Modelling the cosmological co-evolution of supermassive black holes and galaxies

F. Marulli, S. Bonoli, E. Branchini, L. Moscardini, V. Springel

Abstract

Marulli, Bonoli, Branchini, Moscardini, Springel 2008, MNRAS.368.1269M

We model the cosmological co-evolution of galaxies and their central supermassive black holes (BHs) within a semi-analytical framework developed on the outputs of the Millennium Simulation. This model, described in detail in Croton et al. (2006) and De Lucia & Blaizot (2007), introduces a 'radio mode' feedback from Active Galactic Nuclei (AGN) at the centre of X-ray emitting atmospheres in galaxy groups and clusters. Thanks to this mechanism, the model can simultaneously explain: (i) the low observed mass drop-out rate in cooling flows; (ii) the exponential cut-off in the bright end of the galaxy luminosity function; and (iii) the bulge-dominated morphologies and old stellar ages of the most massive galaxies in clusters. In this work we investigate how well this model can also reproduce the physical properties of BHs and AGN. We analyze the scaling relations, the fundamental plane and the mass function of BHs, and compare them with the most recent observational data. Moreover, we extend the semi-analytic model to follow the evolution of the BH mass accretion and its conversion into radiation, and compare the derived AGN bolometric luminosity function with the observed one. While we find, for the most part, a very good agreement between predicted and observed BH properties, the semi-analytic model underestimates the number density of luminous AGN at high redshifts, independently of the adopted Eddington factor and accretion efficiency. However, an agreement with the observations is possible within the framework of our model, provided it is assumed that the cold gas fraction accreted by BHs at high redshifts is larger than at low redshifts.

Model

Dark Matter

We have used the outputs of the **Millennium Simulation**, which followed the dynamical evolution of ~1e10 dark matter (DM) particles with mass 8.6e8 Msun/h in a periodic box of 500 Mpc/h on a side, in a **ΛCDM** "concordance" cosmological framework. The mass resolution is high enough to resolve the DM halo of 0.1 L. galaxies with ~100 particles. The short-range gravitational force law is softened on the co-moving scale 5 kpc/h which may be taken as the spatial resolution limit of the calculation. The cosmological parameters are consistent with determinations from the combined analysis of the 2-degree Field Galaxy Redshift Survey and the first-year WMAP data. The Millennium Simulation was carried out with a special version of the **GADGET-2** code at the Computing Centre of the **Max-Planck Society** in Garching, Germany. We make use of hierarchical merging trees extracted from this simulation which encode the full formation history of DM haloes and subhaloes, previously identified with, respectively, a friends-of-friends group-finder and an extended version of the SUBFIND algorithm. These trees constitute the backbone of our semi-analytic model, which is implemented during the post-processing phase.

Galaxies

We use the galaxy formation model of **Croton et al. 2006** and **De Lucia & Blaizot 2007**. It is assumed that as a DM halo collapses, a fraction of its mass is in the form of baryons and collapses with it. Initially, these baryons are in the form of a diffuse gas with primordial composition, but later they include gas in several phases as well as stars and heavy elements. Conventionally, with the simplifying assumption of an ideal gas which cools isobarically, the **cooling time** of the gas is computed as the ratio of its specific thermal energy to the cooling rate per unit volume. The **star formation** is assumed to occur at a rate proportional to the cold gas mass of the galaxy. Massive stars explode as **supernovae** shortly after star formation events and are assumed to reheat a gas mass proportional to the mass of stars. The amount of gas that leaves the DM halo in a **"super-wind"** is determined by computing whether excess supernovae energy is available to drive the flow after reheating of material to the halo virial temperature. We model the **disk instabilities** using the analytic stability criterion of Mo et al. 1998. In the Millennium Run, **substructures** are followed down to masses of 1.7e10 Msun/h, so that we can properly follow the motion of galaxies inside their hosting DM haloes until tidal truncation and stripping disrupt their subhaloes at this resolution limit. At this point, we estimate a survival time for the galaxies using their current orbit and the dynamical friction formula. After this time, the galaxy is assumed to merge onto the central galaxy of its own halo. **Galaxy mergers** induce starburst, which we describe using the **"collisional starburst"** prescription.

