Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **420,** 2662–2683 (2012) doi:10.1111/j.1365-2966.2011.20236.x

Self-regulated growth of supermassive black holes by a dual jet–heating active galactic nucleus feedback mechanism: methods, tests and implications for cosmological simulations

Yohan Dubois,[1⋆] Julien Devriendt,[1,2] Adrianne Slyz[1] and Romain Teyssier[3,4]

[1]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH*
[2]*Centre de Recherche Astrophysique de Lyon, Université de Lyon I, CNRS UMR 5574, ENS-Lyon, 9 Avenue Charles André, 69561 St-Genis-Laval Cedex, France*
[3]*Institute für Theoretische Physik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[4]*CEA Saclay, DSM/IRFU/SAP, Bâtiment 709, F-91191 Gif-sur-Yvette, Cedex, France*

Accepted 2011 November 18. Received 2011 October 26; in original form 2011 July 30

ABSTRACT

We develop a subgrid model for the growth of supermassive black holes (BHs) and their associated active galactic nucleus (AGN) feedback in hydrodynamical cosmological simulations. This model transposes previous attempts to describe BH accretion and AGN feedback with the smoothed particle hydrodynamics (SPH) technique to the adaptive mesh refinement framework. It also furthers their development by implementing a new jet-like outflow treatment of the AGN feedback which we combine with the heating mode traditionally used in the SPH approach. Thus, our approach allows one to test the robustness of the conclusions derived from simulating the impact of self-regulated AGN feedback on galaxy formation vis-à-vis the numerical method. Assuming that BHs are created in the early stages of galaxy formation, they grow by mergers and accretion of gas at a Eddington-limited Bondi accretion rate. However this growth is regulated by AGN feedback which we model using two different modes: a quasar-heating mode when accretion rates on to the BHs are comparable to the Eddington rate, and a radio-jet mode at lower accretion rates which not only deposits energy, but also deposits mass and momentum on the grid. In other words, our feedback model deposits energy as a succession of thermal bursts and jet outflows depending on the properties of the gas surrounding the BHs. We assess the plausibility of such a model by comparing our results to observational measurements of the co-evolution of BHs and their host galaxy properties, and check their robustness with respect to numerical resolution. We show that AGN feedback must be a crucial physical ingredient for the formation of massive galaxies as it appears to be able to efficiently prevent the accumulation of and/or expel cold gas out of haloes/galaxies and significantly suppress star formation. Our model predicts that the relationship between BHs and their host galaxy mass evolves as a function of redshift, because of the vigorous accretion of cold material in the early Universe that drives Eddington-limited accretion on to BHs. Quasar activity is also enhanced at high redshift. However, as structures grow in mass and lose their cold material through star formation and efficient BH feedback ejection, the AGN activity in the low-redshift Universe becomes more and more dominated by the radio mode, which powers jets through the hot circumgalactic medium.

Key words: methods: numerical – galaxies: active – galaxies: evolution – galaxies: jets – quasars: general.

1 INTRODUCTION

Evidence for the ubiquitous presence of supermassive black holes (BHs) in the centres of galaxies is overwhelming (Kormendy & Richstone 1995). BHs spanning a range of masses from a few $10^6\,M_\odot$ in the centre of galaxies with small bulges like our Milky Way (Schödel et al. 2002) up to several $10^9\,M_\odot$ for elliptical galaxies in the cores of groups and clusters of galaxies (Magorrian et al. 1998) have now been observed. These supermassive BHs are seen not only in the near universe, but also in very luminous quasars

⋆E-mail: yohan.dubois@physics.ox.ac.uk

discovered beyond $z > 6$ (Fan et al. 2003) and suggest that they are already in place during the early stages of galaxy formation. As a consequence, it is now widely accepted that a large variety of galaxies host BHs in their centres, and that these BHs somehow influence the evolution of their host galaxies.

Observations by Magorrian et al. (1998) first pointed out a relationship between the central BHs and their host galaxy bulge mass with a quasi-linear scaling (Laor 2001; McLure & Dunlop 2002; Marconi & Hunt 2003; Häring & Rix 2004). A similar, albeit arguably tighter correlation is also found between the BH mass and the stellar velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Gültekin et al. 2009), or the Sérsic index that measures the concentration of the bulge (Graham & Driver 2007). These correlations define a BH Fundamental Plane, similar to the Fundamental Plane of elliptical galaxies, that links BHs to bulge stellar masses, velocity dispersions and effective radii (Hopkins et al. 2007).

These observations led to the suggestion that the growth of BHs is self-regulated by the energy released during their accretion phase. This would be sufficient to unbind the gas in the galaxy and form powerful outflows (Silk & Rees 1998; King 2003; Wyithe & Loeb 2003). There exists abundant observational evidence of such outflows with direct imaging of X-ray cavities in the vicinity of elliptical galaxies (Boehringer et al. 1993; Owen, Eilek & Kassim 2000; Bîrzan et al. 2004; McNamara et al. 2005; Fabian et al. 2006; Taylor et al. 2006; Dong, Rasmussen & Mulchaey 2010; Dunn et al. 2010) or indirect detection of broad line absorption regions in the spectra of quasars (Chartas, Brandt & Gallagher 2003; Crenshaw, Kraemer & George 2003; Pounds et al. 2003). These observations are supported by numerical models of the microphysics of BH accretion discs that can drive massive hydromagnetic outflows and jets (De Villiers et al. 2005; McKinney 2006; McKinney & Blandford 2009), and large amounts of heat carried by photons that could potentially ionize the surrounding gas. Feedback from BHs is commonly called active galactic nucleus (AGN) feedback as the energy is emitted from the centres of galaxies where BHs reside. However, this generic appellation encompasses various modes of energy release from the central source.

It is commonly believed that two of these modes can describe AGN feedback. These are very similar to the two modes observed for X-ray binaries (Churazov et al. 2005; Merloni & Heinz 2008). The so-called 'radio' mode is associated with a strong radio emission filling X-ray-depressed cavities with relativistic electrons. Most of the energy in the radio mode is driven by a mechanical jet feedback mechanism inflating the radio lobe itself. This mode is equivalent to the 'low/hard' state of X-ray binaries which has a hard X-ray power spectrum with a cut-off at a few 100 keV, and gas accretion at low Eddington rates. The same tendency has been confirmed for supermassive BHs, for which the radio loudness is stronger at lower Eddington accretion ratios (Chiaberge, Capetti & Macchetto 2005; Churazov et al. 2005; Nagar, Falcke & Wilson 2005). This mode is clearly associated with mechanical feedback where most of the energy is powered by the jet mechanism and largely overwhelms the X-ray contribution from the nucleus, as in the well-studied case of M87 (Owen et al. 2000).

A transition to a radio-quiet 'quasar' mode occurs at a few $\sim 10^{-2}$ of the Eddington accretion rate for X-ray binaries (Maccarone 2003). Above this threshold, X-ray binaries enter a 'high/soft' state emitting a soft and thermal X-ray spectrum with almost no trace of a jet mechanism, which is the equivalent of a quasar spectrum. The thermal emission is well described by the standard model of optically-thick and geometrically-thin accretion discs from Shakura & Sunyaev (1973), whereas the launching of the jet for the radio mode comes from optically-thin and geometrically-thick (and radiatively-inefficient) accretion discs modelled with the advection-dominated accretion flow (ADAF) from Narayan & Yi (1994) or with the adiabatic inflow–outflow solutions from Blandford & Begelman (1999).

On the other hand, observations of tidally disrupted galaxies often reveal powerful AGN activity (Barnes & Hernquist 1992). Mergers between galaxies are invoked to compress the interstellar medium (ISM) and provide a fresh flow of material on to the BH in that case. Theoretical models of galaxy formation that associate the growth of BHs with such events have been successful at reproducing many properties of the population of quasars as well as the BH density seen at low redshift (Cattaneo, Haehnelt & Rees 1999; Kauffmann & Haehnelt 2000; Granato et al. 2001; Volonteri, Haardt & Madau 2003). Semi-analytic models of galaxy formation require AGN feedback to suppress the cooling catastrophe in massive galaxies, match the bright end of the galaxy luminosity function, and obtain bulge-dominated galaxies (Bower et al. 2006; Cattaneo et al. 2006; Croton et al. 2006; Bower, McCarthy & Benson 2008; Somerville et al. 2008).

Important steps have recently been incorporated with hydrodynamical simulation to better quantify the negative AGN feedback effect on star formation and gas properties of galaxies. These involved simulating idealized disc galaxies and galaxy mergers (Springel, Di Matteo & Hernquist 2005; Debuhr et al. 2010; Nayakshin & Power 2010), or clusters of galaxies (Cattaneo & Teyssier 2007; Dubois et al. 2009). Self-consistent subgrid models of the AGN feedback heating mode have been introduced in Λ cold dark matter (ΛCDM) cosmological simulations using the smoothed particle hydrodynamics (SPH) technique as implemented in the GADGET code (Sijacki et al. 2007; Booth & Schaye 2009). Alternative approaches based on the injection of cosmic rays have also been explored (Sijacki et al. 2008). These cosmological simulations have successfully reproduced relationships between BH masses and galaxy properties (Sijacki et al. 2007; Di Matteo et al. 2008; Booth & Schaye 2009, 2011), and suppressed the cooling catastrophe in groups and clusters of galaxies (Khalatyan et al. 2008; Puchwein, Sijacki & Springel 2008; McCarthy et al. 2010, 2011). Recently, AGN feedback associated with BH growth has been introduced in adaptive mesh refinement (AMR) cosmological re-simulations of a galaxy cluster with the RAMSES code. These featured either a jet-kinetic mode (Dubois et al. 2010) or a thermal energy input (Teyssier et al. 2011) (see also Dubois et al. 2011), but the cosmic co-evolution of BHs and galaxies has not yet been studied using grid-based techniques.

As limitations of the standard SPH technique to capture Kelvin–Helmholtz instabilities have been pointed out (Agertz et al. 2007), and authors like Mitchell et al. (2009) have shown that it could have severe consequences on the properties of the intracluster gas where AGN-host ellipticals reside, it seems worthwhile to investigate this important issue using a different numerical technique. Therefore, this work uses an Eulerian grid-based approach to model AGN feedback from BHs by a dual jet–heating subgrid model representative of the radio and quasar mechanisms with the self-regulated growth of BHs. We emphasize the importance of these two different modes on the long-term evolution of galaxies, trying to explain why quasars are a common ingredient of the young Universe and why radio jets are more commonly observed in late and massive structures such as clusters of galaxies.

This paper is organized as follows. In Section 2, we detail the numerical technique used for following the BH growth along with its associated AGN feedback, and the standard models employed

for galaxy formation [cooling, star formation, supernova (SN) feedback, etc.]. In Section 3, we describe the set of simulations employed to test the AGN feedback model with the RAMSES code. In Section 4, we present a parameter study of the AGN feedback model and assess the convergence of the results vis-à-vis resolution. Section 5 scrutinizes what drives the domination of the quasar and the radio mode at different epochs. Finally, we comment on the results in Section 6.

2 MODELLING THE PHYSICS OF GALAXY FORMATION

2.1 BH growth and AGN feedback

Sink particles were first introduced by Bate, Bonnell & Price (1995) in a SPH code. Sinks are massive particles that capture gas particles in their surroundings. They mimic the formation of unresolved compact objects, for example, protostellar cores in the ISM, BHs in the ISM, central supermassive BHs in galaxies, etc. Due to the very Lagrangian nature of the sink particle technique, it has been extensively and exclusively used in SPH codes until Krumholz, McKee & Klein (2004) extended its use to grid codes. The version in RAMSES (Teyssier 2002) is strongly inspired by the Krumholz et al. (2004) numerical implementation, and has already been presented in Dubois et al. (2010) and Teyssier et al. (2011), but we reproduce here the details of the numerical implementation to facilitate the discussion of our results.

2.1.1 Seeding galaxies with BHs

There are at least two scenarios for the formation of seed BHs. The first one invokes Population III stars with zero metallicity. These can produce BH remnants as massive as 10^2–$10^3\,\mathrm{M}_\odot$ (Madau & Rees 2001; Heger & Woosley 2002; Schneider et al. 2002) that will eventually rapidly merge in their primordial halo to reach even larger masses. Another channel of BH formation is the direct collapse of matter from haloes with very low angular momentum generating BHs as massive as $10^5\,\mathrm{M}_\odot$ (Loeb & Rasio 1994; Bromm & Loeb 2003; Begelman, Volonteri & Rees 2006). With the kpc-scales typically used in cosmological simulations of galaxy formation, it is pointless to try to follow the formation of these first seeds since this occurs on much smaller scales, but we can take these scenarios as guidelines for a subgrid generation of seed BHs.

BHs represented by sink particles are created in regions where the Jeans criterion is violated, that is, in regions where the maximum level of refinement is reached and where the gas density is large enough to potentially produce a numerical instability, in other words, where

$$\frac{\Delta x}{4} > \lambda_\mathrm{J} = \sqrt{\frac{\pi c_\mathrm{s}^2}{G\rho}}\,. \tag{1}$$

Here Δx is the size of the smallest cell, λ_J the Jeans length, c_s the sound speed and ρ the gas density. According to Truelove et al. (1997), the numerical stability of a gravitationally bound object is ensured if it is resolved with at least four cells. With a mixed composition of matter (DM, gas, stars), Jeans stability is not trivial anymore, but we can reasonably assume that gas is the dominant source of gravitational potential inside dense collapsed objects, like galaxies, in our case.

For numerical stability, each time that the Jeans criterion is violated, we should spawn a sink particle with a mass corresponding to the depleted mass. However, in cosmological simulations, this leads to excessively large sink masses. The reason is that the gas is concentrated in galactic structures that are poorly resolved with kpc-scale resolution. As a result an entire galactic disc can be defined by only a few Jeans-violating cells leading to excessively massive sink particles. To form sufficiently small seed BHs in the centres of the galaxies, we prefer to choose their initial mass, M_seed, thereby introducing a free parameter. We set $M_\mathrm{seed} = 10^5\,\mathrm{M}_\odot$ as the default value of our model in agreement with previous cosmological simulations (e.g. Booth & Schaye 2009). Despite choosing the seed mass, BHs are still spawned only in cells belonging to the maximum level of refinement and that verify equation (1). In Section 4, the importance of the choice of the initial seed mass will be tested. One consequence of this self-controlled formation of the seed BHs is that they are not allowed to accrete gas when the Jeans criterion is violated. They can only accrete gas by a reasonable physical process such as Bondi accretion. With this prescription for initializing the mass of the seed BHs, it is conceivable that gas could be numerically violently Jeans unstable, but this issue is partially solved by the consumption of gas in the star-forming process that temporarily restores gravitational stability.

To avoid formation of sink particles in low-density regions that are Jeans unstable, we set a minimum threshold for the density $\rho > \rho_0$ of gas that can create a new sink, where ρ_0 is the same density threshold that we use for star formation. To make sure that sink BHs do not form before the very first stars form, we check that the local star density ρ_* calculated with a Cloud-in-Cell (CIC) interpolation verifies

$$f_* = \frac{\rho_*}{\rho_* + \rho} > 0.25\,, \tag{2}$$

before a new sink particle is spawned, where ρ is the local gas density. Note that these criteria are very similar to those employed by Bellovary et al. (2010), as they also confine the formation of seed BHs to cold, dense, metal-poor gaseous regions at the centre of galaxies.

To obtain one BH per massive galaxy only, a halo finder is usually run on-the-fly during the simulation to check if candidate galaxies already host a BH (Di Matteo, Springel & Hernquist 2005; Booth & Schaye 2009). We prefer a simpler, more direct, and computationally cheaper approach. To avoid creating multiple BHs inside the same galaxy, we ensure that each time a cell could potentially produce a sink particle (i.e. it verifies equation 1), it is farther than a minimum radius r_min from all other pre-existing sink particles. This distance has to be larger than the typical size of galactic discs and smaller than the typical average intergalactic distance. Test runs suggest that the choice $r_\mathrm{min} = 50\,\mathrm{kpc}$ produces very satisfactory results.

In summary, a sink particle forms out of gas satisfying criteria on: Jeans instability, gas density threshold, stellar fraction threshold, and minimum distance from other BHs. Once the sink particle is created, it is split into several cloud particles with equal mass. Cloud particles are spread over a $4\Delta x$ radius sphere and positioned every $0.5\Delta x$ in (x, y, z). The exact number of cloud particles in this configuration is therefore $n_\mathrm{cloud} = 2109$ per sink. These cloud particles are essentially created to probe the evolution of the region around the BH and provide spatially averaged quantities for the Bondi accretion formula. They move around on the finest time-step scale (corresponding to the highest spatial resolution) and are destroyed and re-created around their sink particles with a given distribution at the beginning of every coarse time-step (corresponding to lowest spatial resolution).

On the other hand, sink particles are only updated every coarse time-step with quantities that have been evolved through the intermediate calculation with cloud particles. They are merged together (mimicking the BH merger) if they stand at a distance closer than $4\Delta x$ from each other. Mass is conserved in this process and momentum vectors of the old sink particles are simply added to compute the momentum of the new sink particle. They are also the source AGN feedback.

Finally, we insist on the fact that BH positions and velocities are updated in the classical way used to update standard particles such as DM particles, that is, using the particle-mesh (PM) solver of RAMSES with CIC interpolation of particle masses into cells. No correction on their positions and velocities is done to force them to stay near their host galaxy (as could be done with a halo-finder approach). Thus, weakly bound objects, such as BHs in galaxy satellites of large groups and clusters, may be stripped from their host galaxy. These BHs behave like star particles that tidal forces compel to populate the stellar halo of massive galaxies.

