Cosmological simulations of black hole growth: AGN luminosities and downsizing

Michaela Hirschmann, Klaus Dolag, Alexandro Saro, Lisa Bachmann, Stefano Borgani and Andreas Burkert

ABSTRACT

In this study, we present a detailed, statistical analysis of black hole growth and the evolution of active galactic nuclei (AGN) using cosmological hydrodynamic simulations run down to redshift $z = 0$. The simulations self-consistently follow radiative cooling, star formation, metal enrichment, black hole growth and associated feedback processes from both Type II/Ia supernovae and AGN. We consider two simulation runs, one with a large comoving volume of $(500 \, \text{Mpc})^3$ and one with a smaller volume of $(68 \, \text{Mpc})^3$ but with a factor of almost 20 higher mass resolution. We compare the predicted statistical properties of AGN with results from large observational surveys. Consistently with previous results, our simulations can widely match observed black hole properties of the local Universe. Furthermore, our simulations can successfully reproduce the evolution of the bolometric AGN luminosity function for both the low-luminosity and the high-luminosity end up to $z = 3.0$, only at $z = 1.5–2.5$, the low-luminosity end is overestimated by up to 1 dex. In addition, the smaller but higher resolution run is able to match the observational data of the low bolometric luminosity end at higher redshifts $z = 3–4$. We also perform a direct comparison with the observed soft and hard X-ray luminosity functions of AGN, including an empirical correction for a torus-level obscuration, and find a similarly good agreement. These results nicely demonstrate that the observed ‘antihierarchical’ trend in the AGN number density evolution (i.e. the number densities of luminous AGN peak at higher redshifts than those of faint AGN) is self-consistently predicted by our simulations. Implications of this downsizing behaviour on active black holes, their masses and Eddington ratios are discussed. Overall, the downsizing behaviour in the AGN number density as a function of redshift can be mainly attributed to the evolution of the gas density in the resolved vicinity of a (massive) black hole (which is depleted with evolving time as a consequence of star formation and AGN feedback).

Key words: methods: numerical – galaxies: active – galaxies: evolution – quasars: general – quasars: supermassive black holes – galaxies: statistics.

1 INTRODUCTION

It is generally accepted that present-day spheroidal galaxies host supermassive black holes (SMBHs) at their centres (Magorrian et al. 1998; Genzel & Eckart 1999). In addition, strong correlations are found between black hole (BH) masses and properties of their host galaxies (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Haring & Rix 2004; Graham & Driver 2007; Gültekin et al. 2009; Burkert & Tremaine 2010; McConnell & Ma 2013) as the bulge mass, the stellar velocity dispersion, the Sérsic index or the number of globular clusters. This is often interpreted as an evidence for a co-evolution between the spheroidal component of host galaxies and their BHs (see Kormendy & Ho 2013 for a recent review). However, these relations were also shown to possibly have a ‘non-causal’ origin and thus, may be just a consequence of...
statistical merging processes (Peng 2007; Hirschmann et al. 2010; Jahnke & Macciò 2011). In addition, some recent observations of BH growth and star formation (SF) in individual objects appear to contradict the picture of a simple one-to-one co-evolution over time (e.g. Bongiorno 2012; Mullaney et al. 2012; Rosario 2012; Hickox et al. 2014). Thus, the details of a connected growth of BHs and their host galaxies remain an unresolved puzzle.

During their lifetime, BHs are assumed to undergo several episodes of significant gas accretion, during which this accretion powers luminous quasars or active galactic nuclei (AGN; Salpeter 1964; Zel’Dovich 1964; Lynden-Bell 1969). By estimating the total energy radiated by AGN over their whole lifetime, it can be shown that nearly all the mass seen in dormant BHs today can be accumulated during the periods of observed bright AGN activity (Soltan 1982). This implies that there is not much room left for ‘dark’ or obscured accretion.

Recent progress in detecting AGN (particularly faint and obscured ones) was achieved by analysing data from X-ray surveys in the hard and soft X-ray range (XMM–Newton, Chandra, ROSAT, ASCA, eROSITA; e.g. Miyaji, Hasinger & Schmidt 2000; La Franca et al. 2002; Barger et al. 2003; Cowie et al. 2003; Fiore et al. 2003, 2012; Ueda et al. 2003; Sazonov & Revnivtsev 2004; Barger & Cowie 2005; Hasinger, Miyaji & Schmidt 2005; Nandra, Laird & Steidel 2005; Ebrero et al. 2009; Aird et al. 2010; Kolodzig et al. 2013). Interestingly, these studies revealed possibly at first sight a puzzling behaviour of AGN concerning the evolution of the comoving number density of AGN: the number density of successively more luminous AGN peaks at higher redshifts than the one of less luminous AGN, with the lowest luminosity AGN showing an almost constant comoving number density. Making the simplified (and naive) assumption that AGN luminosity (i.e. BH accretion) is proportional to BH mass (which we would expect if BHs are accreting at the Eddington rate, $L \propto M^2_B$) would imply that very massive BHs seem to be already in place at very early times, whereas less massive BHs seem to evolve predominantly at lower redshifts. This behaviour is called ‘downsizing’ or ‘antihierarchical’ growth of BHs. The downsizing trend is also seen in the optical (Wolf et al. 2003; Cristiani et al. 2004; Croom et al. 2004; Fan et al. 2004; Hunt et al. 2004; Richards et al. 2006) and the near-infrared (NIR; e.g. Matute et al. 2006).

This observational result seems to be in conflict with the ‘naive’ expectations arising from the currently favoured hierarchical structure formation paradigm based on the cold dark matter model (Peebles 1965; White & Rees 1978; Blumenthal et al. 1985). In this framework, low-mass haloes are expected to form first and more massive haloes to grow later over time via subsequent merging and smooth accretion. However, it is now well known that the evolutionary history of observable galaxies also follows an antihierarchical or downsizing behaviour, with several independent observational indicators suggesting that more massive galaxies formed their stars and had their SF quenched earlier than low-mass galaxies, which continue forming stars to the present day (see Fontanot et al. 2009 for an overview).

To theoretically explore the formation and evolution of BHs and AGN and its (possible) connection to the evolution of their host galaxies, a large amount of studies were published based on semi-analytic models of galaxy formation within which mechanisms for BH formation and evolution were included (e.g. Kauffmann & Haehnelt 2000; Volonteri, Haardt & Madau 2003; Granato et al. 2004; Menci et al. 2004; Bromley, Somerville & Fabian 2004; Monaco & Fontanot 2005; Bower et al. 2006; Croton 2006; Marulli et al. 2008; Somerville & Fabian 2008; Bonoli et al. 2009; Fanidakis et al. 2012; Hirschmann et al. 2012; Menci, Fiore & Lamastra 2013; Neistein & Netzer 2014). It was shown by some of the above studies that an antihierarchical behaviour can be explained within a hierarchical structure formation scenario (Marulli et al. 2008; Bonoli et al. 2009; Fanidakis et al. 2012; Hirschmann et al. 2012; Menci et al. 2013), when considering different models for BH accretion and AGN triggering and/or accounting for dust obscuration. Unfortunately, these different models have not been able to draw a completely uniform picture for the origin of the downsizing behaviour as they differ in many details.

In addition to these semi-analytic models, also a large number of other studies investigated the predictions of the cosmological lambda cold dark matter ($\Lambda$CDM) model for the formation and evolution of SMBHs and AGN. Some works made predictions based on nearly purely analytic models (Efstathiou & Rees 1988; Haehnelt & Rees 1993; Haiman & Loeb 1998) and on semi-empirical models (Shankar, Weinberg & Miralda-Escudé 2009; Shankar, Weinberg & Shen 2010; Shankar et al. 2012).

Several years ago, numerical hydrodynamic simulations of isolated galaxy mergers started to include prescriptions for BH growth and AGN feedback using ‘subgrid’ recipes, where BH accretion according to the Bondi–Hoyle–Littleton formula was assumed (Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005; Hopkins et al. 2006; Robertson et al. 2006; Li et al. 2007; Sijacki et al. 2007; Johansson, Naab & Burkert 2009; Fabjan et al. 2010; Choi et al. 2013; Barai et al. 2014, and references therein). Feedback from AGN, however, was implemented in different ways, ranging from a purely thermal energy injection (e.g. Springel et al. 2005) and a further distinction between radio and quasar mode (by injecting bubbles; e.g. Sijacki et al. 2007, or just up-scaling of the feedback efficiencies; e.g. Fabjan et al. 2010) to mechanical momentum-driven winds (e.g. Choi et al. 2013).

To assess statistical properties of BHs and AGN in a full cosmological context and in a more self-consistent way than it can be achieved by e.g. semi-analytical or semi-empirical models, large cosmological, hydrodynamic simulations were performed by several recent studies (Di Matteo et al. 2008, 2012; McCarthy et al. 2010, 2011; Degraf, Di Matteo & Springel 2010, 2011; Booth & Schaye 2011; DeGraf et al. 2012; Haas et al. 2013; Rosas-Guevara et al. 2013; Khandel et al. 2014) including BH growth and feedback from AGN. Compared to semi-analytical models, they have the advantage that the dynamics of the baryonic component (gas and stars) and the interaction between baryonic matter can be followed directly in a cosmological context, even if the spatial and mass resolution, at present, is not high enough to accurately simulate small-scale physical processes like e.g. the formation of stars and BHs with the associated feedback. Instead, they have to be computed in a simplified manner with subgrid/subresolution models. Nevertheless, they provide a powerful tool to assess the cosmic evolution of statistical AGN and BH properties and thus, go clearly beyond the hydrodynamically isolated merger simulations and/or semi-analytic models.

Di Matteo et al. (2008) and Booth & Schaye (2011), for example, investigated the global history of BH mass assembly and the evolution of the BH–bulge mass relations in cosmological simulations with a maximum box size of $\sim (50\text{Mpc})^3$. Such simulations have also been successful in reproducing the low-luminosity end of the AGN luminosity function (Degraf et al. 2010) or BH clustering properties (Degraf et al. 2011) but have been limited by a box size being not large enough to follow the evolution of luminous AGN with $L_{\text{tot}} \geq 10^{45} \text{ erg s}^{-1}$ (i.e. AGN luminosities above the
exponential cut-off of the luminosity function). In a recent study by Khandai et al. (2014) they have used a larger volume of (100 Mpc $h^{-1}$)$^3$ (where the luminous end is also accessible), but they typically underestimated luminous AGN at $z > 0.5$ and overestimated them at $z < 0.5$. Focusing only on high redshifts, Di Matteo et al. (2012) and DeGraf et al. (2012) investigated the growth of the first quasars appearing in the Universe by performing cosmological simulations with a large box size of (500 Mpc $h^{-1}$)$^3$, however, only run down to redshift $z = 5$ finding a rapid, early growth of BHs driven by cold, infalling gas flows.

