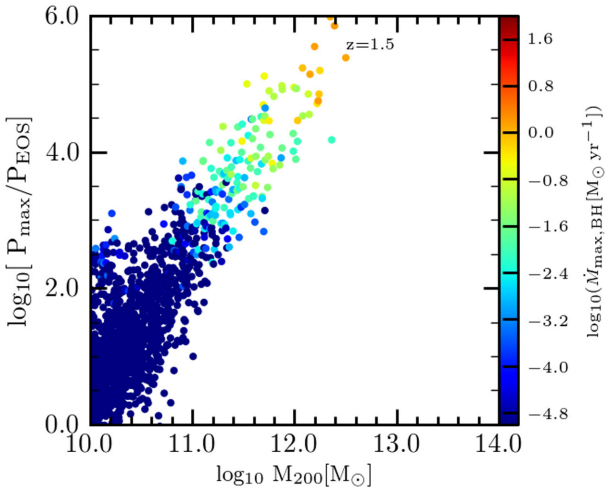

Figure 11. The pressure of gas surrounding the BH as a function of the halo mass. As can be seen from Fig. 9, there are large fluctuations in pressure between each snapshot output. In order to make the trend clear we plot the maximum pressure experienced by each BH in the redshift range 1.35–1.65. The pressure is given in terms of the equation of state threshold pressure $P_{\rm EOS,0}$. Points are coloured by the maximum BH accretion rate in the same redshift interval. As halo mass increases, the effective pressure of the gas surrounding the BH increases rapidly leading to greatly increased BH accretion rates.

(iii) A transition in the impact of the gas heated by the BH feedback. In low-mass haloes, material heated by the BH may simply leave the system without redistributing its energy effectively. In contrast, the energy injected into higher mass haloes may be effectively trapped and shared with the surrounding particles. In Fig. 7, we can clearly identify expanding shells of material in the massive haloes; however, when we trace the final location of particles as a function of halo mass (relative to the halo virial radius), no clear trend with halo mass is evident.

(iv) The ability of the halo to recover after an AGN feedback event. In lower mass haloes, the cooling time of inflowing material is short, and its filamentary geometry tends to limit the effect of a feedback outburst. Examination of individual events suggests that, in lower mass haloes, the outburst only affects the material in the galaxy briefly, and that the system is quickly refuelled by accretion from the surrounding cold streams. Accretion in larger haloes is more fragile, since the cooling time is long and the cool streams evaporate before reaching the central object: once the streams are disrupted by an AGN outburst, they struggle to re-establish themselves (Dubois et al. 2013). While this scenario explains the absence of star formation in the most massive galaxies, it does not account for the relatively weak growth of BHs in smaller haloes.

(v) The pressure dependence of the BH accretion rate. Fig. 8 shows that the accretion rate of the BH depends very strongly on the effective pressure of the surrounding gas. In Fig. 11, we show the halo mass dependence of the pressure. In order to smooth over fluctuations in the accretion rate, we plot the maximum accretion rate reached in the redshift interval $z = 1.35$–1.65 and colour points by the maximum BH accretion rate. A strong correlation can be seen, such that in lower mass haloes, the BH is surrounded by lower pressure gas. This leads to significantly lower BH mass accretion

rates. This trend re-enforces itself: the lower accretion rates lead to slower BH growth and hence to a lower accretion rate at a fixed gas pressure. In contrast, higher mass haloes generate higher pressures around the BH, leading to higher accretion rates and higher BH masses which, in turn, have higher accretion rates at a fixed pressure.

Qualitatively, this picture is confirmed by the images shown in Fig. 7, and in the linked animation (see footnote 3). During the early universe, haloes are able to recover quickly due to the strong filamentary nature of their accretion. Although there are sporadic bursts of energy injection from the BH, these do little to suppress the continuous inflow of gas filaments. At late times, however, the BH is able to couple effectively to the surrounding halo, disrupting filamentary gas accretion and establishing a hot gas envelope that appears to be pressure supported. The disruption of filaments is only part of the BH impact, however. While it may explain the suppression of gas inflow (and hence star formation) in high-mass haloes, it does not provide an explanation for the low accretion rates of BHs in low-mass haloes. The low growth rates of BHs in lower mass haloes can, however, be accounted for by looking at the pressure of the gas surrounding the BH. The pressure shows a strong increase with halo mass, greatly increasing the resulting BH accretion rates.