2.1.2 Accretion rate

Since we do not resolve the accretion discs around BHs, whose sizes are subparsec even for the most massive ones ($\sim 10^{-3}$ pc according to Morgan et al. 2010 from microlensing estimates), we use the most common prescription that these BHs accrete gas at a Bondi–Hoyle–Lyttleton rate (Bondi 1952)

$$\dot{M}_{\rm BH} = \frac{4\pi\alpha G^2 M_{\rm BH}^2 \bar{\rho}}{\left(\bar{c}_{\rm s}^2 + \bar{u}^2\right)^{3/2}} , \tag{3}$$

where α is a dimensionless boost factor ($\alpha \geq 1$), $M_{\rm BH}$ is the BH mass, $\bar{\rho}$ is the average gas density, $\bar{c}_{\rm s}$ is the average sound speed, and $\bar{u}$ is the average gas velocity relative to the BH velocity. One of the major difficulties encountered with the computation of the relative gas velocity is that in cosmological runs, the ISM is poorly resolved and leads to galaxy scaleheights comparable to the resolution which is much larger than the scaleheights of galaxies in nature. Moreover, due to limited sampling of the gravitational force in the galactic disc, BHs can wander in their host galaxy. For this reason, a BH close to the centre of a galaxy can feel the infalling material coming from the halo or the ISM at a relative velocity much higher than the typical velocity inside the bulge. Therefore, because $\bar{u}$ is not a reliably measured quantity, we enforce the relative velocity to be no larger than an average gas velocity dispersion in the ISM which is assumed constant and equal to $u_{\rm max} = 10\,{\rm km\,s^{-1}}$ for our fiducial model (Dib, Bell & Burkert 2006). We will test the effect of varying this maximum allowed value on the properties of BHs.

The average density $\bar{\rho}$ and sound speed $\bar{c}_{\rm s}$ are computed around the BH using the cloud particles for this operation, as mentioned in the previous section. To compute the averages, the cell in which each cloud particle sits is assigned a weight given by a kernel function w, similar to the one used in Krumholz et al. (2004):

$$w \propto \exp\left(-r^2/r_{\rm K}^2\right) , \tag{4}$$

where r is the distance from the cloud particle to the sink particle and $r_{\rm K}$ is the radius defined as

$$r_{\rm K} = \begin{cases} \Delta x/4 & r_{\rm BH} < \Delta x/4 , \\ r_{\rm BH} & \Delta x/4 \leq r_{\rm BH} \leq 2\Delta x , \\ 2\Delta x & r_{\rm BH} > 2\Delta x . \end{cases} \tag{5}$$

The Bondi–Hoyle radius $r_{\rm BH}$ is given by

$$r_{\rm BH} = \frac{G M_{\rm BH}}{c_{\rm s}^2} , \tag{6}$$

where $c_{\rm s}$ is the exact sound speed in the cell where the sink lies.

The accretion rate on to the sink is finally limited by its Eddington rate

$$\dot{M}_{\rm Edd} = \frac{4\pi G M_{\rm BH} m_{\rm p}}{\epsilon_{\rm r}\sigma_{\rm T} c} , \tag{7}$$

where $\sigma_{\rm T}$ is the Thompson cross-section, c is the speed of light, $m_{\rm p}$ is the proton mass, and $\epsilon_{\rm r}$ is the radiative efficiency, assumed to be equal to 0.1 for the Shakura & Sunyaev (1973) accretion on to a Schwarzschild BH.

The accretion rate is computed at each time-step and a fraction $\dot{M}_{\rm BH}\Delta t/n_{\rm cloud}$ of gas mass is depleted from the cell where the cloud particle lies and is added to that cloud particle and to the sink particle. For each coarse time-step of the simulation, cloud particles are re-scattered with equal mass $M_{\rm BH}/n_{\rm cloud}$. As the time-step does not depend on the accretion speed on to BHs and as low-density cells can be close to high-density cells, a BH might remove more mass than is acceptable. To avoid dealing with negative or extremely low gas densities and numerical instabilities arising from this, we do not allow any cloud particle to deplete more than 25 per cent of the gas content in a cell.

With large-scale cosmological simulations and the limited typical kpc-scale resolution, we cannot resolve the scale and the clumpiness of the ISM that require parsec-like resolution (Powell, Slyz & Devriendt 2011). To prevent the collapse of the gas from numerical instabilities and to take into account the mixing of the different phases in the ISM (cold and warm components), we use the polytropic equation of state (EoS) described in Section 2.2. Applying this EoS means that it is impossible to track the 'true' density and sound speed in the ISM; thus, the accretion rate on to BHs must be modified. Early works modelling the accretion rate on to BHs with such a polytropic EoS set the boost factor to a constant value $\alpha = 100$ (Springel et al. 2005; Sijacki et al. 2007; Di Matteo et al. 2008). Here we follow the prescription from Booth & Schaye (2009) who show that $\alpha = (\rho/\rho_0)^2$ for $\rho > \rho_0$ where $\rho_0 = 0.1\,{\rm H\,cm^{-3}}$ is the threshold for star formation, and $\alpha = 1$ for $\rho \leq \rho_0$ gives reasonable results compared to observational predictions.

We insist on the fact that this polytropic EoS (equation 16) has important consequences on the accretion rate on to BHs in high-gas-density regions: equation (3) turns into $\dot{M}_{\rm BH} \propto M_{\rm BH}^2\rho^{5/2}$, and the temperature dependence vanishes. On the other hand, as soon as the cold gas component has been evaporated by star formation or feedback mechanisms giving $\rho \leq \rho_0$ in massive galaxies, the accretion rate of the BH is, by definition, the proper ($\alpha = 1$) Bondi accretion rate. This α boost of the accretion rate is an artificial way of modelling the very fast accretion of gas within cold and gas-rich galaxies at early epochs, where the clumpiness of the ISM due to gas–disc fragmentation is unresolved in large-scale cosmological simulations.

2.1.3 AGN feedback: quasar and radio modes

It is believed that the feedback from AGNs can proceed in two distinct modes. The quasar mode is essentially seen in the high-redshift Universe and proceeds by emitting large amounts of radiation that can photoionize and heat gas. It is assumed in reionization models of the IGM that quasars are an important contribution to the ultraviolet (UV) background (Haardt & Madau 1996). The radio mode

of AGN feedback, on the other hand, proceeds at lower redshifts in the cores of massive galaxy haloes. The typical signatures of this radio mode are inflated cavities with strong magnetic fields and high levels of cosmic-ray energy.

Our aim is to treat self-consistently both modes in the simulation according to very simple prescriptions. It is believed that the radio mode is preferentially triggered during low-accretion-rate episodes, and that the quasar mode occurs when gas accretion takes place at rates comparable to the Eddington limit (Churazov et al. 2005; Merloni & Heinz 2008). We use the ratio of accretion rate to its Eddington limit

$$\chi = \frac{\dot{M}_{\rm BH}}{\dot{M}_{\rm Edd}} \tag{8}$$

as the criterion to determine which of the two AGN modes is active. Following Merloni & Heinz (2008), we take $\chi_{\rm radio} = 10^{-2}$ as the value dividing the radio from the quasar mode. Above $\chi > \chi_{\rm radio}$, the AGN undergoes quasar-like activity with energy mostly emitted by photons. We model this mode by thermal injection of energy. Below $\chi \leq \chi_{\rm radio}$, BHs smoothly accrete gas and provide a radio-mode feedback which is modelled by our kinetic jet implementation. We point out that a similar approach has been taken by Sijacki et al. (2007), but they treat both modes as thermal inputs of energy with different injection radii.

For both modes, we assume that a fraction $\epsilon_{\rm f}$ of the radiated energy, $L_{\rm r}$, is released to the ambient gas:

$$\dot{E}_{\rm AGN} = \epsilon_{\rm f} L_{\rm r} = \epsilon_{\rm f}\epsilon_{\rm r}\dot{M}_{\rm BH}c^2 , \tag{9}$$

where $\epsilon_{\rm f}$ is a free parameter that depends on the mode that is triggered by the accretion. As the energy is continuously released, time-scales for dissipating energy by cooling can be sometimes far smaller than the hydro time-step. This problem is often encountered in SN feedback modelling (Navarro & White 1993) and leads to a null dynamical impact on the surrounding gas. So as to impact the ambient medium, some authors release the AGN heating energy only when a sufficient amount of gas has been accreted (Sijacki et al. 2007; Booth & Schaye 2009). Our modelling of the quasar mode ($\chi > \chi_{\rm radio}$) as a heating mode is very similar to the approach adopted by Booth & Schaye (2009) (see also Teyssier et al. 2011): we store the rest-mass energy of gas accreted on to the BH until it would be enough to raise the temperature of the gas around the BH above $T = 10^7$ K. At that point, we release the energy as thermal energy within a bubble of radius $r_{\rm AGN}$ around the BH with efficiency $\epsilon_{\rm f} = 0.15$.

For the radio mode ($\chi \leq \chi_{\rm radio}$), we model the AGN feedback with a jet-like outflow with the same profile as in Omma et al. (2004) (see also Dubois et al. 2010). Mass, momentum and energy are spread over a small cylinder of radius $r_{\rm J}$ and height $2h_{\rm J}$ multiplied by a kernel window function

$$\psi(r_{\rm cyl}) = \frac{1}{2\pi r_{\rm J}^2}\exp\left(-\frac{r_{\rm cyl}^2}{2r_{\rm J}^2}\right) , \tag{10}$$

where $r_{\rm cyl}$ is the cylindrical radius, and where we impose $r_{\rm J} = h_{\rm J} = r_{\rm AGN}$. The size of the jet, $r_{\rm AGN}$, for the radio mode and the size of the bubble, $r_{\rm AGN}$, for the quasar mode are parameter choices which we will test in Section 4.1.6. Note that once the radius is chosen, it remains fixed for the duration of the simulation. By contrast, Sijacki et al. (2007) use prescriptions where the size of the radio bubbles depends on the amount of energy released and the gas density (see also Barai 2008). This different choice stems from the fact that these authors model the formation of bubbles by assuming they are the result of large radio cocoons inflated by jets, whereas we attempt to directly model the jet.

The mass deposition for the radio mode follows

$$\dot{M}_{\rm J}(r_{\rm cyl}) = \frac{\psi(r_{\rm cyl})}{\|\psi\|}\eta\dot{M}_{\rm BH} , \tag{11}$$

where $\|\psi\|$ is the integrated value of ψ over the whole cylinder, and $\eta = 100$ is an arbitrary value that represents the mass loading factor of the jet on unresolved scales. The value of η adopted here corresponds to a subrelativistic bipolar outflow with velocity $10\,000\,{\rm km\,s^{-1}}$, rather than that which is expected from a radio-loud relativistic jet launched from the BH horizon. Thus, our modelling should be interpreted as a tentative description of the wind arising from a larger region surrounding the BH, and is expected to carry most of the momentum (see Omma et al. 2004 for a more thorough discussion of this issue). Such a choice also allows one to keep the Courant time-step of the simulation under control whist retaining a physically motivated model of the the jet outflow propagation on kpc scales.

Mass is transferred from the central cell (where the BH lies) to all the cells enclosed within the jet. Momentum, q, is deposited in outflowing opposite directions from the centre along the jet axis, according to

$$\|\dot{\boldsymbol{q}}_{\rm J}\|(r_{\rm cyl}) = \dot{M}_{\rm J}(r_{\rm cyl})\|\boldsymbol{u}_{\rm J}\| = \frac{\psi(r_{\rm cyl})}{\|\psi\|}\dot{M}_{\rm BH}\sqrt{2\epsilon_{\rm f}\epsilon_{\rm r}}\eta c\frac{\boldsymbol{j}\cdot{\rm d}\boldsymbol{r}}{\|{\rm d}\boldsymbol{r}\|} , \tag{12}$$

where $\|\boldsymbol{u}_{\rm J}\| = (2\epsilon_{\rm f}\epsilon_{\rm r}/\eta)^{1/2}c$ is the velocity of the jet, ($\|\boldsymbol{u}_{\rm J}\| \simeq 9487\,{\rm km\,s^{-1}}$ for $\epsilon_{\rm f} = 1$ and $\eta = 100$), $\boldsymbol{j}$ is the spin vector of the BH and defines the jet axis, and d$\boldsymbol{r}$ is the distance vector from the centre of the BH. $\boldsymbol{j}$ is computed by adding the different contributions from the neighbouring cells, sampled with the cloud particles, to the total angular momentum

$$\boldsymbol{J} = \sum_{i=1}^{n_{\rm clouds}} m_i\,{\rm d}\boldsymbol{r}_i \times \boldsymbol{u}_i , \tag{13}$$

where m_i and $\boldsymbol{u}_i$ are the mass and velocity, respectively, of the gas in the cell harbouring the cloud particle, so that $\boldsymbol{j} = \boldsymbol{J}/\|\boldsymbol{J}\|$. Finally, the kinetic energy released within a single cell is

$$\dot{E}_{\rm J}(r_{\rm cyl}) = \frac{\dot{\boldsymbol{q}}_{\rm J}^2(r_{\rm cyl})}{2\dot{M}_{\rm J}(r_{\rm cyl})} = \frac{\psi(r_{\rm cyl})}{\|\psi\|}\dot{E}_{\rm AGN} . \tag{14}$$

Integrating the energy over all the cells within the jet, we verify that the energy is strictly equal to $\dot{E}_{\rm AGN}$ given in equation (9). Energy efficiency $\epsilon_{\rm f}$ is a free parameter which we take to have different values depending on the AGN feedback mode, with $\epsilon_{\rm f,r} = 1$ and $\epsilon_{\rm f,q} = 0.15$ our fiducial values for the radio and quasar mode, respectively.

High values of $\epsilon_{\rm f,r} \sim 1$ for the radio mode of AGN feedback are consistent with relativistic magnetohydrodynamic simulations of BH accretion discs (De Villiers et al. 2005; Hawley & Krolik 2006) for maximally spinning BHs. Using analytic arguments, Benson & Babul (2009) argue that balancing the spin-up of BHs caused by gas accretion by angular momentum extraction through a jet naturally leads to an equilibrium high spin value of the BHs, that is, $a/M = 0.93$ corresponding to typical efficiencies $\epsilon_{\rm f,r} \sim 0.1$ and which seem to agree quite well with local observations (Allen et al. 2006). Of course such a picture is simplistic because of potentially rapid change of BH spin during mergers, and the accretion mode of BHs (thin discs or ADAFs) is uncertain. However, it has the merit of yielding a very straightforward prediction of what should be the spin (and, thus the efficiency) of an isolated BH with an ADAF-like

accretion disc, model which is consistent with our wind feedback assumptions.

Unlike in Dubois et al. (2010) and Dubois et al. (2011), in this implementation of kinetic AGN feedback (the radio mode), we allow for a time-delay between the energy release and its mass accretion similar to what Sijacki et al. (2007) do for thermal bubbles in their radio mode. The idea is that the energy is released into a kinetic jet when the BH has grown by more than $\Delta M_{\rm d}$ per cent of its mass. This parameter, $\Delta M_{\rm d}$, gives a relative, but artificial, control on the time-scales of AGN feedback and their duty cycle, by allowing long periods during which the AGN is off and short periods during which the energy is released.

Our approach for this dual radio–quasar mode of AGN feedback is obviously and voluntarily more simplistic than reality. As said before, the quasar mode involves the release of soft X-ray photons for which radiative transfer effects are not negligible. Also even during the quasar mode it is possible to get a faint radio detection, though it is not completely clear whether such a signal would be coming from the remnant of a previously active radio mode or whether it is an intrinsic signal of the quasar mode. Jets are filled with a non-thermal cosmic-ray component and strong magnetic fields that could have important consequences on the dynamics of the jet. Moreover, the transition from the radio mode to the quasar mode very likely takes place at a different accretion ratio χ from the transition from the quasar to the radio mode, reflecting the changing nature of the accretion disc.

2.2 Modelling star formation and stellar feedback

Gas in our simulation radiates energy by atomic collisions in a H/He primordial composition gas (Sutherland & Dopita 1993) down to $T_0 = 10^4$ K, so that it can collapse into DM potential wells to form galaxies. We also account for the enhancement of cooling by metals released in SN explosions from massive stars. The metals are passively advected with the gas and a solar composition of heavy elements is assumed. The minimum temperature T_0 reached is not modified by the presence of metals but they allow for a more efficient cooling. Heating from a UV background is considered following Haardt & Madau (1996) during and after the redshift of reionization which we take to be $z_{\rm reion} = 10.5$.

Star formation occurs in high-density regions with gas density $\rho > \rho_0 = 0.1\,{\rm H\,cm^{-3}}$ using a random Poisson process to spawn star cluster particles, according to a Schmidt–Kennicutt law

$$\dot{\rho}_* = \epsilon_* \frac{\rho}{t_{\rm ff}}\,, \tag{15}$$

where $\dot{\rho}_*$ is the star formation rate (SFR) density, ϵ_* is the star formation efficiency, and $t_{\rm ff}$ is the gas fee-fall time. In these simulations, we set the efficiency of star formation to $\epsilon_* = 0.02$ in good agreement with observational surface density relationships of galaxies (Kennicutt 1998), and local giant molecular clouds (Krumholz & Tan 2007). Each star cluster particle has a mass of $m_* = \rho_0 \Delta x^3$, reaching $3.6 \times 10^5\,h^{-1}\,{\rm M}_\odot$ for our most resolved simulation with $\Delta x = 0.38\,h^{-1}$ kpc. For numerical stability, we check that no more than 90 per cent of the gas in a cell is depleted during the star formation process for numerical stability.