In this work, we extend their studies of the cosmic evolution of statistical properties of BH growth and AGN evolution at lower redshifts, from $z = 5$ down to $z = 0$. We analyse a subset of cosmological, hydrodynamical simulations from the Magneticum Pathfinder simulation set (Dolag et al., in preparation), which includes a very large cosmological box providing a volume of (500 Mpc)$^3$. The simulations include a self-consistent model for BH growth using the GADGET3 code containing several modifications for modelling the growth of BHs with respect to previous studies (e.g. Di Matteo et al. 2008, 2012; Khandai et al. 2014).

In the course of this study, we will particularly focus on the evolution of the AGN luminosity function and the connected anti-hierarchical/downsizing trend in AGN number density evolution with the aim to understand its origin within a hierarchical structure formation scenario. In this context, we test the effect of empirically motivated dust obscuration models and discuss implications of the downsizing trend on the interplay between AGN luminosities, BH mass and Eddington ratios. Overall, in this study, we go beyond previous papers of e.g. Di Matteo et al. (2008), DeGraf et al. (2010) or Khandai et al. (2014) as we examine the evolution of the AGN luminosity function in a factor of ~45 larger cosmological volume with comparably good (even if slightly lower than in Khandai et al. 2014) resolution allowing us to probe particularly luminous AGN with $L_{bol} \geq 10^{45}$ erg s$^{-1}$. In addition, the simulation code is improved in various details, regarding the smoothed particle hydrodynamics (SPH) implementation used as well as the richness of physical processes included, including details of the handling of the BH sink particles. Here the most noticeable difference is reflected in our ability to follow BHs in galaxies which are inside of galaxy clusters.

This study is organized as follows. Section 2 gives an overview of the simulation code and the corresponding model for BH growth and AGN feedback adopted in our simulations. In Section 3, we present some basic properties of present-day and high-redshift BHs and galaxies, which are compared to observational data. We study the evolution of the AGN luminosity function in Section 4 at ($z = 0-5$), considering bolometric, soft and hard X-ray and (for $z = 0$ only) also radio luminosities, and compare our results to observational data. In Section 5, we explicitly show the evolution of the AGN number density and discuss the anti-hierarchical trend in BH growth, its consequences on the connection between BH masses and AGN luminosities and its numerical origin in our simulations. In the end, in Section 6, we summarize and discuss our main results.

### Table 1. Overview of the two simulation runs which are analysed in this study.

| Name       | Box size (Mpc $h^{-1}$) | Resolution level | Initial number particle | $m(\text{dm})$ (M$_\odot$ $h^{-1}$) | $m(\text{gas})$ (M$_\odot$ $h^{-1}$) | $m(\text{stars})$ (M$_\odot$ $h^{-1}$) | Softening length (dm, gas, stars) (kpc $h^{-1}$) |
|------------|-------------------------|------------------|--------------------------|-----------------------------------|---------------------------------------|----------------------------------------|------------------------------------------|
| 500 Mpc/hr | 352                     | hr               | $2 \times 1,564^3$      | $6.9 \times 10^8$                | $1.4 \times 10^9$                    | $3.5 \times 10^7$                      | 3.75, 3.75, 2.0                         |
| 68 Mpc/ahr | 48                      | uhr              | $2 \times 576^3$        | $3.6 \times 10^7$                | $7.3 \times 10^6$                    | $1.8 \times 10^6$                      | 1.4, 1.4, 0.7                           |

### 2 THE SIMULATIONS

In this paper, we analyse a subset of cosmological boxes from the Magneticum Pathfinder simulation set (Dolag et al., in preparation). From this set, we selected a large size cosmological simulation with a comoving box size of (500 Mpc)$^3$, simulated with an initial particle number of $2 \times 1564^3$ (500 Mpc/hr, which is our large ‘high-resolution’ run throughout this study) and one cosmological simulation with a smaller comoving box size of (68 Mpc)$^3$ (only run down to $z = 1$), with a particle number of $2 \times 576^3$ and thus, an increased (‘ultrahigh’) resolution (68 Mpc/ahr). Table 1 gives an overview of the two simulation runs. We have adopted a $\Lambda$CDM model with parameters chosen according to the seven-year Wilkinson Microwave Anisotropy Probe (WMAP) data (Komatsu et al. 2011) with $\Omega_m = 0.272$, $\Omega_\Lambda = 0.259$, $\Omega_\Lambda = 0.728$ and $h = 0.704$. The initial power spectrum follows $n = 0.963$ and is normalized to $\sigma_8 = 0.809$.

#### 2.1 The numerical code

Our simulations are based on the parallel cosmological TreePM-SPH code $\texttt{p-GADGET3}$ (Springel 2005). The code uses an entropy-conserving formulation of SPH (Springel & Hernquist 2002). We are using a higher order kernel based on the bias-corrected, sixth-order Wendland kernel (Dehnen & Aly 2012) with 295 neighbours which together with the usage of a low-viscosity SPH scheme allows us to properly track turbulence within galaxy clusters (Dolag et al. 2005; Donnert et al. 2013). We also allow for isotropic thermal conduction with 1/20 of the classical Spitzer value (Dolag et al. 2004). The simulation code includes a treatment of radiative cooling, heating from a uniform time-dependent ultraviolet (UV) background and SF with the associated feedback processes. The latter is based on a subresolution model for the multiphase structure of the interstellar medium (ISM; Springel & Hernquist 2003).

Radiative cooling rates are computed by following the same procedure presented by Wiersma, Schaye & Smith (2009). We account for the presence of the cosmic microwave background and of UV/X-ray background radiation from quasars and galaxies, as computed by Haardt & Madau (2001). The contributions to cooling from each one of 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca and Fe) have been pre-computed using the publicly available $\texttt{CLOUDY}$ photoionization code (Ferland et al. 1998) for an optically thin gas in photoionization equilibrium.

In the multiphase model for SF (Springel & Hernquist 2003), the ISM is treated as a two-phase medium where clouds of cold gas form from cooling of hot gas and are embedded in the hot gas phase assuming pressure equilibrium whenever gas particles are above a given threshold density. The hot gas within the multiphase model is heated by supernovae and can evaporate the cold clouds. A certain fraction of massive stars (10 percent) is assumed to explode as Type II supernovae (SNII). The released energy by SNII (10$^{51}$ erg) is modelled to trigger galactic winds with a mass loading rate being proportional to the star formation rate (SFR) to obtain a resulting wind velocity of $v_{wind} = 350$ km s$^{-1}$. 

MNRAS 442, 2304–2324 (2014)
Our simulations also include a detailed model of chemical evolution according to Tornatore et al. (2007). Metals are produced by SNIa, by SNII and by intermediate and low-mass stars in the asymptotic giant branch (AGB). Metals and energy are released by stars of different mass to properly account for mass-dependent lifetimes (with a lifetime function according to Padovani & Matteucci 1993), the metallicity-dependent stellar yields by Woosley & Weaver (1995) for SNII, the yields by van den Hoek & Groenewegen (1997) for AGB stars and the yields by Thielemann et al. (2003) for SNIa. Stars of different mass are initially distributed according to a Chabrier initial mass function (Chabrier 2003).

2.2 The BH growth model

Most importantly, our simulations also include a prescription for BH growth and for a feedback from AGN based on the model presented in Springel, Di Matteo & Hernquist (2005) and Di Matteo et al. (2005) including the same modifications as in the study of Fabjan et al. (2010) and some new, minor changes for BH seeding and BH ‘pinning’ which are explained later in this section.

As for SF, the accretion on to BHs and the associated feedback adopts a subresolution model. BHs are represented by collisionless ‘sink particles’ that can grow in mass by accreting gas from their environments, or by merging with other BHs.

The gas accretion rate $\dot{M}_g$ is estimated by using the Bondi–Hoyle–Lyttleton approximation (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Bondi 1952):

$$\dot{M}_g = \frac{4\pi G^2 M_c \rho}{(c_s^2 + v^2)^{3/2}},$$  
(1)

where $\rho$ and $c_s$ are the density and the sound speed of the surrounding (ISM) gas, respectively, $v$ is the velocity of the BH relative to the surrounding gas and $\alpha$ is a boost factor for the density and the sound speed which typically is set to 100 as in most related works (unless a more detailed description as introduced in Booth & Schaye 2009 is used) and accounts for the fact that in cosmological simulations we cannot account the intracluster medium (ICM) properties within the vicinity of the BH. The BH accretion is always limited to the Eddington rate (maximum possible accretion for balance between inwards directed gravitational force and outwards directed radiation pressure): $\dot{M}_e = \min(\dot{M}_g, \dot{M}_{\text{edd}})$. Note that the detailed accretion flows on to the BHs are unresolved, we can only capture BH growth due to the larger scale gas distribution, which is resolved.

Once the accretion rate is computed for each BH particle the mass continuously grows. To model the loss of this accreted gas from the gas particles, a stochastic criterion is used to select the surrounding gas particles to be accreted. Unlike in Springel et al. (2005), in which a selected gas particle contributes to accretion with all its mass, we include the possibility for a gas particle to accrete only with a slice of its mass, which corresponds to 1/4 of its original mass. This way, each gas particle can contribute with up to four generations of BH accretion events, thus providing a more continuous description of the accretion process.

The total released energy $\dot{E}$ is related to the BH accretion rate by

$$\dot{E} = \dot{M}_e c^2,$$  
(2)

where $\epsilon_i$ is the radiative efficiency, for which we adopt a fixed value of 0.2. Here we are using a slightly larger value than the one standardly assumed (= 0.1) for a radiatively efficient accretion disc on to a non-rapidly spinning BH according to Shakura & Sunyaev 1973, (see also Di Matteo et al. 2005; Springel 2005). This choice is motivated by observations of Davis & Laor (2011) who find higher $\epsilon_i$ for higher BH masses and by our numerical resolution.

We assume that a fraction $\epsilon_i$ of this energy is thermally coupled to the surrounding gas so that

$$\dot{E}_i = \epsilon_i \dot{E} c^2,$$  
(3)

is the rate of the energy feedback. $\epsilon_i$ is a free parameter and typically set to 0.15 (as usually done in simulations which follow the metal depending cooling function, see for example, Booth & Schaye 2011). The energy is distributed kernel weighted to the surrounding gas particles in an SPH like manner.