Fig. 9 shows that the suppression of the accretion rate due to the AM does not have a strong halo mass dependence. Over long time-scales, and averaged over the population of objects, the effect is to renormalize the accretion rate relative to the Bondi-like BS09 formula 12. The AM model does however, play an important role in the determining the duty cycle of energy input. As a result of the AM variations, high $\dot{m}_{\rm BH}$ accretion events are interspersed with quiescent periods during which the stellar component of the galaxy grows strongly, as shown in Fig. 6. The history of each individual BH is, however, complex and a combination of all the mechanisms may contribute at some level. The net effect is similar, but not identical, to the mechanisms invoked by semi-analytic models (Bower et al. 2006; Croton et al. 2006); significantly more work is required to explore the similarities and differences, and to capture the behaviour of the simulation in a simple but quantitative model.

8 CONCLUSIONS

We have presented a development of the subgrid accretion model of BS09, based on Springel et al. (2005), that takes into account the AM of the accreting gas using the local SPH kernel to estimate the circular speed at the Bondi radius and to define the circularization radius of the material passing through the Bondi radius. This creates an additional time-scale in the accretion of gas on to a BH which characterizes the transport of material through the disc and the fraction of this material ejected out of the disc before it is accreted by the BH. We incorporate this new time-scale into a revised accretion rate estimate that can be simply implemented as a subgrid model in cosmological simulations.

Two BH accretion models are tested: one using the prescription of BS09 and one with the updated AM model. The cosmological simulations include gas cooling, metal enrichment, star formation and SN feedback which we keep fixed. We show that simulations that do not account for the circulation of gas in the neighbourhood of the BH result in self-regulation of the gas supply by BH growth rather than star formation. As a result BH masses correlate well with the parent halo over a wide range of scales, but self-regulation by the BH results in insufficient star formation to match the observed dependence of stellar mass on halo mass in low-mass haloes.

In contrast, when the AM is taken into account, the accretion time-scale is increased and BH accretion rates are strongly suppressed in haloes less massive than $10^{11.5}\,{\rm M}_{\odot}$. This allows stars to form efficiently and for the balance of gas inflow and outflow to be set by stellar feedback. At higher halo mass, the effective pressure of gas surrounding the BH increases, compensating for the lower accretion rates of the AM model. As a result, a strong division into two regimes of galaxy formation emerges from the model. Below a halo mass of $\sim 10^{11.5}\,{\rm M}_{\odot}$, we find that the presence of the BH has little effect on forming galaxies, but above this halo mass star formation is suppressed and galaxies grow much more slowly in stellar mass compared to the growth of their haloes. The model broadly matches the observed stellar mass fractions of haloes and reproduces the expected correlation between the stellar velocity dispersion and BH mass (McConnell & Ma 2013). The distribution of BH accretion rates also seems compatible with that observed (Aird et al. 2012). The model thus provides a promising prescription for cosmological simulations.

We speculate that there are two critical factors that establish the break in the galaxy SMF. The first is the ability of the BH to accrete the material; accounting for the AM of the material surrounding the BH suppresses accretion from quiescent discs allowing the growth of the galaxy in galactic mass haloes to be regulated by star formation. The second is the response of the halo gas to outbursts of BH activity. This introduces a strong dependence on the ratio of cooling time to dynamical time of the halo through (i) the ability of the inflowing filaments to re-establish themselves following a brief episode of energy injection from the BH and (ii) the pressurization of the halo as a quasi-hydrostatic hot corona is established. In lower mass haloes the cold gas disc surrounding the BH quickly recovers because of the rapid supply of fresh material from the surrounding cosmic web. In high-mass haloes, filamentary accretion is disrupted by the hot corona. As a result, star formation in the most massive galaxies is strongly suppressed and the break in the SMF is established.

The BH model presented here forms the basis of the accretion model used in the EAGLE simulation project (Schaye et al. 2015). The EAGLE simulations use $\beta = 0$, which allows the α factor, discussed in Section 3.1, to be absorbed into the parameter $C_{\rm visc}$. The EAGLE simulations have larger volume than the work presented here allowing an examination of the observable predictions of the model, such as the evolution of the accretion rate distribution and the correlations between episodes of star formation and BH growth in more detail.