We account for the mass and energy release from Type II SNe assuming a Salpeter initial mass function (IMF). Using this IMF, 10 per cent of the stars more massive than $10\,{\rm M}_\odot$ end their life as Type II SNe releasing 10^{51} erg per $10\,{\rm M}_\odot$. Direct thermal input of energy from SNe has been identified as an inefficient way of returning energy back into the ISM because thermal energy is efficiently radiated away by gas cooling in high-density regions (Navarro & White 1993). Approaches to circumvent this include temporarily switching off gas cooling to allow the blast wave to propagate or kinetic energy feedback. The method from Dubois & Teyssier (2008) implements the SN energy input by releasing mass, momentum and kinetic energy locally into the surrounding gas according to a Sedov blast wave solution. The explosion takes place 10 Myr after the birth of a star cluster particle and a fraction of the gas in the cell where the star particle resides is carried into the neighbouring cells with a mass loading factor $f_w = 1$. We assume that Type II SNe release all their mass into the gas (no stellar remnant) with a $y = 0.1$ stellar yield, which is the fraction of primordial gas transformed into heavy elements and released back into the ISM. Our prescription does not take into account the energy and mass release from stellar winds (AGB stars), nor from long-lived Type Ia SNe.

In order to take into account the thermal impact of the heating of the ISM by SNe, we modify the temperature at high density $\rho > \rho_0$ with a polytropic EoS

$$T = T_0 \left(\frac{\rho}{\rho_0}\right)^{p-1}\,, \tag{16}$$

where p is the polytropic index of the gas. The adopted value of $p = 4/3$ is comparable to the value obtained in Springel & Hernquist (2003) through analytic considerations of the multiphase structure of the ISM with stellar heating. This value of $p = 4/3$ does not rigorously ensure that gas will not fragment because of numerical instabilities (Truelove et al. 1997), as with this polytropic index the Jeans length is proportional to the gas density

$$\lambda_{\rm J} = 10.7 \left(\frac{\rho}{0.1\,{\rm H\,cm^{-3}}}\right)^{-1/3} {\rm kpc}. \tag{17}$$

The above formula shows that at very high gas densities the Jeans length can be smaller than our minimum resolution and would cause spurious fragmentation of the gas. Fortunately, the gas cannot infinitely condense because of force resolution sampling and the star formation process that removes gas in the ISM. We did not choose the steeper, but safer, polytropic EoS, $p = 2$, because it leads to very thick galactic discs in simulations with kpc resolution.

3 NUMERICAL ASPECTS

We assume a flat ΛCDM cosmology with total matter (baryons+DM) density $\Omega_{\rm m} = 0.26$, baryon density $\Omega_{\rm b} = 0.045$, dark energy density $\Omega_\Lambda = 0.74$, fluctuation amplitude at $8\,h^{-1}$ Mpc $\sigma_8 = 0.80$ and Hubble constant $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$ consistent with *WMAP* 5-year data (Dunkley et al. 2009). We use several simulations with different box sizes $L_{\rm box}$, numbers of initial DM particles $N_{\rm DM}$, and minimum cell sizes Δx in order to test the convergence of our AGN feedback model with resolution. For a given $L_{\rm box}$ size, we generate our most resolved initial conditions (ICs) and degrade them to obtain lower resolution ICs, so that for the same box size, the ICs will produce the same structures but with different numbers of DM particles.

These simulations are run with the AMR code RAMSES (Teyssier 2002). The evolution of the gas is followed with a second-order unsplit Godunov scheme for the Euler equations. The Riemann solver for flux computation at the cell interface uses a first-order MinMod Total Variation Diminishing scheme to reconstruct the interpolated variables from their cell-centred values. Collisionless particles (DM, stellar and sink particles) are evolved using a PM solver with a CIC interpolation.

Simulations refine the initial mesh by as many as seven levels of refinement, reaching a physical cell size of $\Delta x = 0.38\,h^{-1}$ kpc for

our most resolved ICs and smallest box size $L_{\rm box} = 12.5\,h^{-1}$ Mpc (simulations 256L12noAGN and 256L12JH). Note that the $\ell_{\rm max} = 14$ level of refinement is only reached at $a_{\rm exp} = (1 + z)^{-1} = 0.8$ for our most resolved simulations. Because we enforce a nearly-constant physical resolution (rather than a constant comoving resolution), the highest refinement level triggered for a given redshift increases as the expansion factor grows with time, that is, $\ell_{\rm max} - 2$ at $a_{\rm exp} = 0.2$, $\ell_{\rm max} - 1$ at $a_{\rm exp} = 0.4$, etc. A cell is refined following a quasi-Lagrangian criterion: if more than eight DM particles lie in a cell, or if the baryon mass exceeds eight times the initial DM mass resolution. Lower thresholds for triggering refinement can be adopted to resolve the smallest haloes with 10–1000 DM particles, using sufficient force resolution (O'Shea et al. 2005). However, these can lead to excessive amplification of noise discreetness effects (Romeo et al. 2008).

Our whole set of simulations using different box sizes, resolutions and subgrid physics is summarized in Table 1.

Fig. 1 shows the gas density, temperature and metallicity through a three-colour composite image at two different redshifts for our most resolved simulation of a $50\,h^{-1}$ Mpc box (simulation 256L50JH), with cooling, star formation, SN feedback and our fiducial model of AGN feedback, that is, the dual quasar and radio modes. One can see the formation of hot and metal-rich bubbles at the intersection of filaments where galaxies collapse, already at high redshift $z = 3$, suggesting a pre-heating of haloes before a cluster forms. The bubbles extend farther into the IGM as the simulation proceeds.

4 CONSTRAINING THE AGN FEEDBACK MODEL

In order to test the model for BH growth and its associated AGN feedback, we measure different statistically meaningful quantities and compare them to their observational equivalents. Quantities such as the cosmological density of BHs and the Magorrian et al. (1998) relationships that link BH masses to their host galaxy properties provide good constraints on the co-evolution of BHs and galaxies. We also test the effect of varying the AGN feedback parameters on the cosmic SFR (SFR per unit comoving volume).

The procedure adopted here is in many ways very similar to the one used in Booth & Schaye (2009) to constrain their AGN subgrid model based on thermal energy inputs. As we share some common

Table 1. Simulations performed with different subgrid galactic models, different parameters for the AGN feedback mode and different resolutions. (a) Name of the simulation. (b) Number of DM particles. (c) Mass resolution of a DM particle. (d) Size of the simulation box. (e) Minimum resolution reached at $z = 0$. (f) Presence of feedback from SNe. (g) Presence of AGN feedback: 'BH' stands for the formation and growth of BHs without AGN feedback, 'Jet' stands for the radio mode only, 'Heat' stands for the quasar mode only, and 'Jet/heat' stands for the quasar/radio mode both triggered in the same simulation (see Section 2.1.3 for details). (h) AGN feedback efficiency. (i) AGN energy delay. (j) Maximum relative velocity of the gas to the BH. (k) Mass loading factor of the jet. (l) Initial BH mass. (m) Size of the region for AGN energy input.

(a) Name	(b) $N_{\rm DM}$	(c) $M_{\rm DM}$ ($\rm M_\odot\,h^{-1}$)	(d) $L_{\rm box}$ ($\rm Mpc\,h^{-1}$)	(e) Δx ($\rm kpc\,h^{-1}$)	(f) SN	(g) AGN	(h) $\epsilon_{\rm f}$	(i) $\Delta M_{\rm d}$ (per cent)	(j) $u_{\rm max}$ ($\rm km\,s^{-1}$)	(k) η	(l) $M_{\rm seed}$ ($\rm M_\odot$)	(m) $r_{\rm AGN}$
256L12noAGN	256^3	6.9×10^6	12.5	0.38	Yes	No	–	–	–	–	–	–
256L12JH	256^3	6.9×10^6	12.5	0.38	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx
64L25JH	64^3	3.5×10^9	25	3.04	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx
128L25noAGN	128^3	4.4×10^8	25	1.52	Yes	No	–	–	–	–	–	–
128L25BH	128^3	4.4×10^8	25	1.52	Yes	BH	–	–	10	–	10^5	–
128L25J	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	100	10^5	Δx
128L25Je0.15	128^3	4.4×10^8	25	1.52	Yes	Jet	0.15	0	10	100	10^5	Δx
128L25Je0.01	128^3	4.4×10^8	25	1.52	Yes	Jet	0.01	0	10	100	10^5	Δx
128L25Jm1	128^3	4.4×10^8	25	1.52	Yes	Jet	1	1	10	100	10^5	Δx
128L25Jm10	128^3	4.4×10^8	25	1.52	Yes	Jet	1	10	10	100	10^5	Δx
128L25Jv100	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	100	100	10^5	Δx
128L25Jv1000	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	1000	100	10^5	Δx
128L25Jη10	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	10	10^5	Δx
128L25Jη1000	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	1000	10^5	Δx
128L25Js0.1	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	100	10^4	Δx
128L25Js10	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	100	10^6	Δx
128L25J2dx	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	100	10^5	$2\Delta x$
128L25J4dx	128^3	4.4×10^8	25	1.52	Yes	Jet	1	0	10	100	10^5	$4\Delta x$
128L25H	128^3	4.4×10^8	25	1.52	Yes	Heat	0.15	–	10	–	10^5	Δx
128L25H2dx	128^3	4.4×10^8	25	1.52	Yes	Heat	0.15	–	10	–	10^5	$2\Delta x$
128L25H4dx	128^3	4.4×10^8	25	1.52	Yes	Heat	0.15	–	10	–	10^5	$4\Delta x$
128L25JH	128^3	4.4×10^8	25	1.52	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx
256L25noSNAGN	256^3	5.5×10^7	25	0.76	No	No	–	–	–	–	–	–
256L25noAGN	256^3	5.5×10^7	25	0.76	Yes	No	–	–	–	–	–	–
256L25JH	256^3	5.5×10^7	25	0.76	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx
128L50noAGN	128^3	3.5×10^9	50	3.04	Yes	No	–	–	–	–	–	–
128L50JH	128^3	3.5×10^9	50	3.04	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx
256L50noAGN	256^3	4.4×10^8	50	1.52	Yes	No	–	–	–	–	–	–
256L50JH	256^3	4.4×10^8	50	1.52	Yes	Jet/heat	1/0.15	0/–	10	100/–	10^5	Δx

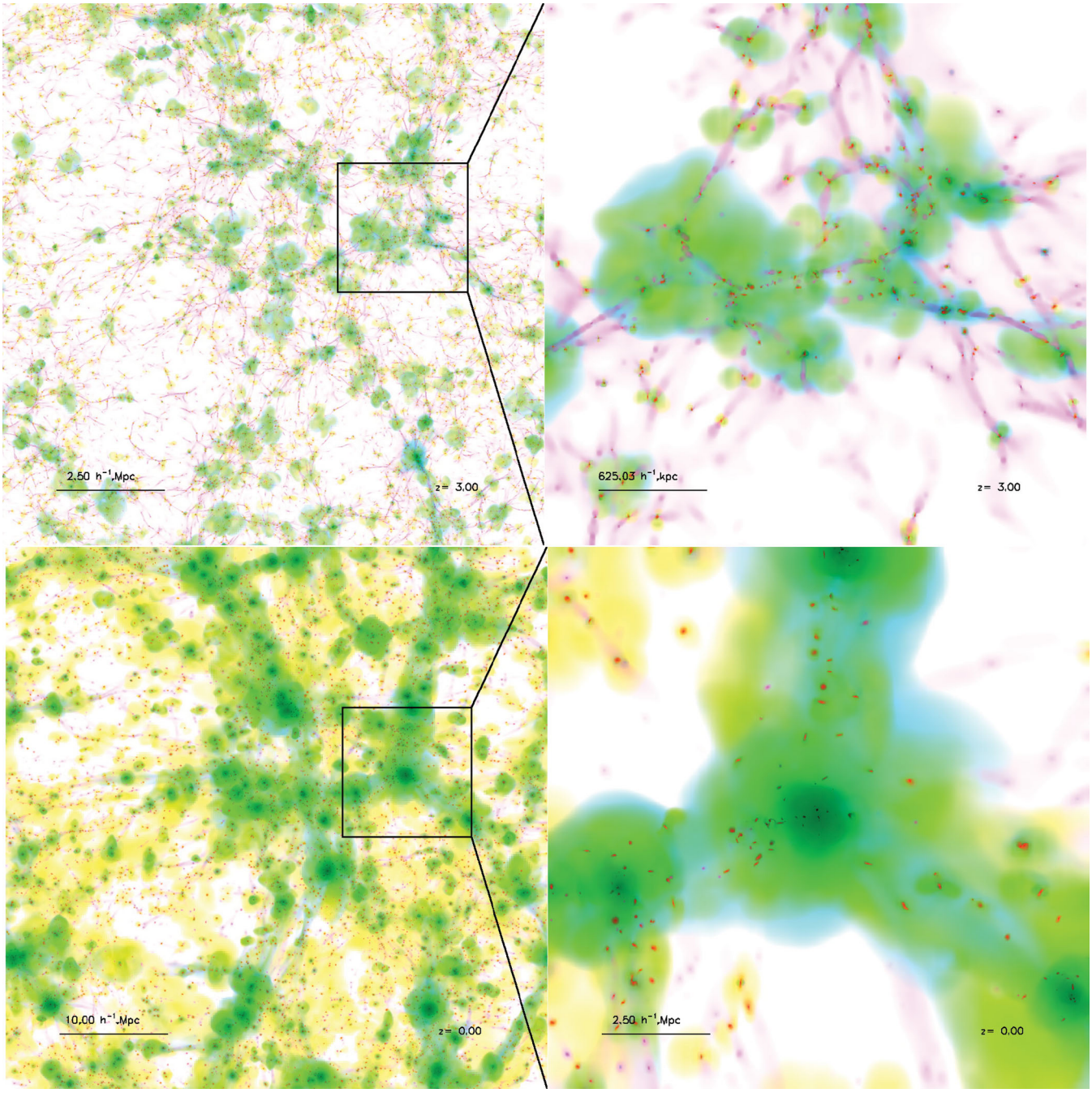

Figure 1. Three colour image of simulation 256L50JH (see Table 1) at $z = 3$ (upper panels) and $z = 0$ (bottom panels) with a zoom on the largest halo (top and bottom right-hand panels). Gas density is colour coded in magenta, gas temperature in cyan and gas metallicity in yellow.

parameters with them, like the accretion rate boost factor α, and the thermal input of energy for the quasar mode of AGN feedback, most of the parameter tests for the quasar mode will not be repeated, except for the AGN bubble size $r_{\rm AGN}$ because SPH and AMR codes treat their gas elements with a different numerical approach. Thus, we set up our thermal (quasar) mode for AGN feedback to the values of Booth & Schaye's (2009) best-fitting model: we assume an $\epsilon_{\rm f} = 0.15$ efficiency, which is also the value employed in Teyssier et al. (2011). In this paper rather than focusing on the quasar mode, we concentrate on testing our radio mode based on bipolar kinetic outflows, which is a modified version of the bipolar kinetic outflows in Dubois et al. (2011) and has never been modelled before in cosmological simulations of galaxy formation.

To test the AGN subgrid model, we fix the 'standard' galactic subgrid models for star formation and SN feedback (see Section 2.2, e.g. star formation threshold $\rho_0 = 0.1\,{\rm H\,cm^{-3}}$, star formation efficiency $\epsilon_* = 0.02$, Salpeter IMF, mass loading factor $f_w = 1$, stellar yield $y = 0.1$, polytropic index $p = 4/3$) and we vary the parameters of the AGN feedback modelling ($\epsilon_{\rm f}$, $\delta M_{\rm d}$, $u_{\rm max}$, η, $M_{\rm seed}$, $r_{\rm AGN}$, mode of AGN feedback). These tests are performed on a $L_{\rm box} = 25\,h^{-1}$ Mpc simulation box with 128^3 DM particles (i.e. simulations with names starting with the prefix '128L25' in Table 1) that are

sufficient to resolve haloes with masses as small as $\sim 10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$ ($\sim$100 particles) and as large as several $\sim 10^{13}\,h^{-1}\,\mathrm{M}_{\odot}$. This choice of box size and resolution is a good compromise between affordable computational resources and resolution requirements. Finally, convergence tests are performed with larger box sizes (simulations 128L50JH and 256L50JH) and more (simulations 256L12JH and 256L25JH) and less resolved simulations (simulation 64L25JH) for our radio/quasar mode with parameters from our best-fitting model (simulation 128L25JH, see Table 1).

In order to compare the simulated BH masses to their host galaxy bulge stellar mass as has been done for observations, we must decompose the stellar surface density profiles into an inner bulge and an outer disc component. Our bulge–disc decomposition is detailed in Appendix A.

Fig. 2 compares the BH comoving density as a function of redshift while varying the parameters of the AGN feedback model. The value of the BH density in our local Universe from Shankar et al. (2004) is overplotted with its 3σ uncertainty. Any AGN feedback model that pretends to model the cosmological growth of BHs should remain close to this data point at $z = 0$. However, the latter does not ensure that their growth is correct, as the observational point is only for redshift $z = 0$ and the data at larger redshifts are more sparse.

Fig. 3 shows the relationships between BH mass and their host galaxy stellar bulge mass as well as with the host's stellar velocity dispersion. These are good constraints for testing the co-evolution of BHs and their galaxy mass content. We represent the average value of the distribution of stellar bulge masses for a given bin of BH mass. Observational fits to BH mass ($M_{\rm BH}$) versus stellar mass ($M_{\rm s}$) (Häring & Rix 2004), and to BH mass ($M_{\rm BH}$) versus stellar velocity dispersion ($\sigma_{\rm s}$) (Tremaine et al. 2002) are overplotted with 3σ uncertainties. Fig. 4 shows the cosmic SFR to test the impact of varying AGN feedback parameters on the history of mass assembly of galaxies.

For these comparisons, we take simulation 128L25J as our reference model for the radio AGN feedback mode (see Table 1), and each parameter of the AGN feedback model is varied one by one. Even though this procedure does not explore the full parameter space, and does not ensure that another set of parameters are possible, it allows us to test the validity of our fiducial model as well as the dependency of the results on the variation of its parameters.