Additionally, we incorporated the feedback prescription according to Fabjan et al. (2010): we account for a transition from a quasar- to a radio-mode feedback (see also Sijacki et al. 2007) whenever the accretion rate falls below an Eddington ratio of $f_{\text{edd}} := M_i/M_{\text{edd}} < 0.1$. During this radio-mode feedback we assume a four times larger feedback efficiency than in the quasar mode. This way, we want to account for massive BHs, which are radiatively inefficient (having low accretion rates), but which are efficient in heating the ICM by inflating hot bubbles in correspondence of the termination of AGN jets. The total efficiency in the radio mode is very close to the value of 0.1 (=0.15 × 0.2 × 4). This is the canonical value, which Churazov et al. (2005) estimated to be needed to balance cooling by AGN feedback.

Note that we also, in contrast to Springel et al. (2005), modify the mass growth of the BH by taking into account the feedback, e.g. $\Delta M_* = (1 - \epsilon_i) \Delta \dot{M}$. Furthermore, we introduced some additional, technical modifications of the original implementation which we will now summarize.

(I) One difference with respect to the original implementation by Springel et al. (2005) concerns the seeding of BH particles. In the implementation by Springel et al. (2005), BH particles are seeded in a halo whenever it first reaches a minimum (total) friends-of-friends (FoF) halo mass, where the FoF is performed on the dark matter particles only. In order to guarantee that BHs are seeded only in haloes representing clearly resolved galaxies, where sufficient SF took place, our implementation performs a FoF algorithm on star particles, grouping them with a linking length of 0.05 times the mean separation of the DM particles.\footnote{Note that this linking length is thus much smaller than that, 0.15–0.20, originally used, to identify virialized haloes.}

In the ‘hr’ simulation presented here, a total stellar mass of roughly $10^{10} M_\odot h^{-1}$ is needed (corresponding to a couple of hundreds of star particles) for a halo to be seeded with a BH particle (starting with a seed mass of $3.2 \times 10^5 M_\odot h^{-1}$). In the ‘ahr’ simulation we are using slightly smaller values due to the better underlying resolution (BH seed masses of $8 \times 10^7 M_\odot h^{-1}$ in galaxies with a minimum stellar mass of $2.5 \times 10^9 M_\odot h^{-1}$). While the BH then grows very fast until it reaches the stellar–BH–mass relation, this recovers the BH feedback within the galaxies which would have been present if resolution had allowed to seed BHs earlier. This also avoids to imprint any stellar–BH–mass relation from the beginning. Finally, we choose the seeded BHs at the position of the star particle with the largest binding energy within the FoF group, instead of at the dark matter particle with the maximum density, as originally implemented.

(II) In the original implementation by Springel et al. (2005), BHs are forced to remain within the host galaxy by pinning them to the...
position of the particle found having the minimum value of the potential among all the particles lying within the SPH smoothing length computed at the BH position. Within a cosmological context an aside effect of this criterion is that, due to the relatively large values of SPH smoothing lengths, a BH can be removed from the host galaxy whenever it becomes a satellite, and is spuriously merged into the BH hosted by the central halo galaxy. We have relaxed this criterion and do not apply any pinning of the BH particles to the minimum potential within the smoothing length. To avoid that the BH particles are wandering away from the centre of galaxies by numerical effects, we take several measures, in addition to the original implementation of the BH treatment: first, we enforce a more strict momentum conservation within the implementation of gas accretion by momentum conservation for the smooth accretion of the gas and then do not model any momentum transfer when swallowing gas. Additionally, we implemented the conservation of momentum and centre of mass when two BH particles are merging.\(^3\)

(III) Moreover, in contrast to the original implementation, we have explicitly included a dynamical friction force, which is switched on unless the underlying simulation has a high enough resolution so that the cosmological simulations can numerically resolve dynamical friction reasonably well. To estimate the typical friction force induced on a BH particle we are using the following approximation of the Chandrasekhar formula (Chandrasekhar 1943):

\[
F_{\text{df}} = -4\pi \left(\frac{GM_s}{v}\right)^2 \rho \ln(\Lambda) \left(\frac{\sigma}{\sqrt{\pi}}\right)^2 \frac{\psi}{\psi} v
\]

where \(G\) is the gravitational constant and \(M_s\) is the mass of the BH. The local density \(\rho\) in the vicinity of the BH as well as for the relative velocity \(v\) is calculated using only the stellar and the dark matter components around the BH. The Coulomb logarithm is calculated as

\[
\ln(\Lambda) = \ln \left(\frac{R_v}{GM_s}\right)
\]

and \(x = v\sqrt{2}/\sigma\), where we estimate \(\sigma\) as one-third of the maximum circular velocity of the hosting subhalo and for \(R\) (as typical size of the system) we used the half-mass radius of the subhalo, hosting the BH. The parameters of the hosting subhalo for each BH particle are updated every time when SUBFIND (Dolag et al. 2009) is executed on the fly.

This way a BH particle remains within the host galaxy, even if it becomes a satellite of a larger halo and, compared to the original scheme, we are able to track BHs also in satellite galaxies in cluster environments. When the BHs are not placed artificially on the minimum of the potential, of course, there is no guarantee (due to numerical noise, two body scattering or when two BHs are merging) that BH particles are staying always exactly at the local potential minimum. But due to the above handling of the dynamical friction, with evolving time during the simulation, BHs sink towards the minimum potential so that typical displacements from the true potential minimum are smaller than the effective gravitational softening. Therefore, they are orders of magnitude smaller than the typical smoothing radius used for estimating the parameters in the accretion model or used for distributing the feedback energy. Fig. 1 shows explicitly a histogram for the distances of BHs from the potential minimum in the 500 Mpc/hr simulation. All BHs are maximal 2 kpc away from the potential minimum which is reasonable given a softening length of 5.2 kpc in this run.

Fig. 2 shows a visualization (of the gas and stellar mass density) of the 500 Mpc/hr run at different redshift steps (\(z = 3-0\), different panels) focusing on a thin (70 Mpc thick) slice. We have additionally indicated the 20 BHs with highest masses (white circles), the highest Eddington ratios (\(f_{\text{edd}} = L/L_{\text{edd}}\), blue circles) and the highest accretion rates (and thus, luminosities, red circles). The sizes of the circles are scaled logarithmically with the different values and normalized to the maximum value of each quantity.

This visualizes that AGN luminosity does not directly trace the mass of a BH as AGN when selected by their BH mass (white) seem to be a better tracer of the underlying matter distribution, whereas when selected by the Eddington ratio (blue) or by their luminosity (red) they are more located in less dense environments, what already indicates the presence of a ‘downsizing’ trend in their evolution.

In particular, one can see that at \(z = 3\) many most massive BHs also have the highest accretion rates as the accretion rate/luminosity is at that time still largely related with BH mass (even if with some scatter, as discussed later in Fig. 12). With evolving time, this is not the case anymore and massive BHs accrete at a very broad range of Eddington ratios so that at those later times, massive BHs hardly coincide with high Eddington ratios and accretion rates. In addition, at low redshifts, there are several objects with high Eddington ratios located close to voids, these are low-mass BHs (not necessarily satellites), which are growing by smooth gas accretion (as they do not have experienced much feedback).

The inlay shows a zoom on to the region where the most massive cluster forms today. Overplotted here are all galaxies with stellar masses above \(10^{10}\)\(\mathrm{M}_\odot\) (white crosses) and all BHs within this slice of the simulation (white diamonds). Fig. 3 shows a zoom on to the redshift zero slice (left-hand panel) and a zoom on to the central, 2.5 Mpc wide region of the cluster. Here, the stellar populations of the galaxies (white) within our ray-tracing visualization can be nicely seen contrasted to the light blue ICM. Note that not all
galaxies contain a BH as the initial seeding is also related to the ongoing SF in a galaxy. The positions of all BHs within this simulation are marked as black diamonds and nicely reflect the ability of our implementation to keep the BHs at the centre of the satellite galaxies (see also Fig. 1).

Figure 2. Shown is a 500 Mpc wide, 70 Mpc thick slice through the cosmological baryonic mass distribution (stellar and gaseous density) of the 500 Mpc/hr simulation at different redshift steps ($z = 3, 2, 1$ and 0). This is the result of a ray-tracing visualization using SPLOTCH (Dolag et al. 2008). The white, blue and red circles indicate the 20 BHs within this slice which have the highest masses, the highest Eddington ratios and the highest accretion rates, i.e. AGN luminosities, respectively. The sizes of the circles are scaled logarithmically with the different values, normalized to the maximum value of each quantity. A zoom on to a region where the most massive cluster forms is shown, together with all galaxies having a stellar mass larger than $10^{10} M_\odot$ illustrated by white crosses, whereas the white diamonds show all BHs.

To show the effect of resolution we additionally present the results for the 68 Mpc/uhr run down to $z = 1$.

3 FUNDAMENTAL GALAXY AND BH PROPERTIES AND THEIR EVOLUTION

In this section, we will discuss some fundamental properties of the simulated galaxies and BHs in the 500 Mpc/hr run, as the stellar mass function, the BH mass function and the BH–stellar mass relation in the present-day Universe and at higher redshifts up to $z = 4$. Fig. 4 shows the stellar mass function at different redshift steps as indicated in the legend. Simulation results (red solid lines for the 500 Mpc/hr run and green dashed lines for the 68 Mpc/uhr run) are compared with observational data from different studies (black symbols: Drory et al. 2004; Bundy et al. 2005; Fontana et al. 2006; Marchesini et al. 2007; Pérez-González et al. 2008; Ilbert et al. 2010; blue dashed lines; Ilbert 2013). At $z = 4$, the amount of galaxies in the 500 Mpc/hr run is slightly
Figure 3. Shown is the 25 Mpc wide zoom on to the galaxy cluster cluster at $z = 0$, where analogous to Fig. 2 all galaxies with stellar mass larger than $10^{10} M_\odot$ are shown as white crosses, and all BHs are shown as white diamonds. The right-hand panel visualizes a further zoom into the cluster. The region shown is 2.5 Mpc wide and correspond to roughly one-third of the virial size of the cluster. In the ray-tracing visualization, the white colours reflect the stellar component, while the light blue colours correspond to the hot phase of the ICM. The black diamonds mark all the BHs in the simulation.