Although the current simulations suggest a promising way forward to include the AM of infalling gas in subgrid models of BH accretion, a great deal of work remains to be done. Larger volume calculations are required to determine accurately the break in the galaxy SMF and to probe the quasar luminosity function, since these are both determined by rare objects. Clearly, one of the most important future steps is to base the choice of the $C_{\rm visc}$ parameter on finer-scale simulations that simultaneously resolve the circularization radius and the multiphase structure and turbulence of the ISM around the BH. For the foreseeable future, such a multiscale approach seems the only feasible route to capture simultaneously the microphysics of BH accretion and the large-scale distribution of galaxies and quasars.

ACKNOWLEDGEMENTS

We thank James Mullaney, Nicolas Tejos, James Aird and Rob Thacker for a careful reading of the paper, useful comments

and discussions. This work would have not been possible without Lydia Heck's technical support. YMR-G gratefully acknowledges financial support from the Mexican Council for Science and Technology (CONACYT) (Studentship no. 213183). CSF acknowledges an ERC Advanced Investigator grant, COSMIWAY (GA 267291). This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. This work was supported by the Science and Technology Facilities Council (ST/F001166/1,ST/L00075X/1), European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement 278594-GasAroundGalaxies, the Netherlands Organization for Scientific Research (NWO) and Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office [AP P7/08 CHARM]. We thank contributors to SCIPY,[4] MATPLOTLIB[5] and the PYTHON programming language.[6]