4.1 Parameter study

4.1.1 Varying efficiency ϵ_f

We test the effect of varying the efficiency $\epsilon_{\rm f}$ of the radio mode on the evolution of the cosmic BH density in Fig. 2(a). We compare three simulations with the implementation of the radio mode identical in every respect except for the efficiency, $\epsilon_{\rm f}$, which varies by approximately factors of 10 from 0.01 (simulation 128L25Je0.01) to 0.15 (simulation 128L25Je0.15) to 1 (simulation 128L25J). We also compare a run with BH growth but no AGN feedback (simulation 128L25BH). Efficiency values lower than $\epsilon_{\rm f} = 1$ produce larger BH densities at $z = 0$ than their observational counterparts. The case for which no AGN feedback energy is released (simulation 128L25BH) produces what we take to be the maximum attainable BH density. Even small amounts of energy ($\epsilon_{\rm f} = 0.15$ and 0.01) prevent BHs from growing to these maximum BH densities. The maximum possible efficiency $\epsilon_{\rm f} = 1$ predicts a slightly larger BH density than the data from Shankar et al. (2004), but is still within 3σ error.

The decrease in efficiency, $\epsilon_{\rm f}$, is compensated by larger accretion rates and more massive BHs, leading to the net result that the total amount of energy released by the AGN feedback is nearly independent of $\epsilon_{\rm f}$ (see also Booth & Schaye 2010). Fig. 5 substantiates this by showing the comoving cumulative AGN energy density as a function of redshift for different efficiencies. Indeed, even though AGN efficiency $\epsilon_{\rm f}$ is allowed to vary by two orders of magnitude, the total amount of energy liberated at $z = 0$ differs by less than a factor 2. This suggests that BHs adjust their masses so that the total energy liberated can blow the gas out from the galaxies and stop the accretion of gas. We do not present the AGN energy density evolution for the other simulations presented in Fig. 2 because aside from the simulations presented in panel (h), they all run with the same $\epsilon_{\rm f}$. Therefore, the AGN energy density, $\epsilon_{\rm AGN}$, can be deduced from their $\rho_{\rm BH}$ since $\epsilon_{\rm AGN} = \epsilon_{\rm f}\epsilon_{\rm r}c^2\rho_{\rm BH}$, where we recall that the radiative efficiency $\epsilon_{\rm r} = 0.1$ for all simulations and c is the speed of light.

Decreasing the efficiency leads to more massive BHs, but, interestingly, the stellar masses of these galaxies and their stellar velocity dispersions are not significantly impacted by the efficiency (Fig. 3a). This is confirmed by the cosmic SFR seen in Fig. 4(a), which shows little difference in SFR for different values of the AGN feedback efficiency, especially for $\epsilon_{\rm f} = 1$ and 0.15. BHs regulate themselves as well as the gas content of their host galaxy by injecting the same quantity of energy, regardless of $\epsilon_{\rm f}$, to unbind the cold gas component. The small decrease in the SFR for $\epsilon_{\rm f} = 0.01$ occurs because BH masses become comparable to their host galaxy masses so the BHs accrete gas instead of letting it form stars. This effect is more obvious when AGN feedback is not allowed but BH growth is permitted (simulation 128L25BH): the SFR is suppressed by one order of magnitude because BHs are more massive than the entire stellar content of their host galaxy and they consume all the fresh gas available.

We also plot in Fig. 4 the cosmic SFR for the simulation (simulation 128L25noAGN) without AGN feedback nor BHs but including our standard subgrid physics (cooling, star formation, SN feedback). In this case, the SFR is systematically higher than any of the simulations including BH growth with or without AGN feedback, and the difference is clearer at low and intermediate redshifts $z = 0$–4. This shows that AGN feedback efficiently suppresses star formation in galaxies, because it prevents gas overcooling and/or ejects large amounts of cold gas back into the circumgalactic medium (CGM).

We remark that the jet velocity depends on the efficiency as $u_{\rm J} \propto \sqrt{\epsilon_{\rm f}/\eta}$, and that the cumulative momentum imparted by all BHs is $Q \propto \sqrt{\eta/\epsilon_{\rm f}}\epsilon_{\rm AGN}$. Since $\epsilon_{\rm AGN}$ is almost independent of $\epsilon_{\rm f}$ (Fig. 5), Q depends only on $\sqrt{\eta/\epsilon_{\rm f}}$. Thus, lower efficiencies produce higher total momentum providing a possible explanation for why self-regulation is weaker for lower efficiency, $\epsilon_{\rm f}$. However, by varying the mass loading parameter, η, in Section 4.1.4, we show that self-regulation is controlled by the AGN feedback energy rather than momentum.

This first set of experiments exploring variations in $\epsilon_{\rm f}$ suggest that AGN feedback is a necessary element for the self-regulation of the growth of BHs, and that a high value of the AGN feedback efficiency in the radio mode $\epsilon_{\rm f} = 1$ is required to obtain realistic results on the co-evolution of BHs and galaxies.

4.1.2 Varying AGN energy delay: $\Delta M_{\rm d}$

We allow for a time-delay, represented by the $\Delta M_{\rm d}$ parameter, before releasing AGN energy in the radio mode. This parameter prescribes that the BH must grow by more than a $\Delta M_{\rm d}$ fraction

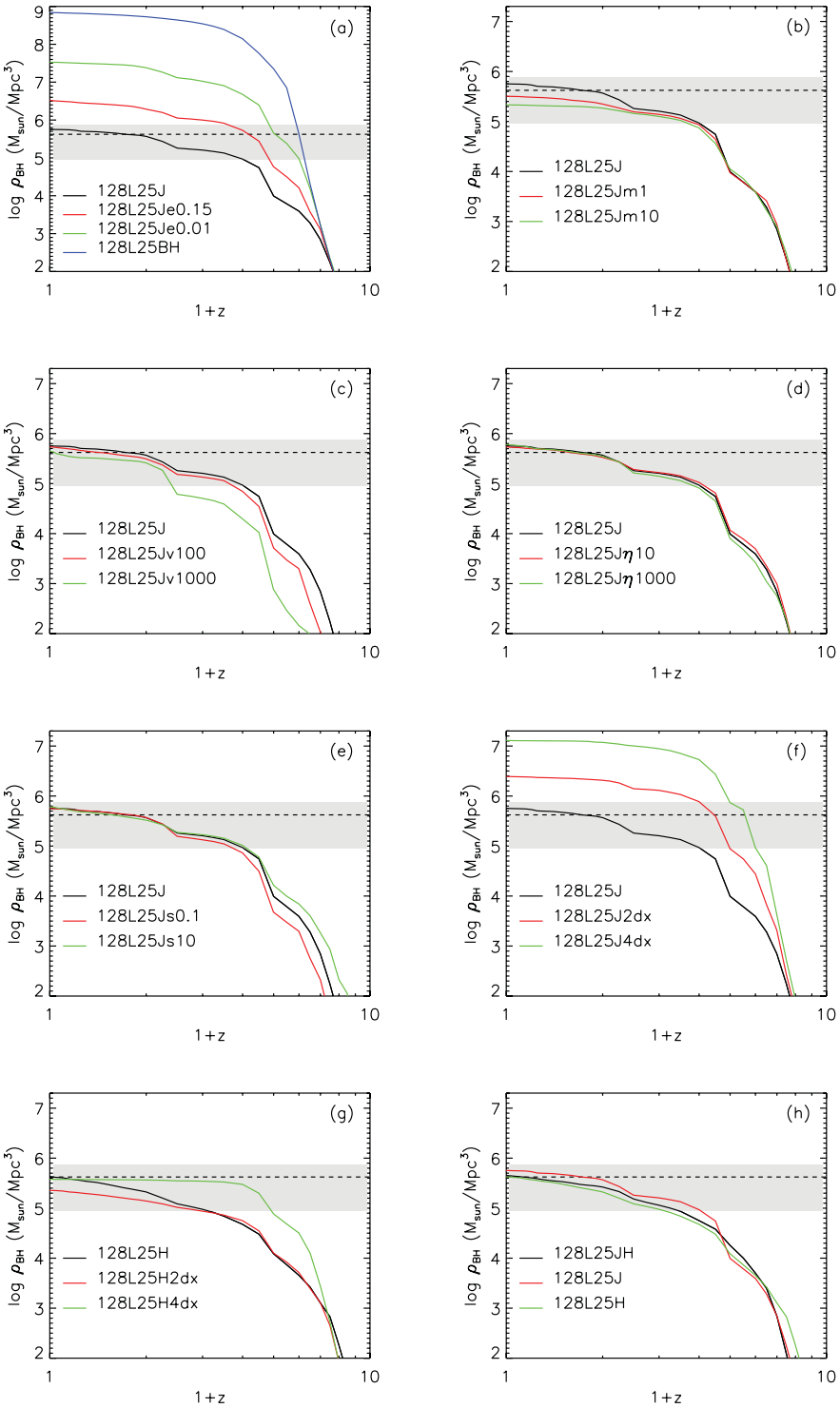

Figure 2. Comoving BH mass density as a function of redshift. (a) Varying the efficiency ϵ_f for the jet mode. (b) Varying the AGN energy delay ΔM_d for the jet mode. (c) Varying the maximum relative velocity u_{max} for the jet mode. (d) Varying the mass loading factor for the jet mode. (e) Varying the BH initial mass M_{seed} for the jet mode. (f) Varying the AGN input size r_{AGN} for the jet mode. (g) Varying the AGN input size r_{AGN} for the heating mode. (h) Varying the mode of the AGN feedback. The dashed line is the average BH mass density in our local Universe with its 3σ uncertainty (grey shaded area) from Shankar et al. (2004).

of its mass before releasing energy into a bipolar kinetic jet. We test three values for the time-delay parameter for the radio mode: $\Delta M_d = 0$ (simulation 128L25J), 1 per cent (simulation 128L25Jm1) and 10 per cent (simulation 128L25Jm10). At high redshift, varying the time-delay has a negligible impact on the evolution of the BH density (see Fig. 2b) because BHs grow close to the Eddington accretion rate $\dot{M}_{Edd}$ (see Section 5). Thus, typical growth time-scales of BHs are extremely short and even with a non-zero ΔM_d, energy is released almost continuously. The BH growth time-scale can be defined by $t_{BH} = M_{BH}/\dot{M}_{BH}$, and since $\dot{M}_{Edd} \propto M_{BH}$, it

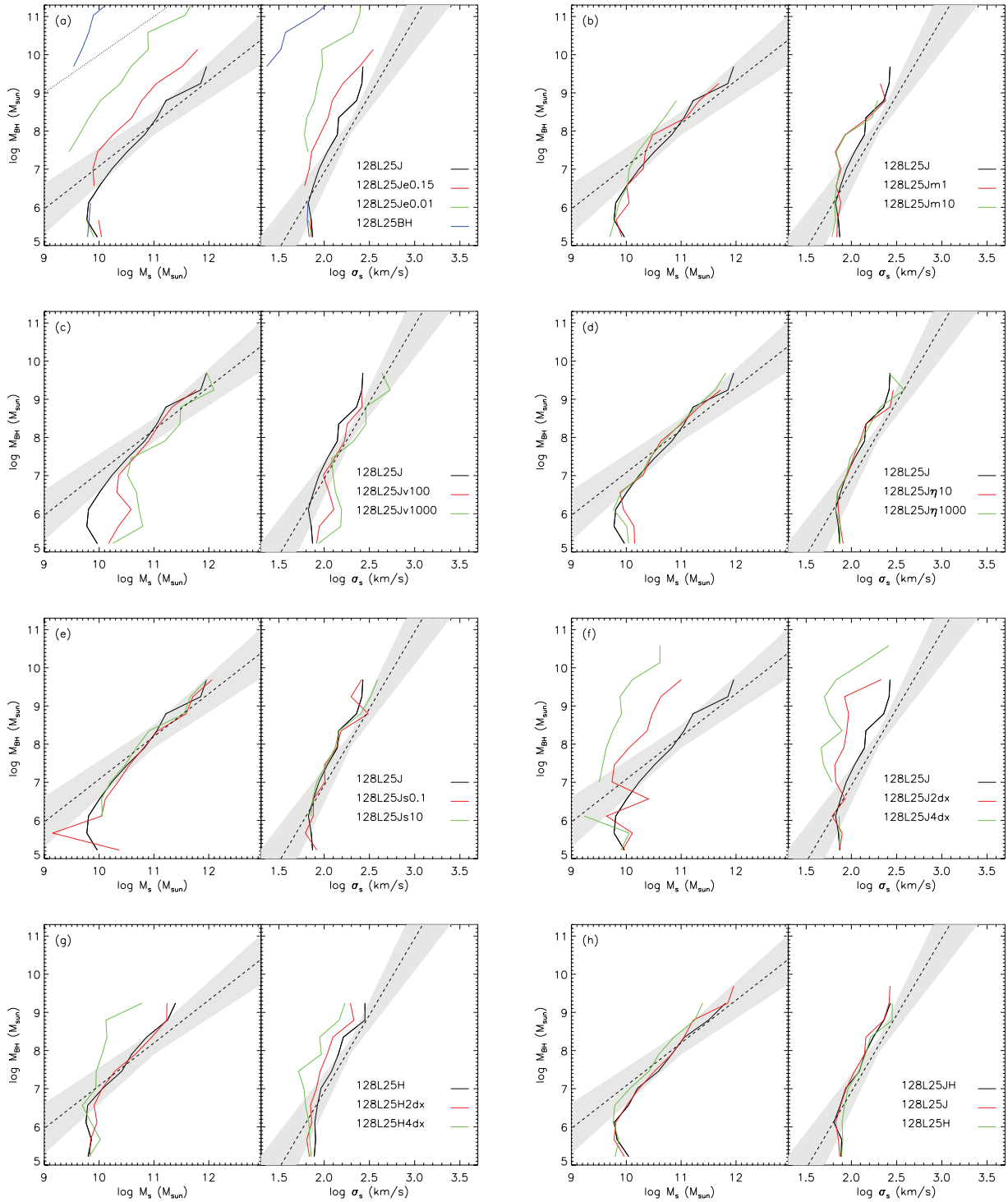

Figure 3. Panels (a)–(h) explore a variation of parameters in the same order as listed in the caption for Fig. 2. In each panel, we plot the BH mass as a function of stellar mass (left-hand plots), or as a function of stellar velocity dispersion (right-hand plots). Measurements are done at $z = 0$. We overplotted the observational laws as the dashed lines from Häring & Rix (2004) for the $M_{\rm BH}$–$M_{\rm s}$ relation and Tremaine et al. (2002) for the $M_{\rm BH}$–$\sigma_{\rm s}$ relation with their 3σ uncertainties. The dotted line in the left-hand panel of (a) indicates the relationship between log $M_{\rm BH}$ and log $M_{\rm s}$ when $M_{\rm BH} = M_{\rm s}$.

simplifies to

$$t_{\rm BH} = \frac{\epsilon_{\rm r}\sigma_{\rm T}c}{4\pi G m_{\rm p}}\chi^{-1} \simeq 45.5\,\chi^{-1}\,{\rm Myr}\,, \qquad (18)$$

depending only on the Eddington accretion ratio χ. As a consequence, BHs accreting gas at high Eddington ratios have a more continuous and hence immediate impact on the surrounding gas than BHs in a low accretion mode.

However, at low redshift, results from simulations with different time-delays diverge. When a delay is permitted (simulations 128L25Jm1 and 128L25Jm10), the final BH density at $z = 0$ is smaller than when energy is continuously deposited (simulation

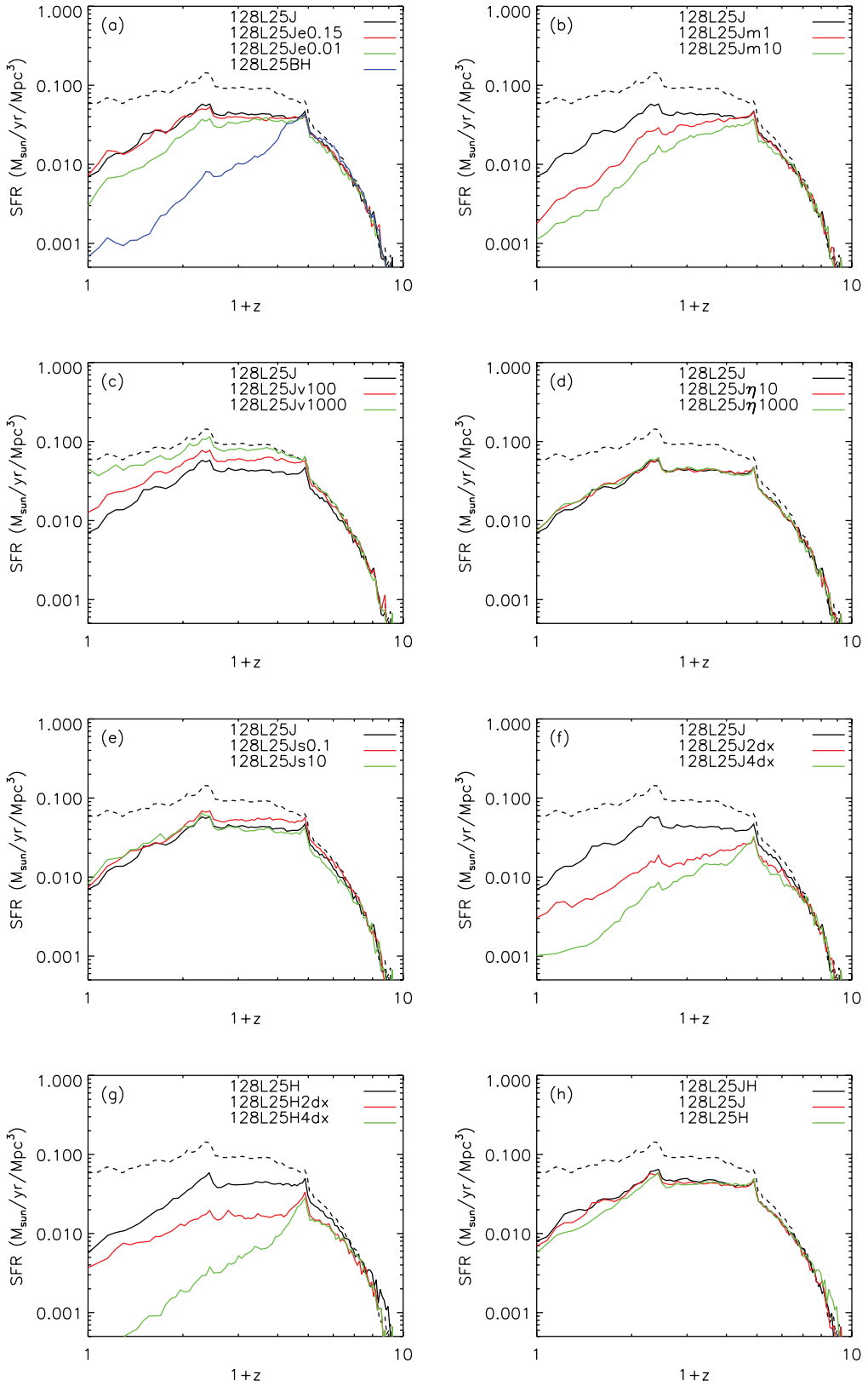

Figure 4. Panels (a)–(h) explore a variation of parameters in the same order as listed in the caption for Fig. 2. Comoving SFR as a function of redshift. The dashed line corresponds to simulation 128L25noAGN which does not include AGN feedback.