Figure 4. Evolution of the stellar mass function in the 500 Mpc/hr (red lines) and in the 68 Mpc/uhr run (green dashed lines) compared to observational data (black symbols; Drory et al. 2004; Bundy, Ellis & Conselice 2005; Fontana et al. 2006; Marchesini et al. 2007; Pérez-González et al. 2008; Ilbert et al. 2010 and blue dashed lines; Ilbert 2013). At $z < 1$ the massive end is overestimated in the simulations due to a too inefficiently working radio-mode feedback. The overestimation of low-mass galaxies is a well-known problem and most likely a consequence of our adopted model for stellar winds assuming a constant wind velocity for the ejected gas, which is a resolution effect: for the 68 Mpc/uhr simulation the low-mass end of the stellar mass function at these high redshifts is consistent with the observational data. Instead, down to $z = 1$ the 500 Mpc/hr simulation results provide a reasonably good match with the observational data, while the low-mass end in the 68 Mpc/uhr run is overestimated by up to 1 dex.
(e.g. Oppenheimer & Davé 2006; Davé, Oppenheimer & Finlator 2011). It was repeatedly shown in literature that energy- or momentum-driven winds can significantly reduce the baryon conversion efficiencies (e.g. Hirschmann et al. 2013, and references therein) and thus, also the low-mass end of the stellar mass function resulting in an improved match with the observational data (e.g. Davé et al. 2013; Puchwein & Springel 2013). In these studies, energy-driven winds, for example, are shown to reduce the low-mass end of the stellar mass function by at least 1 dex compared to a constant wind model which can account for the discrepancy between our 68 Mpc/uhr simulation predictions and the observational constraints. In addition, models including ‘early’ stellar feedback (Stinson et al. 2013; Kannan et al. 2014) or radiation pressure (Hopkins et al. 2013) seem to be particularly efficient in delaying SF in low-mass haloes towards later times (i.e. in breaking the hierarchical formation of galaxies) and thus, predicting low baryon conversion efficiencies in these haloes down to \( z = 0 \) – consistent with observational constraints.

Turning towards lower redshifts (\( z < 1 \)), the massive end of the stellar mass function (\( \log (M_{\text{stellar}}/M_\odot) > 11 \)) is significantly overestimated in the 500 Mpc/hr run, at \( z = 0 \) by more than one order of magnitude for \( 10^{12} M_\odot \) mass galaxies. This is most likely a consequence of the thermal energy injection scheme in the ‘radio mode’ adopted in our model (see also Puchwein & Springel 2013 who also overestimate the massive end of the stellar mass function). Here we may speculate that a mechanical-momentum input from an AGN coupling to the ambient gas via a bipolar wind would be more efficient in limiting the inflow and accretion on to the central BH and also SF and thus, could help making elliptical galaxies ‘red and dead’ by suppressing late SF (Choi et al. 2014). Such mechanisms are for example investigated by Choi et al. (2012), Debühr, Quataert & Ma (2012), Barai et al. (2014) and Dubois et al. (2012) employing simulations of isolated galaxies or galaxy mergers.

### 3.2 The BH mass function

In Fig. 5, the present-day BH mass function of the 500 Mpc/hr run (red line) is compared to observations from Marconi et al. (2004), Shankar et al. (2004, 2009) (black lines and symbols and grey shaded areas) and is found to be in reasonably good agreement for a BH mass range of 7.5 < \( \log (M_*/M_\odot) < 9.5 \). Below, the simulation is underpredicting the amount of low massive BHs by almost one order of magnitude. This is mainly related to a combination of too low resolution (main effect, see dashed coloured lines indicating the 68 Mpc/uhr predictions) and BH seed masses.

At the high-mass end, the amount of massive BHs \( \log (M_*/M_\odot) > 9.5 \) is significantly overestimated by up to 2 dex. The high-mass end is mostly influenced by the choice for the parameter regulating the strength for AGN feedback, but hardly dependent on the efficiency parameter for regulating the radio-mode feedback. Although the latter is supposed to regulate the late time SF and BH growth, in the implementation of the thermal injection scheme in the ‘radio mode’ adopted in our model, is still not efficient enough to lower the high-mass end of the BH mass function (as discussed for the stellar mass function). Therefore, AGN feedback at high redshift is much more efficient in suppressing SF also at late times by pre-heating the environment. Instead of a purely thermal energy injection, we speculate that a mechanical-momentum input from an AGN coupling to the ambient gas via a bipolar wind would be more efficient in limiting the inflow and accretion on to the central BH (e.g. Choi et al. 2012; Barai et al. 2014) and thus, reducing the massive end of the BH mass function.

Turning towards higher redshifts (see coloured solid lines in Fig. 5 indicating the 500 Mpc/hr run), the number density at a given BH mass shows a very rapid evolution until \( z = 1 \). At that redshift, the high-mass end is mostly in place and only the amount of BHs below \( \log (M_*/M_\odot) < 8 \) further increases until \( z = 0 \). This implies that BHs grow most strongly until \( z = 1 \) which is also in agreement when considering the cosmic evolution of the total BH accretion rates, which peak around \( z = 2 \) and fall below 0.1 M\(_\odot\) yr\(^{-1}\) Mpc\(^{-3}\) after \( z = 1 \) (see also Fig. 7).

To demonstrate the effect of resolution, we also show the BH mass functions at \( z = 1–4 \) for the 68 Mpc/uhr run (dashed coloured lines). The low-mass end (\( \log (M_*/M_\odot) < 8 \)) is significantly increased by up to 1 dex compared to the 500 Mpc/hr run as a consequence of a faster mass growth of low-mass BHs in the higher resolution run. This is in line with the increased low-mass end of the stellar mass function in the 68 Mpc/uhr run due to inefficient stellar feedback. Above \( \log (M_*/M_\odot) > 8 \), however, the 68 Mpc/uhr predictions roughly converge against the lower resolution results of the 500 Mpc/hr run.

Compared to observational constraints for the evolution of the BH mass function (derived from integrating the continuity equation; e.g. Merloni & Heinz 2008), our predictions are broadly consistent: also in the observations the massive end of the BH mass function is hardly evolving after \( z = 1 \), while the low-mass end is still increasing. This trend already reflects the antihierarchical behaviour we will discuss in the course of this study in more detail. In addition, at \( z = 2 \), the observed BH mass function is peaking at \( \log (M_*/M_\odot) \sim 8 \) and has a number density of \( \log (\Phi) \sim -3 \) for \( \log (M_*/M_\odot) = 7 \) and a number density of \( \log (\Phi) \sim -5 \) for \( \log (M_*/M_\odot) = 9 \). This is also predicted by the 68 Mpc/uhr run.

"Figure 5. Evolution of the BH mass function in the 500 Mpc/hr simulation (solid coloured lines) and in the 68 Mpc/uhr simulation (dashed coloured lines). The present-day BH mass function (red line) is compared to observations from Marconi et al. (2004) and Shankar et al. (2004, 2009) (black lines and symbols with the grey shaded areas). We find a reasonable agreement between observations and simulations for BH masses between \( 5 \times 10^7 M_\odot < M_* < 3 \times 10^9 M_\odot \) while above, the amount of massive BHs (>\( 3 \times 10^9 M_\odot \)) is overestimated by up to 1 dex (as a consequence of too inefficient radio-mode feedback)."
The slope of the relation is similar to the one of the present-day relation, but the BH masses are mainly found to be undermassive for a given stellar mass.

Between $z = 1$ and 3, the BH-to-stellar mass ratio is slightly increasing for a given stellar mass, i.e. BHs are growing slightly faster than the corresponding galaxy stellar masses. In contrast, afterwards ($z < 1$), this ratio decreases until today, i.e. the relation is just shifted towards more massive galaxy masses, what is most likely due to the strong late growth of massive galaxies. Overall, this is mainly consistent with previous results from cosmological simulations (Di Matteo et al. 2008).

### 3.4 The SFR and BH accretion rate densities

Observations reveal that SFR and BH accretion rate densities ($\dot{\rho}_{\text{stellar}}$ and $\dot{\rho}_{\text{bh}}$, respectively) trace each other over cosmic time with $\dot{\rho}_{\text{stellar}} \sim 2 \times 10^3 \times \dot{\rho}_{\text{bh}}$ (e.g. Zheng et al. 2009, and references therein). In Fig. 7, we show the predicted cosmic evolution of the SFR (orange lines) and BH accretion rate densities (red lines, multiplied with a factor of $2 \times 10^3$) for the 500 Mpc/hr (solid lines) and for the 68 Mpc/uhr run (dashed lines). At high redshifts $z > 1.5$, both the SFR and BH accretion rate densities are larger in the 68 Mpc/uhr run than in the 500 Mpc/hr one. This explains the increased low-mass end of the stellar and BH mass function at high redshifts in the 68 Mpc/uhr run (see Figs 4 and 5), where SF and gas accretion on to BHs is more efficient.

In both runs, SFR and BH accretion rates densities peak between $z = 1–2$ followed by a decline below $0.1 \ M_\odot \ yr^{-1} \ Mpc^{-3}$ after $z = 1$ (for the 500 Mpc/hr run). This is qualitatively consistent with the observational compilation for the SFR densities derived from different wavebands (coloured squares; Hopkins & Beacom 2006). At high redshifts $z > 4$, however, simulated BH accretion rate densities are too low compared to the observed SFR densities, which is a consequence of resolution (see improvement for the 68 Mpc/uhr run shown by the dashed red lines). Between

---

**Figure 6.** Upper panel: present-day relation between BH and stellar mass. Colour-coded contours show the simulated number density, the red line illustrates the mean for the 500 Mpc/hr run. The simulated relation perfectly matches the observational data (symbols illustrate their data set of Haring & Rix (2004), the black solid line shows the fit to the data and the black dashed lines the corresponding 1σ scatter). Bottom panel: evolution of the BH–stellar mass relation in the 500 Mpc/hr (solid lines) and the 68 Mpc/uhr run (dashed lines). For comparison, the observed present-day relation is also illustrated as in the upper panel. In the simulations, the BH–stellar mass relation is in place at $z = 3$ and hardly evolves afterwards.

**Figure 7.** Cosmic evolution of mean SF (orange lines) and BH accretion rate densities multiplied by a factor of $2 \times 10^3$ (red lines) in the 500 Mpc/hr (solid lines) and the 68 Mpc/uhr run (dashed lines). Both, SF and BH accretion rate densities peak between $z = 1–2$ followed by a decline towards lower redshifts. Compared to observed cosmic SFR densities derived from different wavebands (coloured symbols; Hopkins & Beacom 2006), simulations predict too high SFRs at $z < 1$ resulting in too massive galaxies (see Fig. 4).
The contribution to the total bolometric luminosity by radiatively inefficient AGN (their luminosity is calculated according to equation 7, illustrated by the orange, solid lines) is significantly increasing down to \( z = 0 \) and is particularly dominating the low-luminosity end at lower redshifts. This behaviour is mainly driven by an increasing amount of massive BHs accreting lower Eddington ratios as we will discuss in Section 5.2 in more detail.