REFERENCES

Aird J. et al., 2012, ApJ, 746, 90
Anglés-Alcázar D., ÃÜzel F., Davé R., 2013, ApJ, 770, 5
Angulo R. E., Springel V., White S. D. M., Cole S., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 425, 2722
Antognini J., Bird J., Martini P., 2012, ApJ, 756, 116
Balsara D. S., 1995, J. Comput. Phys., 121, 357
Bandara K., Crampton D., Simard L., 2009, ApJ, 704, 1135
Barai P. et al., 2013, MNRAS, 430, 3213
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Birzan L., Rafferty D. A., Nulsen P. E. J., McNamara B. R., Rottgering H. J. A., Wise M. W., Mittal R., 2012, MNRAS, 427, 3468
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Booth C. M., Schaye J., 2009, MNRAS, 398, 53 (BS09)
Booth C. M., Schaye J., 2010, MNRAS, 405, L1
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
Bower R. G., McCarthy I. G., Benson A. J., 2008, MNRAS, 390, 1399
Bower R. G., Benson A. J., Crain R. J., 2012, MNRAS, 422, 2816
Buat-Menard V., Hameury J.-M., Lasota J.-P., 2001, A&A 369, 925
Cannizzo J. K., 2001a, ApJ, 556, 847
Cannizzo J. K., 2001b, ApJ, 561, L175
Crain R. A. et al., 2009, MNRAS, 399, 1773
Creasey P., Theuns T., Bower G. R., Lacey C. G., 2011, MNRAS, 415, 3706
Creasey P., Theuns T., Bower G. R., 2013, MNRAS, 429, 1922
Croton D. J. et al., 2006, MNRAS, 365, 11
Cullen L., Dehnen W., 2010, MNRAS, 408, 669
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
DeBuhr J., Quataert E., Ma C. -P., 2011, MNRAS, 412, 1341
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dubois Y., Devriendt Y., Slyz A., Teyssier R., 2010, MNRAS, 409, 985
Dubois Y., Pichon C., Devriendt J., Silk J., Haehnelt M., Kimm T., Slyz A., 2013, MNRAS, 428, 2885
Durier F., Vecchia D. C., 2012, MNRAS, 419, 465
Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, MNRAS, 436, 315
Gultekin K. et al., 2009, ApJ, 698, 198
Guo Q., White S., Li C., Boylan–Kolchin M., 2010, MNRAS, 404, 1111
Guo Q. et al., 2011, MNRAS, 413, 101
Haas M. R., Schaye J., Booth C. M., Vecchia C. D., Springel V., Theuns T., Wiersma P. C., 2013, MNRAS, 435, 2931
Hernquist L., 1990, ApJ, 356, 359
Hopkins P. F., Quataert E., 2010, MNRAS, 407, 1529
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1
Hopkins P. F., Hernquist L., Cox T. J., Keres D., 2008, ApJS, 175, 356
Jenkins A., 2010, MNRAS, 403, 1859
DavéKereš D., Katz N., Fardal M., Weinberg D. H., 2005, MNRAS, 353, 2
King A. R., Pringle J. E., Livio M., 2007, MNRAS, 376, 1740
King A. R., Zubovas K., Power C., 2011, MNRAS, 415, L6
Komatsu E. et al., 2011, ApJS, 192, 18
Krumholz M. R., McKee C. F., Klein R. I., 2005, ApJ, 618, 757
Ma C. J., McNamara B. R., Nulsen P. E. J., 2013, ApJ, 763, 63
McCarthy I. G., Schaye J., Ponman T. J., Bower R. G., Booth C. M., Dalla Vecchia C., Crain R. A., Springel V., 2010, MNRAS, 406, 822
McCarthy I. G. et al., 2011, MNRAS, 412, 1965
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
McNamara B. R., Nulsen P. E. J., 2007, ARA&A, 45, 117
Magorrian J. et al., 1998, ApJ, 115, 2285
Meier D. L., 2001, ApJ, 548, L9
Mortlock D. J., Patel M., Warren S. J, Hewett P. C., Venemans B. P., McMahon R. G., Simpson C. J., 2012, MNRAS, 419, 390
Moster B. P., Somerville R. S., Maulbetsch C., v. d. Bosch F. C., Maccio' A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Muldrew S. I., Pearce F. R., Power C., 2013, MNRAS, preprint (arXiv:1306.4327)
Mullaney J. R. et al., 2012, ApJ, 753, 2
Nemmen R. S., Bower R. G., Babul A., Storchi-Bergmann T., 2007, MNRAS, 377, 1652
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Oppenheimer B. D., Dave R., 2006, MNRAS, 373, 1265
Power C., Nayakshin S., King A., 2011, MNRAS, 412, 269
Puchwein E., Springel V., 2013, MNRAS, 428, 2966
Romano-Diaz E., Schlosman I., Trenti M., Hoffman Y., 2010, ApJ, 736, 66
Schaye J., 2004, ApJ, 609, 667
Schaye J., Dalla Vechia C., 2008, MNRAS, 383, 1210
Schaye J. et al., 2010, MNRAS, 402, 1536
Schaye J. et al., 2015, MNRAS, 446, 521
Schreiber M. R., Hameury J. -M., Lasota J.-P., 2004, A&A, 427, 621
Sellwood J. A., 2012, ApJ, 751, 44
Shakura N.I., Sunyaev R. A., 1973, A&A, 24, 337
Shlosman I., Begelman M. C., Frank J., 1990, Nature, 345, 679
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Springel V., 2005, MNRAS, 364, 1105
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Stringer M. J., Bower R. G., Cole S., Frenk C. S., Theuns T., 2012, MNRAS, 423, 1596
Tremaine S. et al., 2002, ApJ, 574, 740
van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, MNRAS, 414, 2458
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
White D. M. S., Frenk S. C., 1991 ApJ, 379, 52
Wiersma R. P. C., Schaye J., Theuns T., Smith B. D, 2009a, MNRAS, 393, 99
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574

[4] http://www.scipy.org
[5] http://www.matplotlib.sourceforge.net
[6] http://www.python.org

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

phase_movie.avi (http://www.mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stv2056/-/DC1).

Please note: Oxford University Press are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: CONVERGENCE TESTS

A1 Convergence of V_ϕ in idealized Galaxies

An important component of our approach to modelling BH accretion is to estimate the circulation speed of gas at the Bondi radius using the curl of the velocity field measured from an SPH kernel centred on the BH. In this section we consider the convergence of V_ϕ in idealized galaxies as the number of particles is increased. It should be noted that measurements of the curl of the velocity field are commonly used in SPH simulations to suppress excessive viscosity in rotating discs (e.g. Balsara 1995), and that this aspect of our model is not particularly novel.