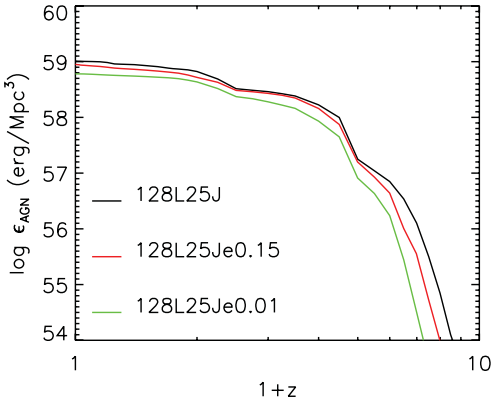

Figure 5. Cumulative comoving AGN energy density as a function of redshift for different AGN feedback efficiencies ϵ_f.

128L25J). This is linked to easier BH self-regulation with larger ΔM_d. Since BHs accrete gas at low redshift in a low-Eddington accretion regime, with non-zero ΔM_d there are larger stretches of time before a significant amount of energy is released. As a result of this accumulation of energy, BHs release a comparable amount of energy to the $\Delta M_d = 0$ case (simulation 128L25J) but in a shorter amount of time, allowing for fewer but stronger AGN luminosity bursts. Thus, BHs can more easily self-regulate with smaller duty cycles (Pope 2011).

The time-delay parametrized by ΔM_d also modifies the relationships between BH masses and their host galaxy properties (Fig. 3b): BHs with masses $M_{BH} > 10^7\,M_\odot$ sit in lower mass galaxies. This effect is stronger for $\Delta M_d = 10$ per cent (simulation 128L25Jm10) for which stellar masses are reduced by an order of magnitude for the most massive galaxies compared to the simulation with a continuous ($\Delta M_d = 0$) injection rate (simulation 128L25J). The maximum stellar velocity dispersions of the host galaxies of the most massive BHs are reduced as a direct consequence of the reduction of the stellar mass in them. The effect is even more apparent for the intermediate BHs ($10^7 < M_{BH} < 10^8\,M_\odot$) in simulation 128L25Jm10, where a clear deviation from the observational fit is observed. This suggests that the stellar velocity dispersion in massive galaxies is essentially controlled by the cold baryon content, and not by the total halo mass, which is hardly modified by the AGN feedback.

Finally, in Fig. 4(b), we see that the intermediate- and low-redshift SFR depends a lot on the ΔM_d parameter, which is not surprising since we saw that it also influences the M_{BH}–M_s relationships. Large ΔM_d more efficiently suppresses the total SFR and has more impact on the gas content in galaxies because they undergo shorter and stronger episodes of AGN feedback. This parameter study teaches us that not only is the amount of energy deposited important, but also the duration of the energy release plays a key role in unbinding the gas content of galaxies. Galactic gas exposed to a small deposit of energy can efficiently return to equilibrium by the gas dynamics and by the short cooling times involved in high-density gas.

4.1.3 *Varying maximum relative velocity: u_{max}*

We measure the effect of varying the maximum allowed velocity u_{max} of the gas relative to the BH in the Bondi formula (equation 3). Increasing our fiducial value $u_{max} = 10\,km\,s^{-1}$ (simulation 128L25J) up to $u_{max} = 1000\,km\,s^{-1}$ (simulation 128L25Jv1000) decreases the overall BH densities (see Fig. 2c), but values are still consistent with the observations. As we would expect, larger values of the relative velocity inhibit the growth of BHs. This effect already comes into play at high redshift, when mergers between galaxies are numerous sometimes resulting in violent excursions of BHs in their host galaxies, leading to large BH velocities relative to the dense gas component.

This spurious effect comes from our inability to resolve the very small scales of the ISM within which BHs should be embedded. Some authors have circumvented this problem by correcting the positions of BHs when they move too far from the gravitational potential well (Volker Springel, private communication). Here we prefer to adopt a more straightforward approach by limiting the maximum gas velocity relative to the BH in the formula for gas accretion (equation 3) rather than changing the position of the BH. This also allows us to follow the BHs that are ejected from their galaxies by strong tidal effects and gravitational friction, as material is stripped from galaxy satellites falling into massive haloes.

Fig. 3(c) shows that different u_{max} produce different BH masses and host galaxy stellar properties. When large relative velocities are permitted (simulations 128L25Jv100 and 128L25v1000), galaxies tend to be more massive, and, as a result, have larger velocity dispersions. Even though the total BH density at $z = 0$ is hardly changed for different u_{max}, the SFR is very sensitive to u_{max} (Fig. 4c). Large u_{max} values tend to nullify the effect of the AGN feedback on the total SFR, and the SFR converges to the case where feedback from AGN is not allowed. As a consequence, $u_{max} = 10\,km\,s^{-1}$ corresponds to a choice that allows for a non-spurious quenching of the accretion rate while keeping the dynamics of the BH particles completely self-consistent.

4.1.4 *Varying jet mass loading factor: η*

The mass loading factor η is a free parameter of the radio mode for AGN feedback that controls the velocity the jet would have if it were propagating into a void. We compare three simulations with $\eta = 10$ (simulation 128L25Jη10), 100 (simulation 128L25J) and 1000 (simulation 128L25Jη1000). As can be seen from Figs 2(d), 3(d) and 4(d), the BH and galaxy properties are very insensitive to the adopted values of the mass loading factor of the jet. The reason is that the jet couples to the gas in its surroundings and AGN feedback becomes ineffective when the energy of the jet becomes comparable to the binding energy of the gas. The only difference introduced by η is that depending on its value, the jet will go more or less quickly into equilibrium with the gas, but as the liberated energy is the same regardless of η, the same amount of gas is impacted by the jet.

We insist on the fact that the total imparted momentum $Q \propto \sqrt{\eta/\epsilon_f}\epsilon_{AGN}$ loses its ϵ_{AGN} dependence because the latter is constant, given that BH densities are constant for different η (Fig. 2d) and ϵ_f is the same ($\epsilon_f = 1$) for the three simulations we are comparing. As a result, $Q \propto \sqrt{\eta/\epsilon_f}$ as it was for the case with varying efficiencies (Section 4.1.1). However, since ϵ_f is a constant, momentum, Q, only depends on η for the set of simulations compared in this section. Figs 2(d), 3(d) and 4(d) show that BHs and their host galaxy properties do not depend on η, or equivalently on the momentum Q. Therefore, we conclude that BHs and their host galaxy properties are only sensitive to jet energies, not their momenta.

4.1.5 *Varying initial BH mass: M_{seed}*

We vary the BH initial seed mass by choosing values as small as $10^4\,M_\odot$ (simulation 128L25Js0.1) and as large as $10^6\,M_\odot$ (simulation 128L25Js10). Simulation 128L25J has $M_{seed} = 10^5\,M_\odot$. As

seen in Fig. 2(e), this has little effect on the final ($z = 0$) BH density. Differences appear at high redshift, when most of the contribution to the BH density comes from BHs with mass close to their initial seed mass. Thus, the choice for the initial seed BH mass has an important impact on the BH density at high redshift but memory of it is rapidly erased when BHs grow to values larger than their initial mass.

The Magorrian relationships are almost unchanged by different choices for $M_{\rm seed}$ (Fig. 3e). However, a careful inspection of the cosmic SFR (Fig. 4e) shows that the SFR is slightly different at high redshift $z = 2$–4: lower seed BH masses result in a larger SFR because it takes more time for BHs to reach a self-regulated equilibrium with their environment.

4.1.6 Varying AGN input size: $r_{\rm AGN}$

We vary the size of the AGN energy input region for both the radio and the quasar mode to deduce its effect on BH self-regulated growth.

First, we explore the impact of varying the extent of the jet in the radio mode through a series of three simulations where the AGN input region varies from $r_{\rm AGN} = \Delta x$ (simulation 128L25J) to $2\Delta x$ (simulation 128L25J2dx) to $4\Delta x$ (simulation 128L25J4dx). Fig. 2(f) shows that large jet sizes for the radio mode lead to large BH densities. Since the energy is spread over larger regions when we increase the jet size, the gas close to the BH is impacted less by larger jets, thereby accreting more easily on to the BH, hindering its self-regulation, and producing larger BH densities.

Fig. 3(f) shows that jets with sizes larger than Δx lead to larger BH masses and lower host galaxy stellar masses, with unrealistic values compared to observations. Also, the SFR is strongly suppressed when larger jet sizes are chosen (Fig. 4f). Thus, large energy injection regions for the radio AGN feedback mode have more impact on galaxy formation because BHs become more massive and as a result inject more energy to the surrounding gas. However, as the BHs are much more massive than what is predicted from the $M_{\rm BH}$–$M_{\rm s}$ relationship, they must release more energy to self-regulate.

For the quasar mode, the behaviour is different. Again, we ran three simulations to explore the effect of changing the size of the energy input region: from $r_{\rm AGN} = \Delta x$ (simulation 128L25H) to $2\Delta x$ (simulation 128L25H2dx) to $4\Delta x$ (simulation 128L25H4dx). Fig. 2(g) shows that the trend of BH density with bubble size is non-linear. Doubling the radius ($r_{\rm AGN} = 2\Delta x$) gives a similar BH density evolution as obtained with $r_{\rm AGN} = \Delta x$ down to redshift $z = 1.5$ but shows a drop in BH density below this redshift. Increasing the radius by four times ($r_{\rm AGN} = 4\Delta x$), on the other hand, leads to a larger BH density at high redshift, which converges to the same value as the $r_{\rm AGN} = \Delta x$ case at $z = 0$.

For the Magorrian relations, choosing $r_{\rm AGN} = 2\Delta x$ rather than $r_{\rm AGN} = \Delta x$ for the quasar mode leads to smaller stellar velocity dispersions but very similar stellar masses. It seems that for the $r_{\rm AGN} = 2\Delta x$ case, even though the SFR is significantly decreased compared to the $r_{\rm AGN} = \Delta x$ case (see Fig. 4g), the decrease in the BH density keeps the BH mass versus host galaxy stellar mass relation unchanged. However, a larger energy injection region ($r_{\rm AGN} = 4\Delta x$) has a dramatic impact on the final galaxy stellar masses and the evolution of the SFR. Both are significantly diminished. Similar behaviour has been found by Booth & Schaye (2009), where increasing the number of neighbouring SPH particles affected by the AGN bubble decreases the SFR and increases the BH density.

Both the radio and quasar modes experience a decline in SFR as the size of the injection region (Figs 4f and g) increases because large energy injection regions extend to less dense regions which are easier to impact. By blowing out the reservoir of hot gas, accretion on to galaxies and hence SFR is suppressed. However, the self-regulation of BHs is somehow very different for the two modes. The difference resides in the very nature of energy deposit. Jets put momentum and kinetic energy into the gas and eventually some of this energy is transformed into thermal energy through shocks, but the more extended the jet, the weaker the shock. Fig. 6 illustrates this effect in the high-redshift Universe: the quasar mode with large $r_{\rm AGN} = 4\Delta x$ (simulation 128L25H4dx) inflates larger and hotter bubbles than the radio mode with the same initial jet extent (simulation 128L25J4dx). As the accretion rate is very sensitive to the temperature of the gas, $\dot{M}_{\rm BH} \propto T^{-1.5}$, it is more difficult for jets than for thermal bubbles to self-regulate the growth of BHs. This is why BH densities for the jet mode with $r_{\rm AGN} = 4\,\Delta x$ are larger than for the heating mode.

As a final remark, these particular numerical experiments demonstrate that the injection of energy through AGN feedback (but it is true for any type of feedback, see Dalla Vecchia & Schaye 2008 for a similar discussion about SN feedback) is a delicate process that cannot be naively decoupled from the gas dynamics up to large distances, and must be handled with great care.

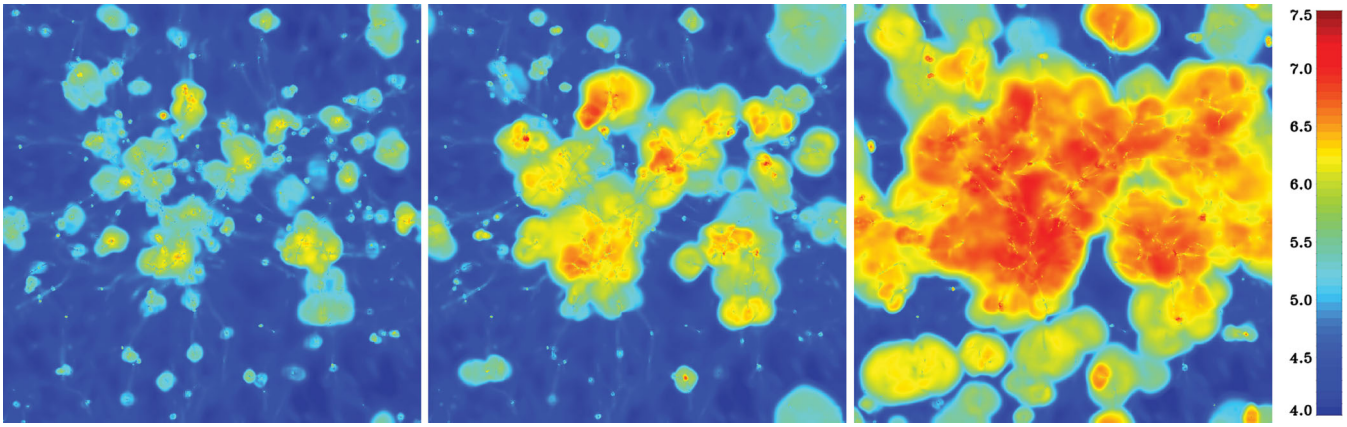

Figure 6. Projected temperatures of simulations 128L25JH (left-hand panel), 128L25J4dx (middle panel) and 128L25H4dx (right-hand panel) at $z = 3$. The colour code is in log (K) units. Images are $6.25\,h^{-1}$ Mpc physical on a side.

4.1.7 Comparing radio mode and quasar mode

We compare the choice of using the radio mode (jet mode) to the quasar mode (heating mode), as well as to a combination of both modes. We found a set of parameters [$\epsilon_{\rm f} = 1$ (radio) and 0.15 (quasar); $\Delta M_{\rm d} = 0$; $u_{\rm max} = 10\,{\rm km\,s^{-1}}$; $\eta = 100$; $M_{\rm seed} = 10^5\,{\rm M_\odot}$; $r_{\rm AGN} = \Delta x$] consistent with observations for the radio mode and quasar mode used individually and use this same set of parameters for the dual radio–quasar mode (simulation 128L25JH). With the dual radio–quasar mode of AGN feedback, the feedback from BHs switches to the radio mode when the Eddington accretion ratio $\chi \leq \chi_{\rm radio}$ and to the quasar mode when $\chi > \chi_{\rm radio}$.

Fig. 2(h) shows that the radio (jet) mode (simulation 128L25J) produces a larger BH density than the quasar (heating) mode (simulation 128L2H), suggesting that the quasar (heating) mode is slightly more efficient at self-regulating BH growth (although a slightly smaller AGN efficiency, $\epsilon_{\rm f}$, of the quasar mode would reduce this difference). As expected, the dual-mode feedback (simulation 128L25JH) gives a BH density value which is between the value for individual modes.

There are some 'jumps' in the evolution of the BH density at $z = 4$ and 1.5 for the radio (jet) mode (Fig. 2h), which are not present for the quasar (heating) mode. This effect is also seen in the SFR evolution (Fig. 4h), both for the radio (jet) and quasar (heating) modes. It comes from the triggering of a new maximum level of refinement at these redshifts. With one more level of refinement, the force on small scales is better resolved and the gravitational potential well is deeper. The gas can therefore get compressed to higher densities, leading to significant boosts in its SFR and the accretion rate on to the BHs. This effect is less pronounced for the quasar (heating) mode where the heating supplies an additional pressure support to the gas, tending to erase this spurious numerical feature.