Above \( z = 3 \), the 500 Mpc/hr simulation run starts to underestimate the amount of AGN for the entire luminosity range by up to two orders of magnitude (at \( z = 5 \)). This can be largely seen as a consequence of a combination of both insufficient resolution for accretion in mainly low-mass BHs and too massive seed BHs (only a minor issue). Instead, for the 68 Mpc/uhhr run (green dashed lines), the amount of low-luminous AGN is significantly increased resulting in a reasonably good agreement with the faintest AGN in the observations at \( z = 4 \).

At \( z = 5 \), in order to provide a fair comparison between observations and simulations, a larger cosmological volume would be needed combined with an increased resolution than currently adopted in the 500 Mpc/hr and the 68 Mpc/uhhr runs. For the present, this, however, represents a great challenge for the currently available computational power. In a recent study of DeGraf et al. (2012), they have performed cosmological simulation (using GADGET2 with a similar model for BH growth) with a large box size of \((500 \text{ Mpc} \cdot \text{h}^{-1})^3\), but only run down to \( z = 5 \). When considering the AGN luminosity function at \( z = 5 \) and 6 they obtain a fairly good agreement to observational data at these high redshifts.

Nevertheless, the overall good agreement between simulations and observations (up to \( z = 3–4 \)) may indicate that the BH growth closely follows the gas density and other physical quantities included in the Bondi–Hoyle accretion formula equation (1) in the resolved vicinity of the BH. This seems to be the case, although simulations are not able to capture the physical processes on small scales, i.e. they do not resolve the inner parts of the galaxy (3.75 kpc \( h^{-1} \)) and thus, adopt a very rough approximation for the accretion process itself. We will discuss this in more detail in Section 5.3.

A previous study of Degraf et al. (2010) also investigated the evolution of the AGN luminosity function in cosmological simulations with similar and larger resolution to ours. However, their simulations have a significantly smaller box size so that they can only probe the low-luminosity end of the AGN luminosity function (up to \( L_{\text{bol}} < 10^{45} \text{ erg s}^{-1} \)), which they find to be in reasonably good agreement with observations. In a very recent work of Khandai et al. (2014) they analyse a simulation with a larger box size of \((100 \text{ Mpc} \cdot \text{h}^{-1})^3\). They can match the low-luminosity end \((L_{\text{bol}} < 10^{45} \text{ erg s}^{-1})\) up to \( z = 4 \), but the luminous end is either significantly overestimated (at \( z < 0.5 \)) or underestimated (at \( z > 0.5 \)).

4.2 X-ray luminosities

By now, we compared our simulation predictions to the observational compilation of Hopkins et al. (2007) (see Fig. 8). In that study, they assume a luminosity dependence of the obscured fraction (the less luminous the more obscured) and the same number of Compton-thick \((N_H > 10^{24} \text{ cm}^{-2})\) and Compton-thin \((10^{23} \text{ cm}^{-2} < N_H < 10^{24} \text{ cm}^{-2})\) AGN. However, there are many aspects of the obscuration corrections that are still being vigorously debated.

Some recent studies suggest that the obscured fraction is dependent on both luminosity and redshift (Hasinger 2008; Fiore
Figure 8. Evolution of the bolometric AGN luminosity function for the 500 Mpc/hr (red solid lines) and the 68 Mpc/uhr run (green dashed lines) at $z = 0–5$. Simulation predictions of the 500 Mpc/hr run match the observational data of Hopkins, Richards & Hernquist (2007) (black, open circles) reasonably well until $z = 1.5–2.5$, even if at $z = 1.5–2.5$ the low-luminosity end is overestimated by up to 1 dex. At higher $z = 3–4$, the amount of AGN is underestimated in the 500 Mpc/hr run due to resolution effects, while the amount of faint AGN in the 68 Mpc/uhr run is increased by 1 dex providing a better match to the observational compilation. The orange lines illustrate the increasing contribution to the total luminosity by radiatively inefficient AGN with decreasing redshift.

et al. 2012), in contrast with the non-redshift-dependent model of Hopkins et al. (2007). There are also uncertainties with respect to the dust correction for the UV luminosity; Hopkins et al. (2007) compute the amount of dust (and therefore extinction), by adopting an $N_{HI}$ distribution from X-ray observations, and a Galactic dust-to-gas ratio. However, it has been shown that AGN absorbers do not have a Galactic dust to gas ratio (Maiolino, Marconi & Oliva 2001; Maiolino et al. 2004). The result is that they probably overestimate the extinction, which might result in slightly higher luminosities for the optically selected quasars. Because of these uncertainties, we additionally attempt in this subsection to correct our model predictions for obscuration and to directly compare them with recent soft (0.5–2 keV) and hard (2–10 keV) X-ray measurements of the AGN luminosity function (Hasinger et al. 2005; Ebrero et al. 2009; Aird et al. 2010).

In contrast to the previous subsection, we do not attempt to correct the observations for obscuration (as is done for the observational compilation in Hopkins et al. 2007), but instead apply an obscuration correction to our models. We convert the modelled, bolometric luminosities into hard and soft X-ray luminosities (0.5–2 keV and >2 keV) using the bolometric correction according to Marconi et al. (2004). In their study, the hard and soft X-ray luminosities $L_{HXR}, L_{SXR}$ are approximated by the following third-degree polynomial fits:

$$\log(L_{HXR}/L_{bol}) = -1.54 - 0.24\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3$$  \hfill (8)

$$\log(L_{SXR}/L_{bol}) = -1.65 - 0.22\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3$$  \hfill (9)

with $\mathcal{L} = \log(L_{bol}/L_{\odot}) – 12$. These corrections are derived from template spectra, which are truncated at $\lambda > 1 \mu m$ in order to remove the IR bump and which are assumed to be independent of redshift (therefore the resulting bolometric corrections are also assumed to be redshift independent).

Additionally, we apply a correction for obscuration to the model luminosities, as suggested by several observational studies (Ueda et al. 2003; Hasinger 2004; La Franca et al. 2005), in which it has been shown that the fraction of obscured AGN is luminosity dependent and decreases with increasing luminosity. While older studies such as Ueda et al. (2003) and Steffen et al. (2003) did not find a clear dependence of obscuration on redshift, several recent observational studies (Ballantyne et al. 2006; Gilli, Comastri & Hasinger 2007; Hasinger 2008) propose a strong evolution of the obscured AGN population (with the relative fraction of obscured AGN increasing with increasing redshift). Here, we follow the study...
of Hasinger (2008), where they compare the same AGN in both the soft and hard X-ray band so that they can derive an approximation for the obscured fraction $f_{\text{obsc}}$ in the soft X-ray band.

The obscured fraction at $z < 2$ is then given by the equation
\[ f_{\text{obsc}}(z, L_{\text{SXR}}) = -0.281(\log(L_{\text{SXR}}) - 43.5) + 0.279(1 + z)^\alpha, \]
with $\alpha = 0.62$ provides the best fit to their observational data. The obscured fraction at $z > 2$ is approximately the same at $z = 2$:
\[ f_{\text{obsc}}(z, L_{\text{SXR}}) = -0.281(\log(L_{\text{SXR}}) - 43.5) + 0.551. \]

Note that according to equations (10) and (11), the obscured fractions can get negative or larger than 1, while we additionally impose that negative values are equal to zero and we set values larger than 1 equal to 1. By calculating the obscured fraction of AGN in the soft X-ray band, we can model the visible fraction of AGN $f_{\text{vis}} = 1 - f_{\text{obsc}}$ and thus, the visible number density of AGN in the soft X-ray range is given by
\[ \Phi_{\text{vis}}(L_{\text{SXR}}) = f_{\text{vis}} \times \Phi_{\text{total}}(L_{\text{SXR}}). \]

The first and second rows in Fig. 9 show the soft X-ray luminosity function at $z = 0.01$–4.8 (different panels) in the 500 Mpc/hr (solid lines) and 68 Mpc/urh (dashed lines) simulation runs without (red and green lines) and with dust obscuration (blue and orange lines) and in observations (black stars and circles; Hasinger et al. 2004; Ebroer et al. 2009). At the high-luminosity end ($L_{\text{SXR}} > 10^{44}$ erg s$^{-1}$), obscuration does not influence the results and thus, the same trends can be seen as already discussed in Fig. 8, i.e. a good match with the observations.

However, turning to the low-luminosity end, at $z < 3.2$, simulation results without dust correction overestimate the amount of moderately luminous AGN by up to 1 dex in the 500 Mpc/hr run and by up to 2 dex in the 68 Mpc/urh run. Instead, the simulation predictions including dust obscuration can achieve a fairly good agreement with the observations. This implies that our results are consistent with the observational conclusions of Hasinger (2008) supporting the adoption of a redshift-dependent dust obscuration.

Note that the adopted dust obscuration model is the same as the one in a recent study of Hirschmann et al. (2012) and similar to the one adopted in Fanidakis et al. (2012), where they use the semi-analytic model of Somerville et al. (2008) and Bower et al. (2006), respectively, to study the evolution of the AGN luminosity function. In these works, they also find a good agreement between models and observational data when adopting a redshift- and luminosity-dependent dust obscuration and comparing their model predictions to observed soft X-ray luminosities.

The third to fifth rows of Fig. 9 illustrate the hard X-ray luminosity functions predicted by the two simulation runs for different redshift ranges $z = 0.01$–3.5, compared with observational data (Ebroer et al. 2009; Aird et al. 2010). The simulation predictions without dust obscuration (red solid and green dashed lines) are able to reproduce the high-luminosity end pretty well, while the amount of fainter AGN is strongly overestimated, particularly in the 68 Mpc/urh run above $z = 1$.

One possible explanation might be again due to obscuration as even 2–10 keV X-ray observations might miss a significant fraction of moderately obscured AGN (25 per cent at $N_H > 10^{21}$ cm$^{-2}$) and nearly all Compton-thick AGN ($N_H > 10^{26}$ cm$^{-2}$; Treister et al. 2004; Ballantyne et al. 2006). From fits to the cosmic X-ray background, Gilli et al. (2007) predict that both moderately obscured and Compton-thick AGN are as numerous as unobscured AGN at luminosities higher than $log(L_{0.5–2 keV}) > 43.5$ [ergs s$^{-1}$], and four times as numerous as unobscured AGN at lower luminosities ($log(L_{0.5–2 keV}) < 43.5$ [ergs s$^{-1}$]).