Table A1. A list of the simulations for an idealized disc galaxy. The simulations have identical initial conditions, differing only in the number of particles used to realize the galaxy.

Name	Particle mass	σ_{V_ϕ}/km s^{-1}	σ_{c_s}/km s^{-1}
IDEAL-EAGLE-RES	1.4×10^6	1.94	2.41
IDEAL-EAGLE-RES5	2.5×10^5	1.87	2.73
IDEAL-EAGLE-RES25	5.6×10^4	1.63	4.10

In order to investigate the convergence of the estimate of V_ϕ, we required well-controlled simulations in which we are able to minimize differences arising in other subgrid physics components such as the star formation implementation, cooling and the treatment of the ISM. The difficulty of obtaining meaningful convergence results in the presence of subgrid physics is discussed in depth in Schaye et al. (2015). We set up a suite of simulations of idealized

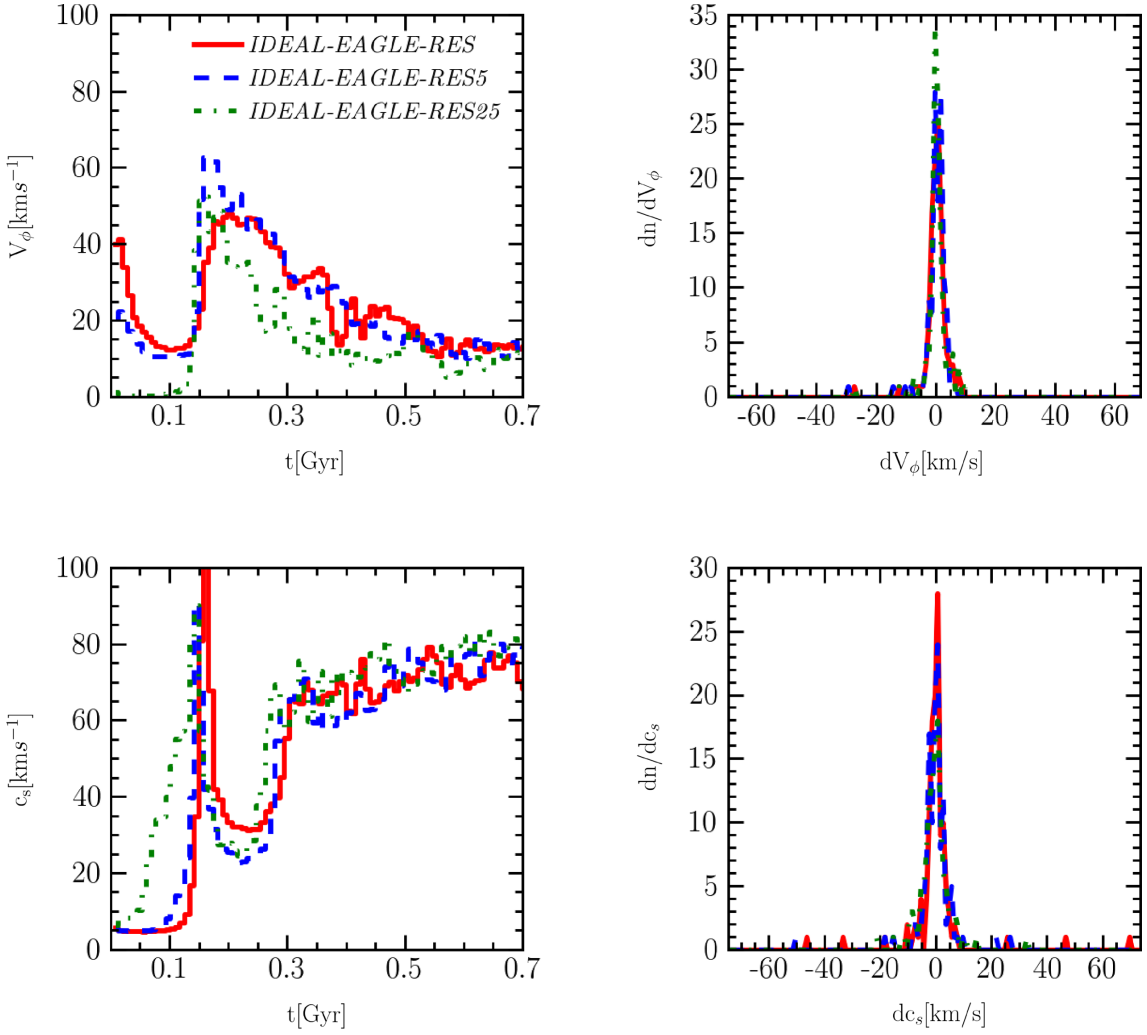

Figure A1. The evolution of V_ϕ and c_s in an idealized galaxy simulation carried out with increasing particle number. The simulations do not include cooling or feedback and all assume the same gravitational softening length. The left-hand panels shows the evolution of V_ϕ and c_s as a function of time. The galaxies quickly relax into a stable configuration in which the circulation speed measured within the BH smoothing length stabilizes to values around 10 km s^{-1}, while c_s converges to values around 80 km s^{-1}. Despite the differences in particle number, similar values are found in each of the simulations. The right-hand panels show the distribution of change V_ϕ and c_s between consecutive time steps for different resolutions. The distributions show little dependence on particle number.