Fig. 3(h) shows that the stellar mass in the quasar (heating) mode (simulation 128L25H) is slightly reduced compared to the radio (jet) mode (simulation 128L25J), which is another signature of a slightly more efficient feedback in the quasar (heating) mode. The mode of AGN feedback has very little effect on the BH mass versus stellar velocity dispersion. A combination of the two modes shows deviations, especially at low $M_{\rm s}$ and $\sigma_{\rm s}$, from observed BH mass versus stellar mass relationships and observed BH mass versus stellar velocity dispersion relations. However, in Fig. 9 (shown later) we show that this is a consequence of the limited numerical resolution, and that we converge to the observational measurements with increased resolution.

Fig. 4(h) shows that the SFRs for the different AGN feedback modes are almost undistinguishable. A small difference can be seen at low redshift $z = 0$–1 where the quasar (heating) mode seems to more efficiently suppress the total SFR than the radio (jet) mode. The latter explains the difference seen in the BH mass versus stellar mass relationships (Fig. 3h).

Finally, this parameter study has allowed us to choose the best-fitting parameters for our dual radio–quasar AGN feedback model compared to observations, namely the parameters used in simulation 128L25JH (see Table 1).

4.2 Resolution study

In order to test the convergence of our models, we vary the resolution, by changing the DM mass, cell size and box size. We run five different simulations with our fiducial model for AGN feedback (simulations 256L12JH, 256L25JH, 128L25JH, 64L25JH, 256L50JH and 128L50JH), with three different box sizes $L_{\rm box} = 12.5$, 25 and $50\,h^{-1}$ Mpc, and four different resolutions $\{M_{\rm DM} = 3.5 \times 10^9\,{\rm M_\odot}$, $\Delta x = 3.04\,h^{-1}$ kpc$\}$ (simulations 64L25JH and 128L50JH), $\{M_{\rm DM} = 4.4 \times 10^8\,{\rm M_\odot}$, $\Delta x = 1.52\,h^{-1}$ kpc$\}$ (simulations 128L25JH and 256L50JH), $\{M_{\rm DM} = 5.5 \times 10^7\,{\rm M_\odot}$, $\Delta x = 0.76\,h^{-1}$ kpc$\}$ (simulation 256L25JH) and $\{M_{\rm DM} = 6.9 \times 10^6\,{\rm M_\odot}$, $\Delta x = 0.38\,h^{-1}$ kpc$\}$ (simulation 256L12JH) (see Table 1 for details).

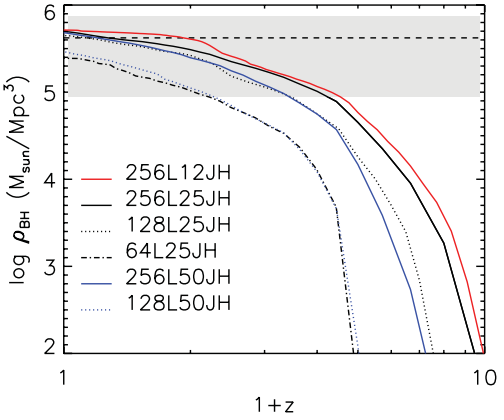

Figure 7. Comoving BH mass density as a function of redshift for different box sizes and resolutions. The grey shaded area is the BH mass density in our local Universe with its 3σ uncertainty from Shankar et al. (2004).

The BH densities shown in Fig. 7 slowly converge to the same value at $z = 0$ when the resolution is increased. Even though low-resolution simulations (simulations 64L25JH and 128L50JH) are within the observational 3σ error bars, they tend to underestimate the BH density at all redshifts compared to more resolved simulations. Intermediate-resolution simulations with $\{M_{\rm DM} = 4.4 \times 10^8\,{\rm M_\odot}$, $\Delta x = 1.52\,h^{-1}$ kpc$\}$, which correspond to runs 128L25JH and 256L50JH, have already converged at $z = 0$ and differ only by $\sim$10 per cent at $z = 0$ from the simulation (simulation 256L25JH) with one additional refinement level. However, at high redshift, the difference is larger because galaxies in the field with intermediate BH masses contribute more to the total BH density (Fig. 8), and some of these galaxies are not resolved in simulations 128L25JH and 256L50JH with lower resolution.

The same convergence can be seen in the BH mass versus stellar mass relationships (Fig. 9, left-hand panel). When the resolution is increased, these relationships converge to the same value close to the observations from Häring & Rix (2004). There is a departure from the observational constraint for the least massive galaxies that reside in haloes that are barely resolved. For PM codes, DM haloes with more than $\sim$1000 DM particles are followed with sufficient force resolution (O'Shea et al. 2005; Heitmann et al. 2008). Thus, the break observed at the low galaxy mass end corresponds to this low limit.

For example, simulation 256L25JH has a DM mass resolution $M_{\rm DM} = 5.5 \times 10^7\,h^{-1}\,{\rm M_\odot}$, which gives a minimum DM halo mass $M_{\rm h,min} \sim 8 \times 10^{10}\,{\rm M_\odot}$ that corresponds to a total gas content of $f_{\rm b} M_{\rm h,min} \sim 10^{10}\,{\rm M_\odot}$. A non-negligible fraction of the baryon content is locked into stars. Assuming 25 per cent for this fraction, we obtain $M_{\rm s,min} \sim 2.5 \times 10^9\,{\rm M_\odot}$, that is, the value of the bulge stellar mass where the break appears in the $M_{\rm BH}$–$M_{\rm s}$ relationship. Indeed, the same behaviour is seen when relating the BH mass to the stellar velocity dispersion of their host galaxy: low-mass BHs hosted by galaxies with low stellar velocity dispersion show a significant

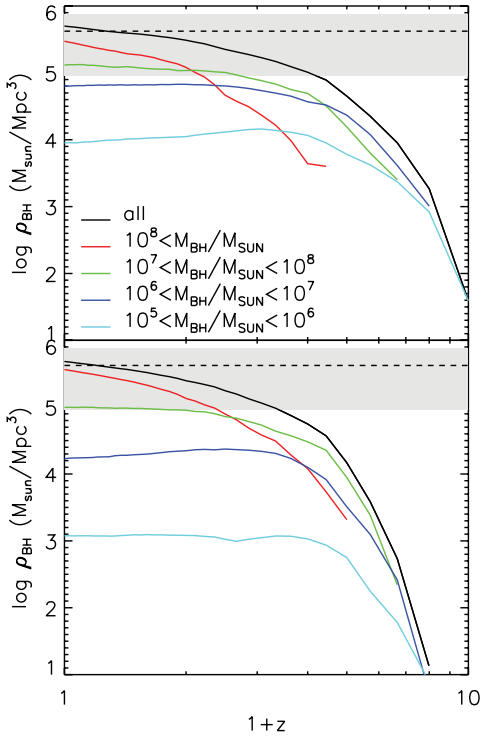

Figure 8. Comoving BH mass density as a function of redshift with contributions from different BH mass ranges. The upper panel is for simulation 256L25JH and the bottom panel is for simulation 256L50JH. The grey shaded area is the BH mass density in our local Universe with its 3σ uncertainty from Shankar et al. (2004).

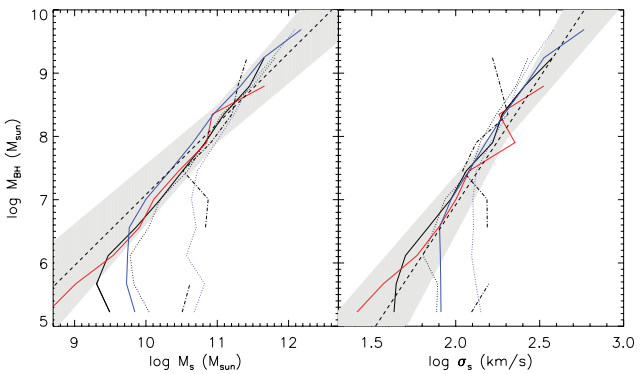

Figure 9. BH mass as a function of stellar mass (left-hand panel panel), or as a function of stellar velocity dispersion (right-hand panel) at $z = 0$ for different box sizes and resolutions. The colour code is the same as for Fig. 7. We overplotted the observational laws as the dashed lines from Häring & Rix (2004) (left-hand panel) and Tremaine et al. (2002) (right-hand panel) with their 3σ uncertainties.

deviation from the constraints of Tremaine et al. (2002) (Fig. 9, right-hand panel).

Fig. 10 shows the cosmic SFR for simulations with our fiducial model for AGN feedback and for different resolutions and box sizes. It appears that this quantity is strongly dependant on DM mass resolution at high redshift because of two effects. When increasing the resolution, smaller haloes which contribute a lot to the total SFR at high redshift are resolved. Furthermore, structures that were already resolved at lower resolution collapse earlier when resolution is increased, thus, forming stars earlier. This effect is well constrained and, if one has sufficient resolution, the numerical solution converges to its analytical prediction (Rasera & Teyssier 2006).

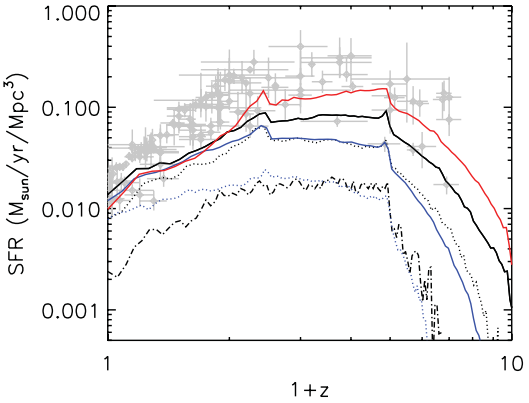

Figure 10. Comoving SFR as a function of redshift for different box sizes and resolutions. The colour code is the same as for Fig. 7. The light grey circles with error bars correspond to observational points from Hopkins & Beacom (2006).

We point out that Tree codes or Tree-PM codes, such as the GADGET code (Springel 2005), have a better force resolution in the early Universe than PM codes. The former therefore can follow the formation of smaller haloes (down to 10 DM particles) with equivalent initial conditions. This is why, in general, the convergence in the SFR with such codes is more rapidly obtained, even though 10 particles per halo are not sufficient to properly treat the gas dynamics.

Even though the convergence for a given box size is not reached, the SFR for the box size 25 h^{-1} Mpc slowly converges at $z = 0$. It increases by a factor of 4 from simulation 64L25JH to 128L25JH, and by a factor of 2 from simulation 128L25JH to 256L25JH. Changing box size at constant resolution has also some non-negligible effect on the SFR at low redshift. The reason is that at low redshift the SFR is essentially dominated by massive galaxies. Thus, with larger box sizes, very massive clusters of galaxies are more likely to be present (Davis et al. 2011). Finally, we conclude that getting convergence for the SFR is extremely difficult, because it requires both a large box size and a small DM mass resolution, which can only be achieved with tremendous computational power.

It is interesting to note that the DM resolution effect on SFR is also present for the simulations without AGN feedback (Fig. 11), which proves that this effect is uncorrelated to AGN feedback but only to DM mass resolution and box size. AGN feedback is most efficient at suppressing the SFR at low redshift when massive structures such as groups and clusters of galaxies are formed. At high redshift, the SFR is dominated by galaxies in the field which are not progenitors of groups or clusters of galaxies; thus, the effect of AGN feedback is less visible.

There is an interesting behaviour of the simulation without AGN feedback and without SN feedback (simulation 256L25noSNAGN, dotted line in Fig. 11) compared to the simulation without AGN feedback but with SN feedback (simulation 256L25noAGN, dashed line in Fig. 11). At high redshift, SN feedback reduces the SFR, because galaxies form large-scale galactic winds (Springel & Hernquist 2003; Dubois & Teyssier 2008) that remove some baryons from them and prevent some gas from collapsing into them. However, as time goes by, structures become more massive and the ram-pressure confinement from the halo becomes higher, preventing galactic winds from escaping the discs. As a result they develop galactic fountains (Dubois & Teyssier 2008). The SN feedback also

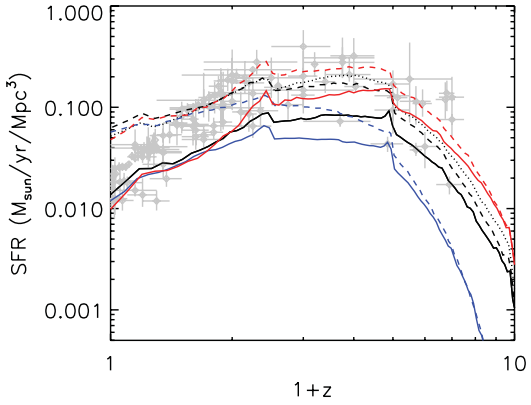

Figure 11. Comoving SFR as a function of redshift for different box sizes $L_{\rm box} = 12.5\,h^{-1}$ Mpc (red), $L_{\rm box} = 25\,h^{-1}$ Mpc (black) and $L_{\rm box} = 50\,h^{-1}$ Mpc (blue) including SN and AGN feedback (solid line), including SN feedback and no AGN feedback (dashed line) and without SN or AGN feedback (dotted line). The light grey circles with error bars correspond to observational points from Hopkins & Beacom (2006).

enriches the gas with metals and enhances the gas cooling rates, developing stronger accretion flows and SFRs at late times, when metals are confined in haloes (Dubois et al. 2011). Thus, the feedback of SN has a negative impact on SFR at high redshift but a positive effect at low redshift due to metal enrichment.

5 REDSHIFT EVOLUTION OF BH PROPERTIES

The way BHs acquire their mass and how they liberate energy to the gas is important for constraining their co-evolution with their host galaxy. Observations of the relationships between BH masses and galaxy masses at high redshift are extremely difficult because the luminosity of host galaxies of massive BHs is dominated by the AGN component. Despite these difficulties, an increasing amount of data is suggesting that the $M_{\rm BH}$–$M_{\rm s}$ relationship shows some positive evolution with redshift (McLure et al. 2006; Peng et al. 2006; Shields et al. 2006; Salviander et al. 2007; Bennert et al. 2010; Decarli et al. 2010; Merloni et al. 2010). Simulations can provide insight into the time-evolution of these relationships.

Fig. 12 shows the BH mass versus bulge stellar mass at different redshifts for the 25 h^{-1} Mpc (256L25JH) and the 50 h^{-1} Mpc (256L50JH) simulations. At high redshift, the BHs are more massive than the $z = 0$ Magorrian relation would predict, given the stellar mass of their host galaxy. This is supported by observations (see Merloni et al. 2010). We evaluate this deviation from the $z = 0$ relationships by measuring the median of the distribution of the $M_{\rm BH}/M_{\rm s}$ measurements for simulations 256L25JH and 256L50JH in Fig. 13. We explore the effect of removing from the sample BHs with masses smaller than a mass threshold to see if there is a BH mass for which the deviation is most pronounced. Like the observations (Decarli et al. 2010; Merloni et al. 2010) for which BH masses are larger than a few 10^7 M$_\odot$, we observe a positive trend with redshift of the $M_{\rm BH}/M_{\rm s}$ ratio. In our simulations, the ratio is larger for more massive BHs, but the trend is independent of the BH mass threshold. Quantifying the trend, we find $M_{\rm BH}/M_{\rm s} \propto (1+z)^{\alpha_{\rm s}}$ with $\alpha_{\rm s} = 0.42 \pm 0.09$ for simulation 256L25JH and $\alpha_{\rm s} = 0.42 \pm 0.06$ for simulation 256L50JH when fitting our simulation data on BHs with masses larger than $>5 \times 10^7$ M$_\odot$ between redshifts $z = 0$ and 3. This is in relatively good agreement with the value $\alpha_{\rm s} = 0.68 \pm 0.12$ measured in the observational data by Merloni et al. (2010),

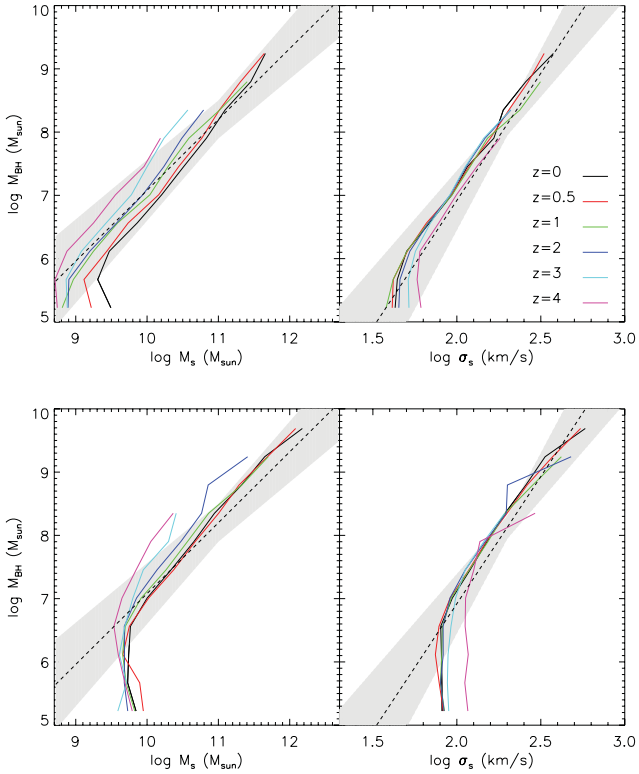

Figure 12. For each plot: BH mass as a function of stellar mass (left-hand panels), or as a function of stellar velocity dispersion (right-hand panels). Measurements are done at different redshifts as labelled in the upper right-hand panel. The top plot corresponds to simulation 256L25JH and the bottom plot to simulation 256L50JH. We overplotted the observational laws as the dashed lines from Tremaine et al. (2002) and Häring & Rix (2004) with their 3σ uncertainties.

and with numerical simulations from Di Matteo et al. (2008) ($\alpha_{\rm s} = 0.5$) and Booth & Schaye (2011) ($\alpha_{\rm s} = 0.52 \pm 0.05$).