In addition, in a study of Merloni (2014), they investigate a complete sample of 1310 AGN from the XMM-COSMOS survey between $z = 0.3$–3.5 and classify them to be obscured in the hard X-ray band due to the shape of the X-ray spectrum. Their results confirm earlier results that at each redshift there is a clear decrease of the fraction of obscured AGN with increasing luminosity. In addition (and in contrast to an optical classification), nuclear obscuration in the hard X-ray band also reveals a dependence on redshift, where the fraction of obscured AGN increases with increasing redshift.

For this reason, we considered the corresponding observed, (hard X-ray) obscured fractions of AGN at a given redshift and luminosity (table 1 in Merloni 2014) in our simulations. Due to the limited redshift range, we have assumed that obscured fractions at $z < 0.3$ are the same as at $z = 0.3$ and that the obscured fractions at $z > 3.5$ stay also constant. This results in a fairly good match to the observed low-luminosity end at $z = 0.01$–1.5 (blue solid and orange dashed lines in Fig. 9). At higher redshifts $z > 1.5$, moderately luminous AGN are still somewhat too numerous in the simulations (even if improved), particularly in the 68 Mpc/urh run. Nevertheless, at $z \sim 0$–3.5, our simulation results are widely consistent with the existence of a redshift- and luminosity-dependent obscured fraction of AGN in the hard X-ray band as suggested by observational data of Merloni (2014).

4.3 The radio luminosity function at $z = 0$

In the last subsections, we were focusing on radiative AGN luminosities above $10^{31}$ erg s$^{-1}$ with the main contribution from radiatively efficient AGN. Radiatively less efficient AGN, however, are often observed to produce energetic, powerful radio jets and as a consequence hot X-ray cavities and thus, heat the ISM. Therefore, in this section we explicitly investigate the present-day radio luminosity function by distinguishing between low excitation ('radio-mode', LERG) and high excitation ('quasar-mode', HERG) radio galaxies with Eddington ratios $f_{\text{edd}} < 0.01$ and $f_{\text{edd}} > 0.01$, respectively. This allows us to study particularly all the AGN accreting at very low Eddington ratios and to quantify their number density.

To derive a radio luminosity, we follow a recent observational study of Best & Heckman (2012) who investigate a sample of 18 286 radio-loud AGN by combining data from different surveys and derive the radio luminosity functions for the first time separately for the LERGs and HERGs. In their fig. 6, they show histograms for the ‘total’ Eddington ratios $f_{\text{edd}, \text{tot}} = (L_{\text{rad}} + L_{\text{mech}})/L_{\text{edd}}$ of the HERGs and LERGs, which are peaking at ratios of $\sim$1.6 and $\sim$3, respectively. We adopted Gaussian distributions to model these ‘total’ Eddington-ratio distributions for HERGs and LERGs with a standard deviation of 0.3 and 0.4, respectively. For each AGN, we randomly draw a ‘total’ Eddington ratio so that we can derive the mechanical jet luminosity $L_{\text{mech}}$ for the HERGs and the LERGs:

\[ L_{\text{mech, HERGs}} = f_{\text{Edd, tot, HERGs}} \times L_{\text{edd}} - L_{\text{rad}} \]
\[ L_{\text{mech, LERGs}} = f_{\text{Edd, tot, LERGs}} \times L_{\text{edd}} - L_{\text{rad}}. \]
Figure 9. First and second rows: evolution of the soft X-ray ($0.5−2$ keV) AGN luminosity function at $z = 0−4.8$. Third, fourth and fifth rows: evolution of the hard X-ray ($2−10$ keV) AGN luminosity function at $z = 0−3.5$. Blue (orange) and red (green) solid (dashed) lines illustrate the results of the 500 Mpc/hr (68 Mpc/uhr) run with and without accounting for dust obscuration, respectively. For soft and hard X-ray luminosities at $z = 0−1.6$ and at $z = 0−3.5$, respectively, we find a good match with observational data (black symbols; Hasinger et al. 2005; Ebrero et al. 2009; Aird et al. 2010) when accounting for dust obscuration.
5 ANTIHIERARCHICAL GROWTH OF BHs

5.1 Comoving number density evolution of AGN

In Section 4 we have shown that a combination of the 500 Mpc/hr and the 68 Mpc/ahr simulation runs can reproduce the evolution of the AGN luminosity function from $z = 4$ to $0$ reasonably well. This implies that our simulations should also be able to naturally reproduce the observed ‘antihierarchical’ or ‘downsizing’ trend in BH growth within the framework of a hierarchical structure formation scenario. As described in Section 1, the term ‘downsizing’ refers to the observational result that the number density of luminous AGN peaks before successively less luminous AGN.

This is visualized in Fig. 11 showing the redshift evolution of the AGN number densities binned in different bolometric luminosity bins (different coloured lines as indicated in the legend). The observational data of Hopkins et al. (2007) (open stars with the dashed lines and the grey shaded area) reveal a time evolution of the peaks of the luminosity curves being characteristic for the downsizing trend. At $z < 3$, the 500 Mpc/hr run (filled squares with solid lines) matches the observations reasonably well (only slightly overpredicts the amount of moderately luminous AGN), while at higher redshifts, the 68 Mpc/ahr run (filled circles with dotted-dashed lines) agrees better with the observations (due to higher resolution), but only for moderately luminous AGN (due to small box size).

This clearly shows that the simulations can correctly capture the observed time evolution of the peaks of the different luminosity curves, and thus, the corresponding downsizing trend. The strong decrease of luminous AGN (see yellow and red curves) at low redshift $z < 2$ is mainly due to the decreasing cold gas content in the vicinity of the massive BHs, i.e. if a merger happens (which is typically supposed to induce nuclear activity), it is preferentially a ‘dry’ merger not triggering any significant AGN activity. Instead, moderately luminous AGN probably are either low-mass BHs with high accretion rates (close to the Eddington rate) or massive BHs, which are – due to the small amount of cold gas which is left – accreting at low rates and which have most likely been very luminous in the past. This will be discussed and analysed in more detail in the next sections.

5.2 Implications on the connection between BH mass and AGN luminosity

In this section, we examine consequences and implications of the downsizing trend in the AGN evolution on the connection between...
masses of active BHs, the Eddington ratios and the bolometric luminosities of AGN.

5.2.1 The bolometric luminosity–BH mass plane

Fig. 12 shows a 2D histogram (AGN number density is colour-coded) of the bolometric AGN luminosity–BH mass plane at different redshifts $z = 0$–3 for the 500 Mpc/hr simulation run. To guide the eye, the red lines indicate an accretion at the Eddington rate (solid line), at 1/10 (dashed line) and at 1/100 (dotted-dashed line) of the Eddington rate. At high redshifts, BHs accreting at or close to the Eddington rate dominate. Instead, towards lower redshifts the contribution of massive BHs accreting at very low Eddington ratios significantly increases explaining the decline of the number density of luminous AGN with decreasing redshift (see the red and yellow lines in Fig. 11).

Figure 12. 2D histograms (number density is colour coded) of the bolometric luminosity–BH mass plane at different redshifts $z = 3$–0 for the 500 Mpc/hr simulation run. To guide the eye, the red lines indicate an accretion at the Eddington rate (solid line), at 1/10 (dashed line) and at 1/100 (dotted-dashed line) of the Eddington rate. At high redshifts, BHs accreting at or close to the Eddington rate dominate. Instead, towards lower redshifts the contribution of massive BHs accreting at very low Eddington ratios significantly increases explaining the decline of the number density of luminous AGN with decreasing redshift (see the red and yellow lines in Fig. 11).

AGN triggered by mergers, with a first phase of high accretion close to the Eddington rate, followed by a power-law decline in the accretion rate. We expect that this is broadly traced within our cosmological simulation and plan to study the detailed light curves in a higher resolution simulation in a follow-up study.

Finally, Fig. 12 shows that the number density of luminous AGN ($L_{\text{bol}} > 10^{45}$ erg s$^{-1}$) is significantly smaller (almost completely suppressed) at $z = 0$ than it is at $z = 2$. This is mainly due to the available gas content, which gets reduced particularly in massive galaxies over cosmic time. This is also consistent with observational results of e.g. Steinhardt & Elvis (2010) and Kelly & Shen (2013), who find that the probability of massive BHs accreting close to the Eddington rate is significantly reduced at low redshift.

5.2.2 The evolution of the Eddington ratios

Fig. 13 shows the redshift evolution of the Eddington ratio distributions for the 500 Mpc/hr (solid, coloured lines) and the 68 Mpc/uch simulation run (dashed colour lines). The vertical dotted line visualizes the adopted division between the radio and quasar mode in the simulation code. Consistent with our previous analysis, and often adopted luminosity limits of observational samples, we consider solely AGN with bolometric luminosities larger than $10^{44}$ erg s$^{-1}$.

The simulated fraction of AGN which are accreting at small Eddington ratios ($f_{\text{edd}} < 0.01$) increases strongly with decreasing redshift and thus, the peaks of the distribution curves are clearly shifted towards smaller Eddington ratios with decreasing redshift in both simulation runs – a typical implication of the downsizing trend: at $z = 4$, the distribution peaks at $f_{\text{edd}} \sim 1$, while at $z = 0$ the peak is located around $f_{\text{edd}} \sim 5 \times 10^{-4}$. This illustrates the same trend we discussed for Fig. 12: at later times BHs spend less time accreting at the Eddington rate, but mainly reside in a decline dominated ‘blowout’ accretion phase.
Eddington ratio distributions at different redshifts $z = 1$. In contrast, the most massive BHs are preferentially accreting within a broad range of Eddington ratios 0.01 $< f_{\text{edd}} < 1$. To summarize, the majority of AGN at $z = 0$ are hardly radiating at or near the Eddington limit, which is in qualitative agreement with several observational studies (Vestergaard 2003; Kollmeier et al. 2006; Kelly et al. 2010; Schulze & Wisotzki 2010; Kelly & Shen 2013).

As a further consequence, the number of BHs accreting close to the Eddington rate is low at $z = 0$ (log ($dN/df_{\text{edd}}$) $\sim -4.5$ Mpc$^{-3}$ dex$^{-1}$), while at high redshifts ($z = 1$) a large number of BHs (log ($dN/df_{\text{edd}}$) $\sim -3.5$ Mpc$^{-3}$ dex$^{-1}$) are preferentially accreting within a broad range of Eddington ratios 0.01 $< f_{\text{edd}} < 1$. To summarize, the majority of AGN at $z = 0$ are hardly radiating at or near the Eddington limit, which is in qualitative agreement with several observational studies (Vestergaard 2003; Kollmeier et al. 2006; Kelly et al. 2010; Schulze & Wisotzki 2010; Steinhardt & Elvis 2010; Kelly & Shen 2013). Kelly et al. (2010), for example, show that the Eddington ratio distribution (using broad-line quasars between $z = 1$) peaks at an Eddington ratio of $f_{\text{edd}} = 0.05$.