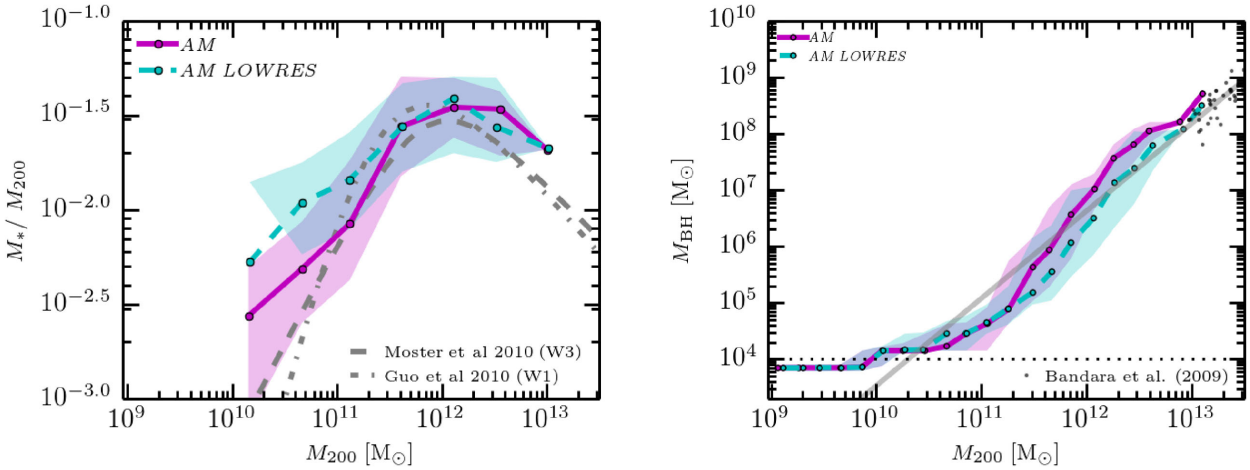

Figure A2. The right-hand panel shows the correlation of stellar mass fraction with M_{200}, while the left-hand panel shows the $M_{\rm BH}$–M_{200} relation. The simulation AM is shown as a purple solid line while the simulation AM LOWRES, with lower resolution in mass, is included as cyan dotted line. Although there is some residual offset in BH masses in the right-hand panel, the stellar mass fractions (left-hand panel) shows good convergence in haloes that are well resolved in both simulations.

disc galaxies generated from the same initial conditions as used by Springel et al. (2005), Dalla Vecchia & Schaye (2008) and Stringer et al. (2012). These papers showed that even such idealized galaxies are very sensitive to resolution for identical physical assumptions. In order to minimize these issues, we disabled cooling and feedback from SN and AGN. Our aim is to undertake the same calculation more accurately as we increase the number of particles, and we therefore use the same gravitational softening length in each run. Even under these conditions, convergence is not guaranteed since the amplification noise in the spiral density waves will not be identical (e.g. Sellwood 2012).