The increase in the $M_{\rm BH}/M_{\rm s}$ ratio reflects the different accretion modes on to the BH and the gas content and properties at different redshifts. Fig. 14 illustrates how massive BHs grow with time, with very fast accretion of gas at high redshift due to the presence of a cold and dense ISM. The accretion proceeds by bursts accompanied by large releases of AGN energy that temporarily delay the accretion on to the BH. At high redshift, two kinds of accretion occur: the accretion of a diffuse component that can eventually shock and virialize the gas in the halo and accretion of dense filaments of gas (Kereš et al. 2005; Ocvirk, Pichon & Teyssier 2008; Brooks et al. 2009; Dekel et al. 2009). The gaseous filaments feed galaxies so that their BHs can grow to larger masses, pre-heat their protocluster environment and remove gas, thereby halting the adiabatic contraction in the protocluster cores.

For simulations 256L50JH and 256L50noAGN, Fig. 15 shows the fraction of baryons in galaxies in the form of a cold gas component with gas density larger than >0.1 H cm^{-3} for different stellar masses at different redshifts. It appears that more massive galaxies have lower gas fractions that decline with time. This can be explained by two effects. First, galaxies efficiently consume their gas to form stars without replenishing their cold gas content quickly enough to maintain a constant specific SFR. Secondly, AGN feedback reduces the amount of cold gas available in galaxies by ejecting dense material into the CGM. Hence, AGN feedback coupled to a vigorous consumption of gas via star formation reduces the gas content in

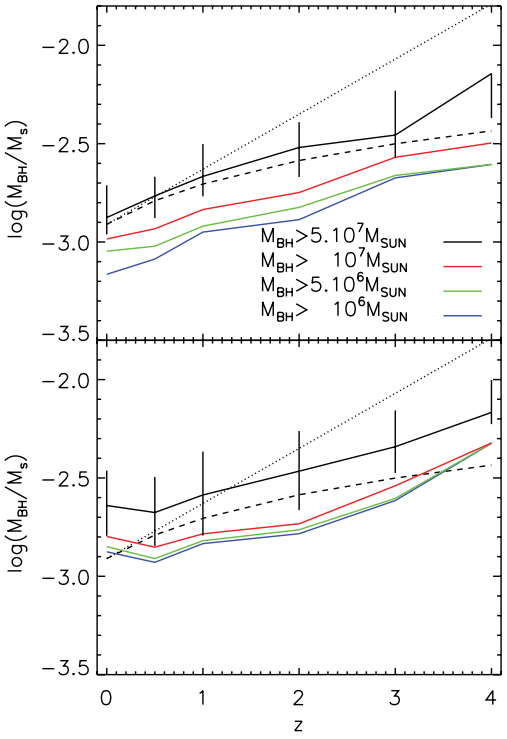

Figure 13. Median value of the BH mass ($M_{\rm BH}$) over its host stellar bulge mass ($M_{\rm s}$) as a function of redshift for simulations 256L25JH (upper panel) and 256L50JH (bottom panel). Different colours correspond to different BH mass cut-offs used to get the median value of $M_{\rm BH}/M_{\rm s}$. The vertical error bars correspond to first and third quartile limits of the distribution of points. The dotted line is the trend from a fit to observations in Decarli et al. (2010) and the dashed line is the trend from Merloni et al. (2010).

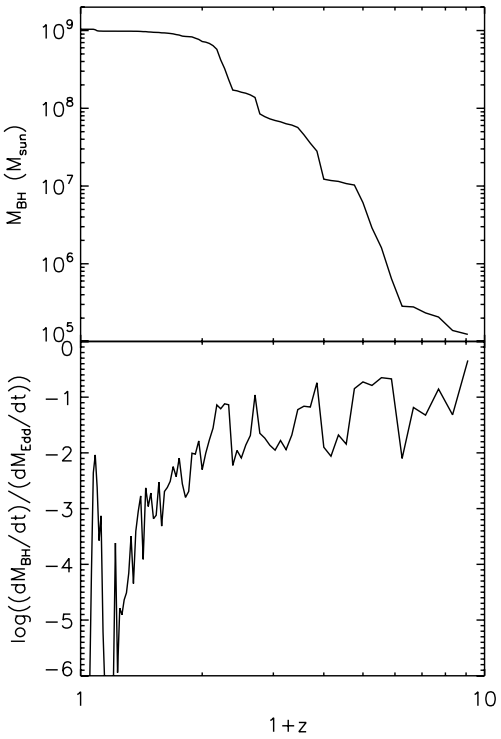

Figure 14. BH mass as a function of redshift for the most massive BH in simulation 256L12JH (upper panel) and the logarithm of the ratio of its accretion rate to the Eddington accretion limit (bottom panel).

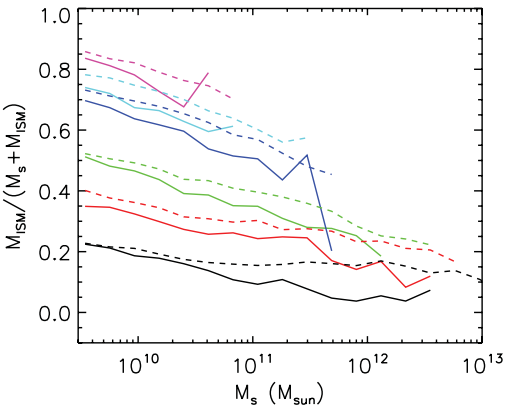

Figure 15. Average gas mass in the disc ($M_{\rm ISM}$) over the total mass of gas plus stars ($M_{\rm ISM} + M_{\rm s}$) as a function of the total stellar mass ($M_{\rm s}$) for different redshifts $z = 4$, 3, 2, 1, 0.5 and 0 from the top to bottom for simulations 256L50JH (solid lines) and 256L50noAGN (dashed lines).

galaxies, resulting in a much lower accretion rate at low redshift, where BHs enter a low-Eddington accretion regime (Fig. 14).

The behaviour of the accretion rate shown in Fig. 14 for a single BH is common to BHs of a large range of masses. We represent the number-weighted diagram of the BH accretion luminosity $L_{\rm acc} = \dot{M}_{\rm BH}c^2$ versus the BH mass at different redshifts in Fig. 16 for the 50 h^{-1} Mpc simulation (simulation 256L50JH). At high redshift ($z = 4$), the accretion proceeds in a high-accretion regime, where most BHs accrete very close to their Eddington accretion rate and release their energy in a 'quasar' mode. Some of the most massive BHs have already entered a low-accretion regime suggesting that the most massive objects are the first to self-regulate their gas content. Then, as the simulation evolves, more and more BHs enter the low-accretion regime providing a 'radio' mode of AGN feedback. The core of the distribution which is located at $\sim 10^{-1}$–$1\dot{M}_{\rm Edd}$ at $z = 4$ is at $\sim 10^{-2}\dot{M}_{\rm Edd}$ at $z = 2$, $\sim 10^{-3}\dot{M}_{\rm Edd}$ at $z = 1$, and $\sim 10^{-4}\dot{M}_{\rm Edd}$ at $z = 0$, with very fewer and fewer Eddington-limited BHs at lower redshifts. We observe a lower limit trend that goes like $M_{\rm BH}^2$ produced by a combination of the minimum density and maximum temperature reached in the CGM very close to the galaxy hosting the BH. This lower bound evolves with time because of both a rarefaction of the gas and an increase in temperature as haloes get more shock-heated as they grow in mass (Birnboim & Dekel 2003; Dekel & Birnboim 2006; Birnboim, Dekel & Neistein 2007).

These diagrams suggest that the feedback from AGNs at high redshifts is essentially dominated by a quasar mode ($\chi > 10^{-2}$), whereas at low redshift a radio mode ($\chi \leq 10^{-2}$) prevails within the core of massive structures (Dubois et al. 2010).

Fig. 17 shows the AGN luminosities $L_{\rm AGN} = \epsilon_{\rm f}\epsilon_{\rm r}\dot{M}_{\rm BH}c^2$ for different redshifts and two different box sizes 25 and 50 h^{-1} Mpc (simulations 256L25JH and 256L50JH). Recall that for the dual quasar–radio AGN mode, $\epsilon_{\rm f}$ depends on the accretion rate to Eddington ratio χ. The bright end of the high redshift ($z = 4$) AGN luminosity function is dominated by the quasar mode because most of the BHs accrete gas at a high Eddington rate (see Fig. 16). At intermediate redshifts, $z = 2$ and 1, the bright end of the AGN luminosity is marginally dominated by the quasar mode, but the transition from the quasar mode to the radio mode appears at larger luminosities. Emitting the largest amounts of energy, massive BHs start to strongly deplete the gas content of their host galaxies and enter a radio-mode AGN regime. At $z = 0$, most of the BHs have

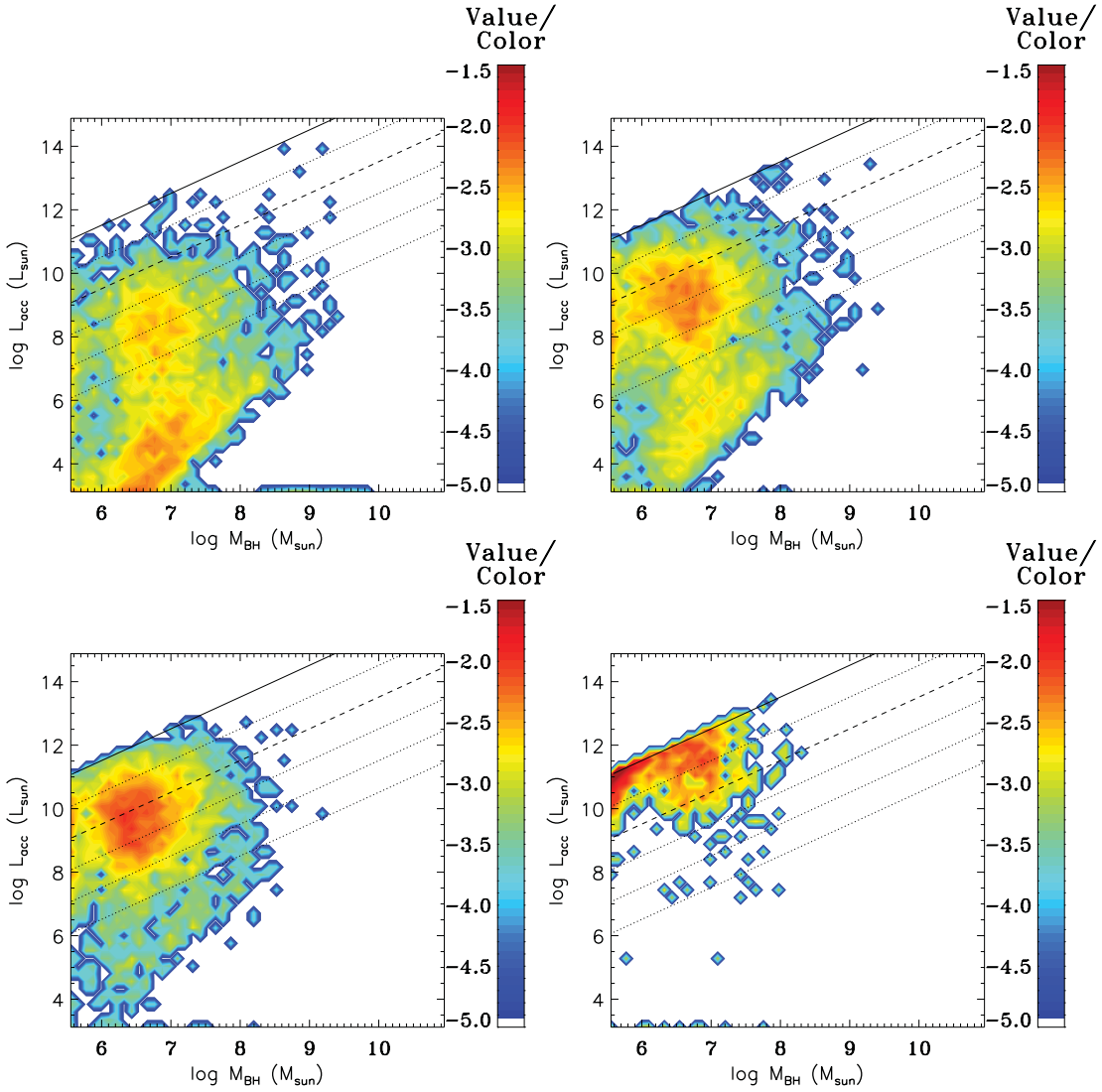

Figure 16. Number-weighted histogram as a function of BH mass and their accretion luminosity for simulation 256L50JH. The solid line corresponds to the Eddington limit. The dashed line separates our fiducial radio and quasar modes. The dotted lines correspond to $10^{-1}\dot{M}_{\rm Edd}$, $10^{-3}\dot{M}_{\rm Edd}$, $10^{-4}\dot{M}_{\rm Edd}$ and $10^{-5}\dot{M}_{\rm Edd}$ from the top to bottom. The upper left-hand plot is for $z = 0$, upper right-hand plot for $z = 1$ bottom left-hand plot for $z = 2$ and bottom right-hand plot for $z = 4$.

reached a very quiescent phase for gas accretion and quasar-mode feedback is almost imperceptible. The remaining AGN quasars are intermediate-mass BHs (see Fig. 16), with most of the very massive BHs in a radio mode.

It is worth noting that as a result of DM mass resolution, the AGN luminosity functions show extrema. However, the slope of the bright end seems relatively independent of the resolution.

6 DISCUSSION AND CONCLUSIONS

We have designed a new self-consistent model of BH growth and AGN feedback for hydrodynamical cosmological simulations with a dual jet–heating mechanism that accounts for the radio mode and the quasar mode. BHs are seeded at an early stage of the formation of galaxies. They grow by successive mergers and by accretion of gas at a Bondi (1952) rate. Some of the rest-mass-accreted energy is converted into energy for the AGN feedback mechanism. At high accretion rates defined by the Eddington ratio $\chi > \chi_{\rm radio}$, BHs enter a quasar mode, with energy liberated as thermal energy. At lower accretion rates $\chi \leq \chi_{\rm radio}$, a radio mode is triggered with injection of mass, momentum and kinetic energy within a bipolar jet.

The parameters of the radio mode for the AGN feedback model are tested to reproduce the observations of the $z = 0$ cosmic BH density and the Magorrian et al. (1998) $M_{\rm BH}$–$M_{\rm s}$ and $M_{\rm BH}$–$\sigma_{\rm s}$ relationships. We find the following behaviour of BH growth upon varying the parameters of the radio AGN feedback mode:

(i) AGN feedback efficiencies lower than $\epsilon_{\rm f} = 1$ lead to larger and unrealistic BH masses that overshoot the observational predictions for the cosmic BH density and the Magorrian relationships. BHs grow to larger masses to inject the same total amount of energy in order to self-regulate their growth.

(ii) Introducing a time-delay in the AGN feedback, as opposed to an instantaneous energy deposit, increases the effectiveness of AGN feedback effect on the gas and the growth of BHs, as a more important energy release occurs over a shorter period of time. However, this increased efficiency tends to exaggerate the effect of AGN

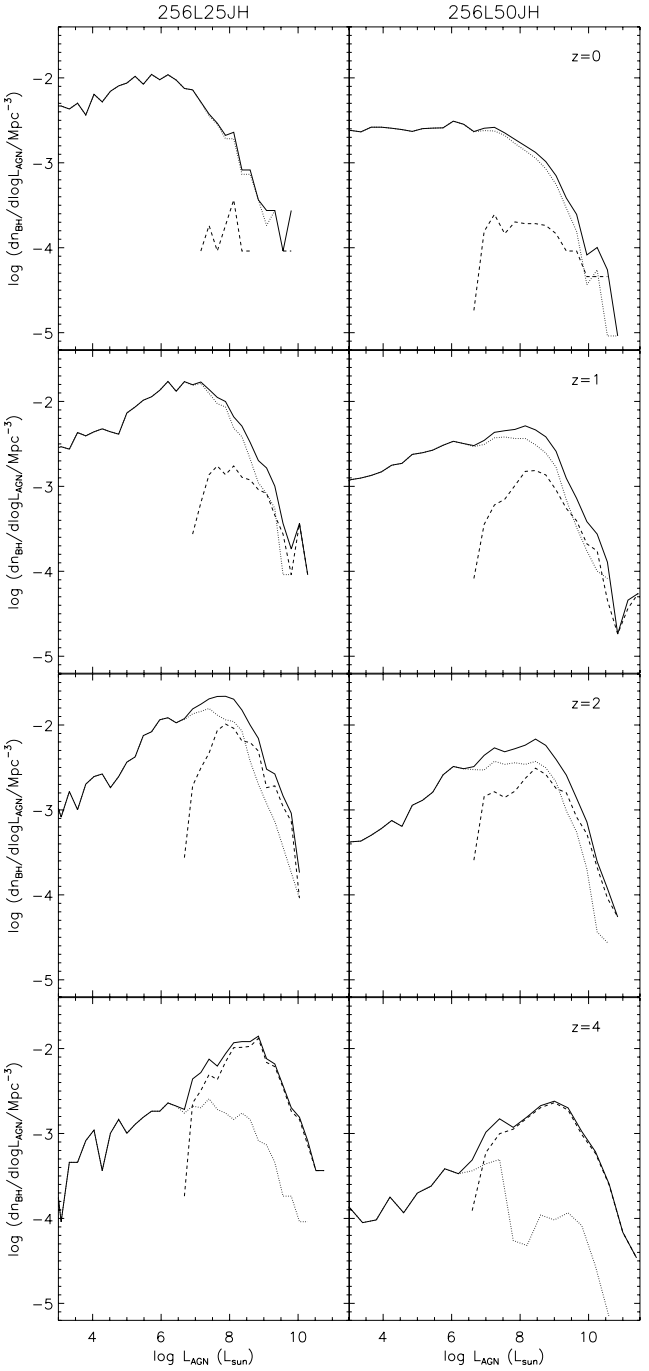

Figure 17. AGN luminosity functions at different redshifts for simulations 256L25JH (left-hand panels) and 256L50JH (right-hand panels). The total luminosity functions (solid lines) are decomposed into the contribution from the radio mode (dotted lines) and from the quasar mode (dashed lines).

feedback, that is, it underestimates the BH density and overestimates the $M_{\rm BH}/M_{\rm s}$ ratios.