In Fig. 14 we present a quantitative comparison of simulated Eddington ratios in the 500 Mpc/hr (solid lines) and the 68 Mpc/xhr run with observational data: the simulated mean Eddington ratios of AGN with luminosities above $10^{43}$ erg s$^{-1}$ are plotted versus redshift (see solid and dashed black lines with grey shaded areas in both panels) and compared to observational data of Shen & Kelly (2012) (black, filled circles). In both simulations and observations, the mean Eddington ratios are decreasing with decreasing redshift and we find a nearly perfect agreement between them at $z < 4$.

Additionally, in the simulations (only explicitly shown for the 500 Mpc/hr run), we have divided the AGN sample into different BH mass bins (solid, coloured lines in the top panel of Fig. 14) and different AGN luminosity bins (coloured lines in the bottom panel of Fig. 14). This illustrates that on average, for the entire redshift range since $z < 4$, lowest mass BHs (blue line in the top panel) and most luminous AGN (red line in the bottom panel) accrete at the highest Eddington ratios ($0.1 < f_{\text{edd}} < 1$). In contrast, the most massive BHs (red line in the top panel) and the lowest luminous AGN (blue line in the bottom panel) have the lowest mean Eddington ratios. This behaviour is consistent with Fig. 12.

Finally, we want to point out that the interplay between BH mass, AGN luminosity and Eddington ratio, we have discussed in this section, is in good qualitative agreement with previous studies of Fanidakis et al. (2012) and Hirschmann et al. (2012) who also reproduced the downsizing trend in BH growth, but by using semi-analytic models. This confirms that such a connection between BH masses and AGN luminosities seems to be a necessary condition for reproducing a downsizing trend in BH growth.

5.3 Physical origin of the downsizing trend in our simulations

A common problem of all currently existing, large cosmological simulations is that they are not able to capture the detailed physical processes of gas accretion in the close vicinity of the BH due to too low resolution as a result of the (up to now) limited computational power. Besides, the physical description of BH accretion is also more complicated than typically assumed in large-scale simulations (e.g. in reality, there is no spherical symmetry King & Nixon 2013 and the Bondi accretion seems to be also a poor prescription for cold
Irrespectively of the redshift, the gas temperature and relative velocity is increasing, while the gas density is decreasing with increasing BH mass. In addition, the gas density at a given BH mass is decreasing with decreasing redshift.

(chaotic) gas accretion, Gaspari, Ruszkowski & Oh 2013). Thus, a very rough approximation for the accretion process (following the Bondi accretion formula) is generally adopted when performing large-scale simulations as we have described earlier in Section 2.2.

Nevertheless, as we have seen in the course of this study, simulations seem at least to be able to reproduce the right trends of how BHs assemble their mass and appear as an AGN despite their relatively low spatial and mass resolution (compared to the physical scales at which the gas accretion is actually happening). In other words, the general scheme in our simulations is able to capture the essence of BH growth ‘in reality’. Therefore, to better understand the physical reason for automatically reproducing the downsizing trend in our simulations, we directly investigate the different gaseous quantities, as the density, the sound speed, which is proportional to the gas temperature and the relative velocity (between the gas particles and the BH), which are contributing to the estimation of the Bondi accretion rate in the code given by equation (1) by averaging over all the gas particles within the numerically resolved accretion region of the BH. Note that ‘numerically resolved accretion region’ is supposed to describe the SPH-like smoothing length of the BH sink particles.

Fig. 15 shows the mean (solid lines) and the corresponding 1σ scatter (dotted lines) of the gas density (top row), of the gas temperature (middle row) and of the relative gas velocity (bottom row) within the numerically resolved accretion region versus BH mass at $z = 2.8$ (green lines), $z = 1$ (orange lines) and $z = 0$ (red lines) for the 500 Mpc/hr (solid lines) and the 68 Mpc/uhr run (dashed lines). The left-hand column corresponds to AGN with $\log(L_{\text{bol}}) > 43$ [erg s$^{-1}$], while in the middle and right-hand column we additionally distinguish between AGN in the quasar and the radio mode, respectively. Irrespectively of the redshift, the gas temperature and relative velocity is increasing, while the gas density is decreasing with increasing BH mass. In addition, the gas density at a given BH mass is decreasing with decreasing redshift.
Turning now to the redshift dependence of the average gas quantities (coloured lines in the left-hand column of Fig. 15), we find for massive BHs ($M_\bullet > 10^7 M_\odot$) a significant decrease of the mean gas density by almost three orders of magnitudes from $z = 2.8$ to 0, which most likely causes a decrease of the relative accretion rates at a given BH mass towards lower redshifts (even if the gas temperature, and thus the sound speed, is also slightly decreasing towards lower redshifts). This can, therefore, explain the decrease in the number density of very luminous AGN with $\log(L_{5100}) \geq 46$ [erg s$^{-1}$] (i.e. very massive BHs $M_\bullet \geq 10^7 M_\odot$ accreting at high Eddington ratios) since $z \sim 2$ (see yellow and red curves in Fig. 11). A further consequence of the decreasing accretion rates of massive BHs with decreasing redshift is that these objects have now preferentially moderate luminosities and thus, can contribute to the weakly increasing number density of *moderately luminous* AGN with evolving time (see black and dark/light blue lines in Fig. 11).

Instead, for low-mass BHs ($M_\bullet < 10^7 M_\odot$), the density, the gas temperature and the relative velocity (between the gas and the BH) are all slightly decreasing down to $z = 1$ in the 500 Mpc/hr run. However, after $z = 1$, the gas density is hardly decreasing anymore, the gas temperature stays constant, and the relative velocity is slightly decreasing suggesting that the accretion for low-mass BHs is approximately constant or even slightly increasing with decreasing redshift. This ‘not-decreasing’ accretion in low-mass BHs with decreasing redshift (after $z = 1$) can also contribute to the weak increase of the number densities of moderately luminous AGN (see black and dark blue lines in Fig. 11) – in addition to the massive BHs accreting at low Eddington ratios.

Considering the higher resolved 68 Mpc/uvhr run, the gas density in such low-mass BHs is higher, while the relative velocity is lower than in the 500 Mpc/hr run. This leads to a stronger and faster growth of low-mass BHs, as e.g. seen for the low-mass end of the BH mass function. This is also responsible for the increase of the low-luminous end of the AGN luminosity function in the 68 Mpc/uvhr run. In addition, these trends suggest that the $\alpha$ parameter in equation (1) needs to be adjusted when going to higher and higher resolution, as also discussed in Booth & Schaye (2011) and Vogelsberger et al. (2013).

Overall, the redshift and mass evolution of the different gaseous quantities contributing to the Bondi accretion can be a result of the interplay of radio- and quasar-mode accretion as adopted in our model: at high redshifts $z > 2$ for the whole mass range and at lower redshifts for low-mass BHs ($M_\bullet \leq 10^7 M_\odot$), the gaseous quantities are driven by the evolution in the quasar mode (see middle column of Fig. 15). This is the dominating accretion mechanism for these regimes and thus, is responsible for the overall decrease of the mean densities, temperatures and relative velocities of low-mass BHs towards low redshifts. Instead, for low redshifts $z \leq 1$ and more massive BHs ($M_\bullet \geq 10^7 M_\odot$), the accretion in the radio mode is dominating the gaseous quantities (right-hand column of Fig. 15).

The AGN feedback (in the radio mode) together with the gas consumption due to SF in massive galaxies over cosmic time leads to a significantly decreasing gas density with increasing BH mass and decreasing redshift. This is not the case for an accretion at the quasar mode, where at $z = 0$ the gas density stays constant with increasing BH mass and for the most massive BHs, the gas density tends to be even slightly larger at $z = 0$ than at $z = 1$.

To summarize, we find that the downsizing trend in the AGN number density evolution and thus, the BH accretion, in our simulations is mainly caused by ‘global’ gas physical properties in the vicinity of the resolved accretion region and not necessarily by the ‘local’ gas properties within the (unresolved) Bondi accretion radius. Interestingly, this is in agreement with detailed, high-resolution (spatial resolution of 200 pc) 3D simulations of BH accretion, which indicate that the BH accretion is mainly driven by large-scale gas properties beyond 1 kpc (Novak, personal communication; Novak, Durier & Babul in preparation) and not by the local gas properties.

Essentially, we have demonstrated that a combination of both a decreasing mean gas density of BHs with increasing BH mass and decreasing redshift and a slightly lower mean relative velocity of low-mass BHs at lower redshifts (the latter being of minor importance) is causing the downsizing trend in the AGN number density evolution.

### 6 SUMMARY AND DISCUSSION

In this study, we have analysed a subset of hydrodynamic, fully cosmological simulations from the Magneticum Pathfinder simulation set (Dolag et al., in preparation) focusing on the evolution of statistical properties of BH growth and AGN luminosities from $z = 5$ down to $z = 0$, in particular the observed downsizing trend in BH growth. The simulations are based on an improved SPH code GADGET-3 (Springel 2005) where we are additionally using a higher order kernel based on the bias-corrected, sixth-order Wendland kernel (Dehnen & Aly 2012) with 295 neighbours which together with a low-viscosity SPH scheme allows us to properly track turbulence within galaxy clusters (Dolag et al. 2005; Donnert et al. 2013).

The code also contains radiative gas cooling, a spatially uniform UV background and a multiphase model for SF with the associated feedback processes (Springel & Hernquist 2003). The simulations also include a detailed model of chemical evolution according to Tornatore et al. (2007), where metals are released by SNII, SNIa explosions and AGB stars. We also allow for isotropic thermal conduction with 1/20 of the classical Spitzer value (Dolag et al. 2004). In addition, the prescriptions for BH growth and feedback from AGN are based on the models presented in Springel et al. (2005) and Di Matteo et al. (2005), but contain the same modifications for a transition from a quasar- to a radio-mode feedback as in the study of Fabjan et al. (2010). It also contains some new changes (as discussed in Section 2.2), where the most noticeable difference is reflected in our ability to follow BHs properly in galaxies which are inside of galaxy clusters.

We have considered two simulation runs, one with a large comoving volume of (500 Mpc)$^3$ and one with a smaller volume of (68 Mpc)$^3$ but a higher resolution with initial conditions based on the WMAP7 cosmology. We now summarize our main results.