The idealized simulations consist of a static dark matter halo of mass $10^{12}\,h^{-1}\,{\rm M}_\odot$ with a Hernquist mass distribution (Hernquist 1990) and an exponential disc of stars and gas. The halo has a dimensionless spin parameter $\lambda = 0.33$. The disc contains 4 per cent of both the total mass and the total AM and the bulge contains 0.014 per cent of the total mass. The initial gas fraction of the disc is 0.3 and the rest consists of stars. The disc and bulge scaleheights are set to 10 per cent of the radial disc scalelengths and the vertical gas distribution is estimated to be in hydrostatic equilibrium using an iterative procedure. The initial masses of particles in the IDEAL-EAGLE-RES simulation were set to be equal to the value of the cosmological simulations. The simulation was then repeated with increasing particle number, as shown in Table A1. We follow the galaxies' evolution for 700 Myr, by which time the initial transients have decayed.

We focus on the effects of resolution on the evolution of V_ϕ and $c_{\rm s}$. The left-hand panel of Fig. A1 shows the evolution of V_ϕ (top) and $c_{\rm s}$ (bottom). During the initial transients, V_ϕ reaches high values of $\sim 60\,{\rm km\,s^{-1}}$ and $c_{\rm s}$ of $\sim 100\,{\rm km\,s^{-1}}$ more or less independent on resolution. As the disc relaxes, values of V_ϕ become much smaller, while $c_{\rm s}$ remains high. After 500 Myr, the measured values of V_ϕ and $c_{\rm s}$ are almost independent of resolution. We use these simulations to investigate the noise in the measurement of V_ϕ by examining the fluctuations in V_ϕ (and $c_{\rm s}$). In the absence of violent feedback events we expect V_ϕ to vary smoothly between time steps. The distribution of V_ϕ for each simulations is shown in the upper-right panel of Fig. A1 (the distribution of variations in $c_{\rm s}$ are shown in the lower panel). In order to avoid undue influence from outliers, we quantify the width of the distribution as half of the 16th–84th percentile range (for a Gaussian distribution is equivalent to the standard deviation). Values are given in Table A1. The distribution has a similar width regardless of the particle number.

A2 Convergence of the galaxy formation model

Because of the small relative scales of the BH and its accretion disc relative to the scale of the simulation, the most challenging aspect of the calculation is to adequately describe the properties of the multiphase gas in the region around the BH. Because of this, it would not be surprising if the parameters of the model were functions of scale. Nevertheless, we present a comparison of two versions of simulation AM carried out at different resolutions in the standard 25 Mpc volume. Specifically, we run the simulation AM LOWRES with a lower mass resolution by a factor of 8 (with 2×180^3 particles and with mass 5.88×10^7 and $1.18 \times 10^7\,h^{-1}\,{\rm M}_\odot$ for dark matter and gas particles, respectively).

Fig. A2 shows the stellar mass fraction–M_{200} relation and the $M_{\rm BH}$–M_{200} relation for these simulations in purple and cyan for well-resolved haloes in each simulation. The stellar mass fractions have similar values of the medians and scatter. From this measure, the convergence seems good. In the case of the $M_{\rm BH}$–M_{200} relation, the medians take somewhat smaller values for the low-mass BHs in the simulation AM LOWRES, but the two resolutions show a large overlap in terms of the scatter and appear to converge at high and low masses. The convergence is nevertheless better than for the accretion disc particle scheme (Power et al. 2011) when it is applied at this resolution (Muldrew et al. 2013).

APPENDIX B: THE EFFECT OF CHANGING EFFECTIVE VISCOSITY PARAMETER $C_{\rm visc}$

In this appendix, we show the effect of changing the value of the effective viscosity parameter on the stellar mass fraction in the simulated galaxy population and on the $M_{\rm BH}$–M_{200} relation. We increase $C_{\rm visc}$ by two orders of magnitude from 6.3×10^4 to 6.3×10^6 (our fiducial value is 2.1×10^6, purple line). The Fig. B1 shows the stellar mass fraction and the $M_{\rm BH}$–M_{200} relation when $C_{\rm visc}$ takes these values. We have included the simulations NO-AGN ($C_{\rm visc} = \infty$,