(iii) The mass loading factor of the jet, that is, the velocity of the jet, has a negligible impact on the results.

(iv) The choice of the seed BH mass is relatively unimportant except at high redshift when BH masses are comparable to the total stellar mass of galaxies. However, the choice of the seed mass is quickly forgotten as BHs self-regulate their growth.

(v) The size of the jet for the radio mode, as well as the size of bubbles for the quasar mode, must be chosen carefully. Large jets and bubbles deposit energy far away from the galaxy and have a harder time impacting the dense gas, and, thus, self-regulating the growth of BHs. We find that the size of the energy injection region must be as close as possible to the minimum physical resolution of the code.

(vi) We find that the radio mode alone requires a larger energy efficiency $\epsilon_{\rm f,r} = 1$ than the quasar mode alone $\epsilon_{\rm f,q} = 0.15$ to match the data from observations. The thermal mode couples more efficiently to the gas than the kinetic mode, and has more impact on the baryon content of galaxies.

We also tested the convergence and the robustness of our model to the effect of finite numerical resolution using different box sizes, DM particle masses and minimum cell sizes. We obtain a satisfying convergence of the cosmic BH density, and $M_{\rm BH}$–$M_{\rm s}$ and $M_{\rm BH}$–$\sigma_{\rm s}$ relationships at $z = 0$ as long as the DM mass resolution is $M_{\rm DM} \leq 4.4 \times 10^8\,{\rm M}_\odot$ and the minimum physical cell size is $\Delta x \leq 1.52\,h^{-1}$ kpc. However, we have seen that the convergence of the cosmic SFR at high redshift is not reached because it is dominated by small haloes.

These simulations have demonstrated the ability of AGN feedback to efficiently suppress the amount of stars and cold gas. The removal of cold baryons begins at high redshift as soon as the first progenitors of massive galaxies collapse. At low redshift, the effect of AGN feedback is amplified by the presence of more massive objects for which the impact of AGN feedback is stronger. The presence of AGN feedback is a necessary ingredient of galaxy mass budgets as it leads to a better fit to observed SFRs that slowly converge at low redshift.

We have shown that quasars are increasingly important at high redshift. Since cold gas is more abundant in galaxies at high redshift, BHs accrete gas at higher rates, and, as a consequence, the quasar mode of AGN feedback is more often triggered. As the gas is consumed by star formation, and the accretion of new fresh gas is quenched by both the shock-heating of massive structures and the feedback from SNe and AGNs, the accretion on to BHs proceeds at lower rates. Thus, as time goes by, the radio mode of AGN feedback becomes more and more dominant, and triggers jets in massive structures such as groups and clusters of galaxies. Radio-active galaxies comprise only a fraction of observed AGNs (Best et al. 2005; Smolčić et al. 2009, 2011). As the amount of radio emission varies a lot from object to object and depends on the gas accretion rate/mode on to the BH, working out the reasons of quantitative (dis)agreement between our predictions and observations in detail, taking into account limitations on both sides, is a complex issue, beyond the scope of this paper. We therefore defer such an analysis to future work.

Our results are in good agreement with previous cosmological numerical simulations of the self-regulated growth of BHs. The quasar mode employed here is extremely similar to that modelled with a different numerical technique by Booth & Schaye (2009), inspired by the model of Sijacki et al. (2007). Our conclusions in terms of $M_{\rm BH}$–$M_{\rm s}$ relationships, and BH density evolution, are, indeed, extremely comparable. This is an important conclusion of this paper: cosmological codes that treat gravity and gas dynamics with different approaches produce extremely similar results for the co-evolution of BHs and galaxies with an identical set of subgrid physics for galaxy formation. We have implemented a radio mode for AGN feedback and found a set of parameters that give results in good agreement with both observations and results produced by

the well-tested quasar mode. It proves that a pure kinetic mode is also capable of reducing the amount of baryons in galaxies and self-regulate the growth of BHs. A simple combination of both modes, the dual radio–quasar mode, reproduces the observations as well as any one of the two single modes, and provides a more realistic approach to the treatment of AGN feedback.

ACKNOWLEDGMENTS

We thank J. Magorrian, L. Miller and A. Babul for useful discussions. YD is supported by an STFC Postdoctoral Fellowship. The simulations presented here were run on the TITANE cluster at the Centre de Calcul Recherche et Technologie of CEA Saclay on allocated resources from the GENCI grant c2009046197, and on the DiRAC facility jointly funded by STFC, the Large Facilities Capital Fund of BIS and the University of Oxford. This research is part of the Horizon-UK project. The research work of JD and AS is supported by Adrian Beecroft, the Oxford Martin School and STFC.

REFERENCES

Agertz O. et al., 2007, MNRAS, 380, 963
Allen S. W., Dunn R. J. H., Fabian A. C., Taylor G. B., Reynolds C. S., 2006, MNRAS, 372, 21
Barai P., 2008, ApJ, 682, L17
Barnes J. E., Hernquist L., 1992, ARA&A, 30, 705
Bate M. R., Bonnell I. A., Price N. M., 1995, MNRAS, 277, 362
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bellovary J. M., Governato F., Quinn T. R., Wadsley J., Shen S., Volonteri M., 2010, ApJ, 721, L148
Bennert V. N., Treu T., Woo J., Malkan M. A., Le Bris A., Auger M. W., Gallagher S., Blandford R. D., 2010, ApJ, 708, 1507
Benson A. J., Babul A., 2009, MNRAS, 397, 1302
Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Birnboim Y., Dekel A., Neistein E., 2007, MNRAS, 380, 339
Bîrzan L., Rafferty D. A., McNamara B. R., Wise M. W., Nulsen P. E. J., 2004, ApJ, 607, 800
Blandford R. D., Begelman M. C., 1999, MNRAS, 303, L1
Boehringer H., Voges W., Fabian A. C., Edge A. C., Neumann D. M., 1993, MNRAS, 264, L25
Bondi H., 1952, MNRAS, 112, 195
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Booth C. M., Schaye J., 2010, MNRAS, 405, L1
Booth C. M., Schaye J., 2011, MNRAS, 413, 1158
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
Bromm V., Loeb A., 2003, ApJ, 596, 34
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Cattaneo A., Teyssier R., 2007, MNRAS, 376, 1547
Cattaneo A., Haehnelt M. G., Rees M. J., 1999, MNRAS, 308, 77
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Chartas G., Brandt W. N., Gallagher S. C., 2003, ApJ, 595, 85
Chiaberge M., Capetti A., Macchetto F. D., 2005, ApJ, 625, 716
Churazov E., Sazonov S., Sunyaev R., Forman W., Jones C., Böhringer H., 2005, MNRAS, 363, L91
Crenshaw D. M., Kraemer S. B., George I. M., 2003, ARA&A, 41, 117
Croton D. J. et al., 2006, MNRAS, 365, 11
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Davis O., Devriendt J., Colombi S., Silk J., Pichon C., 2011, MNRAS, 413, 2087
De Villiers J., Hawley J. F., Krolik J. H., Hirose S., 2005, ApJ, 620, 878
Debuhr J., Quataert E., Ma C., Hopkins P., 2010, MNRAS, 406, L55
Decarli R., Falomo R., Treves A., Labita M., Kotilainen J. K., Scarpa R., 2010, MNRAS, 402, 2453
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A. et al., 2009, Nat, 457, 451
Di Matteo T., Springel V., Hernquist L., 2005, Nat, 433, 604
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33
Dib S., Bell E., Burkert A., 2006, ApJ, 638, 797
Dong R., Rasmussen J., Mulchaey J. S., 2010, ApJ, 712, 883
Dubois Y., Teyssier R., 2008, A&A, 477, 79
Dubois Y., Devriendt J., Slyz A., Silk J., 2009, MNRAS, 399, L49
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, MNRAS, 409, 985
Dubois Y., Devriendt J., Teyssier R., Slyz A., 2011, MNRAS, 417, 1853
Dunkley J. et al., 2009, ApJS, 180, 306
Dunn R. J. H., Allen S. W., Taylor G. B., Shurkin K. F., Gentile G., Fabian A. C., Reynolds C. S., 2010, MNRAS, 404, 180
Fabian A. C., Sanders J. S., Taylor G. B., Allen S. W., Crawford C. S., Johnstone R. M., Iwasawa K., 2006, MNRAS, 366, 417
Fan X. et al., 2003, AJ, 125, 1649
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Gebhardt K. et al., 2000, ApJ, 539, L13
Graham A. W., Driver S. P., 2007, ApJ, 655, 77
Granato G. L., Silva L., Monaco P., Panuzzo P., Salucci P., De Zotti G., Danese L., 2001, MNRAS, 324, 757
Gültekin K. et al., 2009, ApJ, 698, 198
Haardt F., Madau P., 1996, ApJ, 461, 20
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hawley J. F., Krolik J. H., 2006, ApJ, 641, 103
Heger A., Woosley S. E., 2002, ApJ, 567, 532
Heitmann K. et al., 2008, Comput. Sci. Discovery, 1, 015003
Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Krause E., 2007, ApJ, 669, 67
Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Khalatyan A., Cattaneo A., Schramm M., Gottlöber S., Steinmetz M., Wisotzki L., 2008, MNRAS, 387, 13
King A., 2003, ApJ, 596, L27
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Krumholz M. R., Tan J. C., 2007, ApJ, 654, 304
Krumholz M. R., McKee C. F., Klein R. I., 2004, ApJ, 611, 399
Laor A., 2001, ApJ, 553, 677
Loeb A., Rasio F. A., 1994, ApJ, 432, 52
Maccarone T. J., 2003, A&A, 409, 697
McCarthy I. G. et al., 2010, MNRAS, 406, 822
McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, MNRAS, 412, 1965
McKinney J. C., 2006, MNRAS, 368, 1561
McKinney J. C., Blandford R. D., 2009, MNRAS, 394, L126
McLure R. J., Dunlop J. S., 2002, MNRAS, 331, 795
McLure R. J., Jarvis M. J., Targett T. A., Dunlop J. S., Best P. N., 2006, MNRAS, 368, 1395
McNamara B. R., Nulsen P. E. J., Wise M. W., Rafferty D. A., Carilli C., Sarazin C. L., Blanton E. L., 2005, Nat, 433, 45
Madau P., Rees M. J., 2001, ApJ, 551, L27
Magorrian J. et al., 1998, AJ, 115, 2285
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Merloni A., Heinz S., 2008, MNRAS, 388, 1011
Merloni A. et al., 2010, ApJ, 708, 137
Mitchell N. L., McCarthy I. G., Bower R. G., Theuns T., Crain R. A., 2009, MNRAS, 395, 180
Morgan C. W., Kochanek C. S., Morgan N. D., Falco E. E., 2010, ApJ, 712, 1129
Nagar N. M., Falcke H., Wilson A. S., 2005, A&A, 435, 521
Narayan R., Yi I., 1994, ApJ, 428, L13

Navarro J. F., White S. D. M., 1993, MNRAS, 265, 271
Nayakshin S., Power C., 2010, MNRAS, 402, 789
O'Shea B. W., Nagamine K., Springel V., Hernquist L., Norman M. L., 2005, ApJS, 160, 1
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Omma H., Binney J., Bryan G., Slyz A., 2004, MNRAS, 348, 1105
Owen F. N., Eilek J. A., Kassim N. E., 2000, ApJ, 543, 611
Peng C. Y., Impey C. D., Rix H., Kochanek C. S., Keeton C. R., Falco E. E., Lehár J., McLeod B. A., 2006, ApJ, 649, 616
Pope E. C. D., 2011, MNRAS, 414, 3344
Pounds K. A., Reeves J. N., King A. R., Page K. L., O'Brien P. T., Turner M. J. L., 2003, MNRAS, 345, 705
Powell L. C., Slyz A., Devriendt J., 2011, MNRAS, 414, 3671
Puchwein E., Sijacki D., Springel V., 2008, ApJ, 687, L53
Rasera Y., Teyssier R., 2006, A&A, 445, 1
Romeo A. B., Agertz O., Moore B., Stadel J., 2008, ApJ, 686, 1
Salviander S., Shields G. A., Gebhardt K., Bonning E. W., 2007, ApJ, 662, 131
Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30
Schödel R. et al., 2002, Nat, 419, 694
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, MNRAS, 354, 1020
Shields G. A., Menezes K. L., Massart C. A., Vanden Bout P., 2006, ApJ, 641, 683
Sijacki D., Springel V., di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Sijacki D., Pfrommer C., Springel V., Enßlin T. A., 2008, MNRAS, 387, 1403
Silk J., Rees M. J., 1998, A&A, 331, L1
Smolčić V. et al., 2009, ApJ, 696, 24
Smolčić V., Finoguenov A., Zamorani G., Schinnerer E., Tanaka M., Giodini S., Scoville N., 2011, MNRAS, 416, L31
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Taylor G. B., Sanders J. S., Fabian A. C., Allen S. W., 2006, MNRAS, 365, 705
Teyssier R., 2002, A&A, 385, 337
Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195
Tremaine S. et al., 2002, ApJ, 574, 740
Truelove J. K., Klein R. I., McKee C. F., Holliman J. H., II, Howell L. H., Greenough J. A., 1997, ApJ, 489, L179
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559
Wyithe J. S. B., Loeb A., 2003, ApJ, 595, 614

APPENDIX A: BULGE AND DISC DECOMPOSITION

We decompose the bulge and the disc in our simulated galaxies with a similar procedure to what is used for decomposing observed stellar luminosity profiles. The difference is that we do this decomposition directly on the stellar density on not on the luminosity. Our major assumption is that the relation between stellar surface luminosity and stellar surface density is linear. The procedure is as follows:

(i) Galaxies with a minimum of 100 star particles are identified with a galaxy finder based on the Most massive Sub-node Method described in Tweed et al. (2009), which allows for a clean separation of structures and substructures. This is particularly important for galaxy mergers, or for massive galaxies with extended stellar haloes and galaxy satellites.

(ii) A stellar surface density profile is computed from the stars that belong to a given galaxy as defined by the galaxy finder, with the galaxy seen face-on. The face-on view of the galaxy is defined by the line of sight along the angular momentum axis defined by the rotation of the stars.

(iii) A double decreasing exponential of the form $\propto \Sigma_i \exp r/r_i$ (the 'i' subscript is for the bulge Σ_b, r_b, or for the disc Σ_d, r_d) is fitted to the stellar surface density profile using a least χ^2 to separate the bulge from the disc component.

(iv) The bulge mass is taken to be $M_\mathrm{b} = 2\pi r_\mathrm{b}^2 \Sigma_\mathrm{b}$; the disc mass is taken to be $M_\mathrm{d} = 2\pi r_\mathrm{d}^2 \Sigma_\mathrm{d}$.

(v) If only one component fits the stellar density profile, we directly use the stellar mass given by the galaxy finder.

We point out that, usually, cruder strategies are employed to separate the so-called 'bulge' component from a galaxy's total stellar mass, by assuming that the half-mass equals the bulge mass (Sijacki et al. 2007), or using the total halo stellar mass as a proxy for the bulge stellar mass (Di Matteo et al. 2008; Booth & Schaye 2009). Even though our method is more consistent with the observational definition of a bulge, our tests suggest that at these kpc resolutions, other definitions lead to similar conclusions.

Fig. A1 shows the stellar surface density profile obtained for the most massive galaxy in simulation 256L25JH. We can clearly observe an inner bulge component with size $r_\mathrm{b} = 8.12$ kpc, and a disc component with larger characteristic radius $r_\mathrm{b} = 21.85$ kpc. This galaxy has a bulge mass $M_\mathrm{b} = 3.71 \times 10^{11}\,\mathrm{M}_\odot$ and a disc mass $M_\mathrm{d} = 4.32 \times 10^{11}\,\mathrm{M}_\odot$. The sum of the two fits with the exponential forms is a good approximation of the stellar density profile and gives a total stellar mass $M_\mathrm{s} = 8.03 \times 10^{11}\,\mathrm{M}_\odot$ similar to the total stellar mass obtained with the galaxy finder $M_\mathrm{gal} = 8.34 \times 10^{11}\,\mathrm{M}_\odot$. We must stress that the disc component for such massive galaxies is not relevant, because this outer component of a galaxy corresponds to a large stellar halo rather than a rotating disc of stars. However, in this paper, we do not discuss the properties of stellar discs or stellar haloes of galaxies. Our main focus is to separate the bulge component from the total distribution of stars.

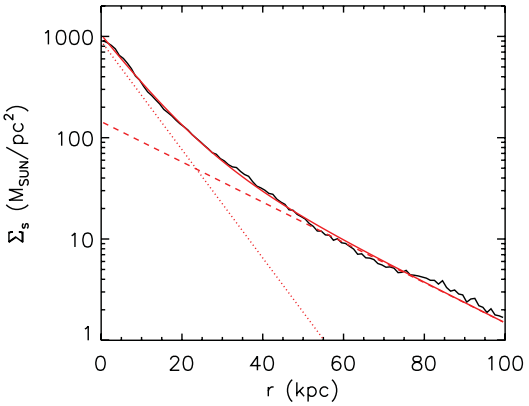

Figure A1. Stellar surface density (black solid line) for the most massive galaxy at $z = 0$ in simulation 256L25JH with the bulge (red dotted line) and disc (red dashed line) decomposition from the best-fitting model (red solid line).

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