(1) Consistent with previous studies (e.g. Di Matteo et al. 2008), our simulations are in reasonably good agreement with the observed present-day BH–stellar mass relation and the BH mass function, although the high-mass end is overestimated due to too inefficient radio-mode feedback. The latter point is also true for the stellar mass function, i.e. the simulations tend to produce too many too massive galaxies hosting too massive BHs after $z = 1$. The main growth of BHs is found to occur before $z \sim 1$ when considering the evolution of the BH mass function, which is largely in place at $z \sim 1$ in qualitative agreement with observational estimates for the evolution of the BH mass function (Merloni & Heinz 2008). This is also consistent with the cosmic evolution of BH accretion rates, which peak around $z \sim 1.5$ and start to strongly decline towards lower redshifts.

(2) Our large-volume simulation can successfully reproduce the observed bolometric AGN luminosity function up to $z \sim 3$ (Hopkins et al. 2007), even if moderately luminous AGN are slightly
overestimated at $z = 1.5 – 2.5$. For moderately luminous AGN this is consistent with a previous study of Degraf et al. (2010), but the high-luminosity end ($\log (L_{\text{bol}}/L_{\odot}) \geq 45$ erg s$^{-1}$) matches – for the first time – the observational data. At higher redshifts (up to $z \sim 5$), however, the large-volume simulation predicts a too small amount of AGN due to resolution effects. Turning to the run with a smaller volume but higher resolution helps to increase the amount of moderately luminous AGN in reasonably good agreement with observations. Even if the smaller volume, higher resolution simulation predicts more moderately luminous AGN than the large-volume run, it converges against the observational data for more luminous AGN ($\log (L_{\text{bol}}/L_{\odot}) > 45$ erg s$^{-1}$).

(3) Besides the bolometric luminosities, we have additionally extended our simulation predictions by dust obscuration (in post-processing) to compare them directly with recent soft and hard X-ray measurements of the AGN luminosity function. When adopting an empirically motivated dust obscuration model, where obscuration of AGN is dependent on both, redshift and luminosity, we find a reasonably good agreement between the large-volume simulation and observations up to $z \sim 3$ and up to a redshift $z = 4–5$ for the small volume, but higher resolution simulation.

Dividing our AGN in low and high excitation state, we are also able to reproduce observed number densities of present-day LERGs/HERGs radio galaxies, implying that the simulations can successfully capture the correct amount of radiatively efficient and inefficient AGNs.

(4) As a consequence of points (2) and (3), a combination of both simulation runs (small and large volume) can successfully reproduce the comoving number density evolution of AGN up to $z = 4–5$, and therefore, naturally predict the characteristic, observed downsizing trend within a hierarchical structure formation scenario. The strong decline of luminous AGN with decreasing redshift (since $z = 2$) can be attributed to the gas density in the vicinity of a (massive) BH, which gets successively depleted with evolving time due SF and AGN feedback. The gas inflow towards a massive BH is inhibited and its luminosity starts to fade (and can then contribute to the amount of moderately luminous AGN). This also explains why massive BHs preferentially accrete at low Eddington ratios after $z = 2$.

The continuous increase of moderately luminous AGN over time can, instead, be seen a result of both massive BHs in their ‘fading’ phase and lower mass BHs as they tend to have constant or slightly increasing accretion rates towards lower redshifts (after $z = 1$) due to slightly decreasing mean relative velocities between the BH and the surrounding gas with decreasing redshift.

(5) Finally, we find that the downsizing behaviour implies that the peaks of the Eddington ratio distributions are shifted towards successively smaller Eddington ratios with decreasing redshift – a trend which has also been found by observational studies. The number of BHs accreting close to the Eddington rate is more than one order of magnitude smaller at $z = 0$ than at high redshifts ($z = 1–3$), where the BHs are preferentially accreting within a broad range of Eddington ratios $0.01 < f_{\text{edd}} < 1$. Besides, our simulations can successfully predict the observed evolution of the mean Eddington ratio for AGN more luminous than $L_{\text{bol}} > 10^{43}$ [erg s$^{-1}$]. These implications of the downsizing behaviour are in qualitative agreement with predictions from semi-analytic models (see e.g. Fanidakis et al. 2012; Hirschmann et al. 2012).

Despite the overall success of the simulations being in reasonably good agreement with observational data we should keep in mind that this is attained despite of the rather crude, approximate and numerically limited description of BH accretion, energy extraction and thermalization (even if high-resolution simulations of BH accretion indicate an independence of the BH accretion on local gas properties up to 1 kpc around the BH; Novak, Durier and Babul in preparation). In addition, the simulations presented in this work fail to reproduce the massive end of the BH and the stellar mass functions at low redshift due to too inefficient radio-mode feedback.

But nevertheless, this study nicely demonstrates that the observed downsizing trend in BH growth is not contradictory with a hierarchical structure formation model (consistent with the predictions of semi-analytic models as shown in Bonoli et al. 2009; Fanidakis et al. 2012; Hirschmann et al. 2012), but instead is a ‘natural’ outcome due to specific baryon processes. Besides the ‘downsizing’ trend, there exist many other unresolved questions about BH growth and AGN evolution which can be easily assessed in our simulations. Therefore, in forthcoming studies, we plan to examine the connection between AGN and their host galaxy properties, AGN clustering properties and typical AGN light curves (of differently luminous AGN) which may particularly illuminate our current understanding of how strongly galaxies and BHs are co-evolving and which are the main trigger mechanisms for AGN activity (if there are any). Particularly the latter is – up to now – a major unresolved open issue in understanding BH growth and AGN evolution.

ACKNOWLEDGEMENTS

We thank Andrea Merloni, Francesco Shankar, James Aird and Jacobo Ebbero for providing us with observational data and Andrea Merloni, Greg Novak and Rachel Somerville for fruitful discussions.

This research was supported by the DFG Cluster of Excellence ‘Origin and structure of the Universe’. We are especially grateful for the support by M. Petkova through the Computational Center for Particle and Astrophysics (C²PAP). Computations have been performed at the at the ‘Leibniz-Rechenzentrum’ with CPU time assigned to the Project ‘pr86re’ as well as at the ‘Rechenzentrum der Max-Planck-Gesellschaft’ at the ‘Max-Planck-Institut für Plasmafysik’ with CPU time assigned to the ‘Max-Planck-Institut für Astrophysik’. MH acknowledges financial support from the European Research Council under the European Community’s Seventh Framework Programme (FP7-2007-2013)/ERC grant agreement no. 202781. SB acknowledges financial support from the European Commission’s Framework Programme 7, through the Marie Curie Initial Training Network Cosmo-Comp (PITN-GA-2009-23856), the PRIN-MIUR-2009 grant ‘Tracing the growth of structures in the Universe’ and the PD51-INFN grant.

REFERENCES

Aird J. et al., 2010, MNRAS, 401, 2531
Ballantyne D. R., Shi Y., Rieke G. H., Donley J. L., Papovich C., Rigby J. R., 2006, ApJ, 653, 1070
Barai P., Viel M., Murante G., Gaspari M., Borgani S., 2014, MNRAS, 437, 1456
Barger A. J., Cowie L. L., 2005, ApJ, 635, 115
Barger A. J., Cowie L. L., Capak P., Alexander D. M., Bauer F. E., Brandt W. N., Garmire G. P., Hornschemeier A. E., 2003, ApJ, 584, L61
Best P. N., Heckman T. M., 2012, MNRAS, 421, 1569
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1985, Nature, 313, 72
Bondi H., 1952, MNRAS, 112, 195
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Bongiorno A., 2012, MNRAS, 427, 3103
Maiolino R., Oliva E., Ghinassi F., Pedani M., Mannucci F., Mujica R., Juarez Y., 2004, A&A, 420, 889
Marchesini D. et al., 2007, ApJ, 656, 42
Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169
Marulli F., Bonoli S., Branchini E., Moscardini L., Springel V., 2008, MNRAS, 385, 1846
Matute I., La Franca F., Pozzi F., Gruppioni C., Lari C., Zamorani G., 2006, A&A, 451, 443
Menci N., Fiore F., Perola G. C., Cavaliere A., 2004, ApJ, 606, 58
Menci N., Fiore F., Lamastra A., 2013, ApJ, 766, 10
Merloni A., 2014, MNRAS, 437, 3550
Merloni A., Heinz S., 2008, MNRAS, 388, 1011
Miyaji T., Hasinger G., Schmidt M., 2000, A&A, 353, 25
Monaco P., Fontanot F., 2005, MNRAS, 359, 283
Mullany J. R. et al., 2012, MNRAS, 419, 95
Nandra K., Laird E. S., Steidel C. C., 2005, MNRAS, 360, L39
Neistein E., Netzer H., 2014, MNRAS, 437, 3373
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
Padovani P., Matteucci F., 1993, ApJ, 416, 26
Peebles P. J. E., 1965, ApJ, 142, 1317
Peng C. Y., 2007, ApJ, 671, 1098
Pérez-González P. G. et al., 2008, ApJ, 675, 234
Puchwein E., Springel V., 2013, MNRAS, 428, 2966
Richards G. T. et al., 2006, AJ, 131, 2766
Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, ApJ, 645, 986
Rosario D. J. et al., 2012, A&A, 545, A45
Rosas-Guevara Y. M. et al., 2013, arXiv:preprints
Salpeter E. E., 1964, ApJ, 140, 796
Sazonov S. Y., Revnivtsev M. G., 2004, A&A, 423, 469
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
Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20
Shankar F., Weinberg D. H., Shen Y., 2010, MNRAS, 406, 1959
Shankar F., Marulli F., Mathur S., Bernardi M., Bournaud F., 2012, A&A, 540, A23
Shen Y., Kelly B. C., 2012, ApJ, 746, 169
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Soltan A., 1982, MNRAS, 200, 115
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., 2002, MNRAS, 333, 649
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, ApJ, 596, L23
Steinhardt C. L., Elvis M., 2010, MNRAS, 402, 2637
Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, MNRAS, 428, 129
Thielemann F.-K. et al., 2003, Nucl. Phys. A, 718, 139
Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050
Treister E. et al., 2004, ApJ, 616, 123
Tremaine S. et al., 2002, ApJ, 574, 740
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123, 305
Vestergaard M., 2003, ApJ, 599, 116
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wiersma R. C. P., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Wolf C., Wisotzki L., Borch A., Dye S., Kleinheinrich M., Meisenheimer K., 2003, A&A, 408, 499
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181
Zel’Dovich Y. B., 1964, Sov. Phys. Dokl., 9, 195
Zheng X. Z. et al., 2009, ApJ, 707, 1566

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