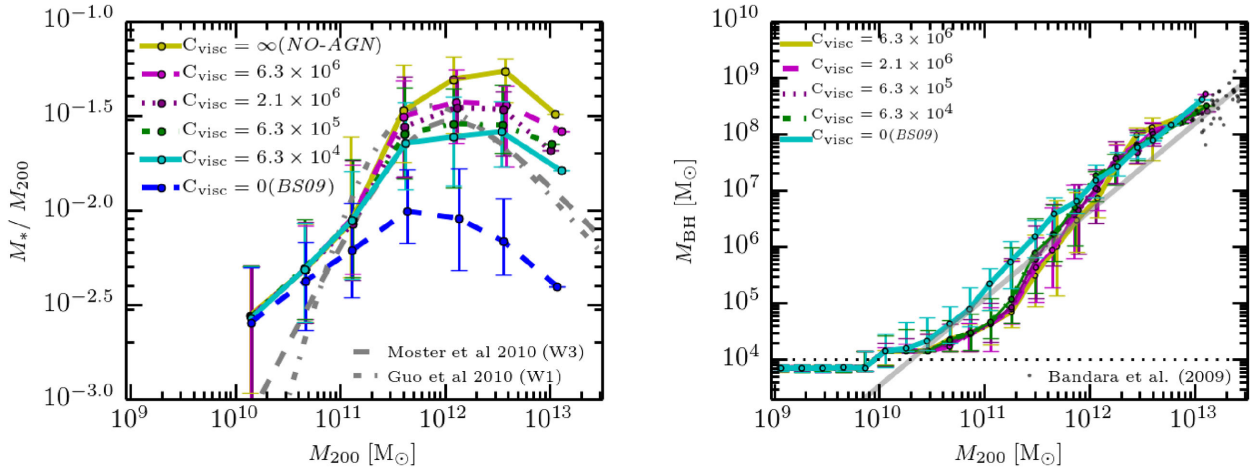

Figure B1. The left-hand panel shows the stellar mass fraction–M_{200} relation and the right-hand panel shows the M_{BH}–M_{200} relation for simulations similar to AM ($C_{visc} = 2.1 \times 10^6$), but with varying C_{visc} as is indicated in the legends. The simulations NO-AGN ($C_{visc} = \infty$, yellow solid colour) and BS09 ($C_{visc} = 0$, blue dashed colour) show extreme cases of the effects of C_{visc}. The smaller the C_{visc}, the higher the suppression in the stellar mass fraction is above a critical mass halo. The critical halo mass decreases slightly as C_{visc} decreases. The M_{BH}–M_{200} relation is not affected above this critical mass, but as C_{visc} decreases slightly less massive BHs are hosted by haloes below this critical mass.

yellow colour) and BS09 ($C_{visc} = 0$, blue colour) as extreme cases of C_{visc}.

Decreasing C_{visc} (corresponding to a higher disc viscosity or thicker disc) results in the haloes being regulated by BH accretion at slightly lower stellar masses and as a result the stellar mass fraction (left-hand panel) breaks at a critical halo mass; however, how sharp the turnover in the stellar mass fraction is depends on C_{visc}. The extreme case of this effect is shown in the simulation BS09 ($C_{visc} = 0$) where the stellar mass fraction lies below the abundance matching results. For less extreme values, the dependence on C_{visc} is weak, re-enforcing the idea that the break in the mass fraction is set by the impact of periods of AGN heating (and hence by the halo's cooling and dynamical time-scales) rather than on the details of the BH accretion model.

Looking at the M_{BH}–M_{200} relation (right-hand panel), for haloes above mass $10^{12}\,M_\odot$ variations in C_{visc} have little effect: in this regime, BHs are able to self-regulate their growth regardless of the details of the accretion model. Haloes below this critical mass tend to host less massive BHs, although the dependence on value of C_{visc} is again weak.

Hence, the main results are relatively insensitive to variations of C_{visc} regarding the critical halo mass in which the break of stellar mass fraction occurs, reproducing a turnover for haloes with mass larger than the critical mass. The sharpness of the turnover depends on the choice of C_{visc}. Large changes in C_{visc} are required to have significant impact because the suppression by AM is only effective when $(C_{visc})^{1/3} V_\phi > c_s$ (see. equation 11). Although the changes in C_{visc} affect the sharpness of the turnover in the mass fraction, this allows adjustment of the location of the break in the stellar mass fraction seen in cosmological volume simulations.

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