Black Holes and AGN

Quasar mode

In our model BHs accrete mass after a galaxy merger both through coalescence with another BH and by accreting cold gas, the latter being the dominant accretion mechanism. For simplicity, the BH coalescence is modelled as a direct sum of the progenitor masses, thus ignoring gravitational wave losses. We assume that the gas mass accreted during a merger is proportional to the total cold gas mass present, but with an efficiency which is lower for smaller mass systems and for unequal mergers:

$$\Delta M_{BH,Q} = \frac{f'_{BH}\, m_{cold}}{1+\left(\frac{280 km/s}{V_{vir}}\right)^2}$$

$$f'_{BH} = f_{BH}\frac{m_{sat}}{m_{central}}$$

Thus, any merger-induced perturbation to the gas disk (which might come from a bar instability or a merger-induced starburst) can, in principle, drive gas onto the central BH. However, the fractional contribution of minor mergers is typically quite small, so that accretion driven by major mergers is the dominant mode of BH growth in our scenario. This kind of accretion is also closely associated with starbursts, which occur concurrently. We do not model feedback from the quasar activity in the current model, but it can be approximately represented by an enhanced effective feedback efficiency for the supernovae associated with the intense starburst.

Radio mode

When a static hot halo has formed around a galaxy, we assume that a fraction of the hot gas continuously accretes onto the central BH, causing a low-energy `radio' activity in the galaxy centre. The BH mass accretion rate during these phases is postulated to scale as follows:

$$\dot{M}_{BH,R} = k_{AGN}\left(\frac{M_{BH}}{10^8 M_\Theta}\right)\left(\frac{f_{hot}}{0.1}\right)\left(\frac{V_{vir}}{200 km/s^{-1}}\right)^3$$

This accretion rate is typically orders-of-magnitude below the Eddington limit. In fact, the total mass growth of BHs in the radio relative to the quasar mode discussed below is negligible. It is also assumed that the radio mode feedback injects energy efficiently into the surrounding medium, which can reduce or even stop the cooling flow in the halo centres. The mechanical heating generated by this kind of BH mass accretion induces a modified infall rate of the following kind:

$$\dot{m}'_{cool} = \dot{m}_{cool} - \frac{L_{BH}}{0.5 V_{vir}^2}$$

$$L_{BH} = 0.1 \dot{M}_{BH} c^2$$

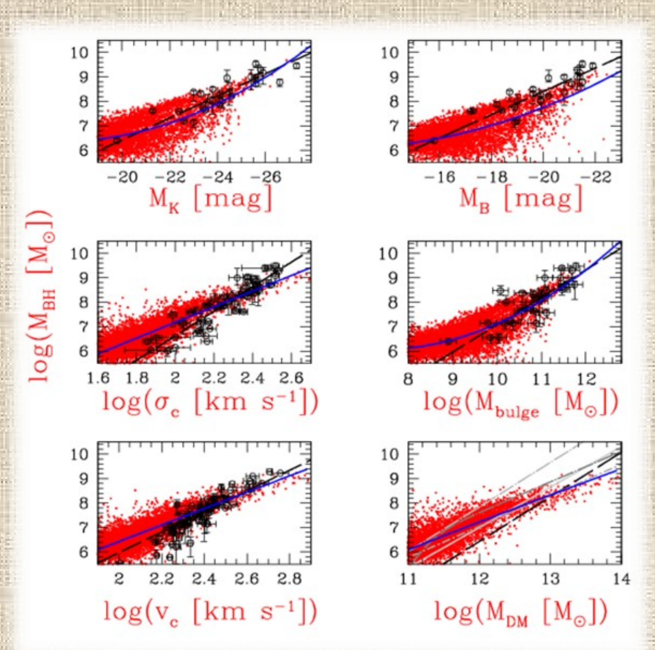
In this scenario, the effectiveness of radio AGN in suppressing cooling flows is greatest at late times and for large values of the BH mass, which is required to successfully reproduce the luminosities, colours and clustering of low-redshift bright galaxies.

We use the following definitions to parameterize the bolometric luminosity emitted by accretion onto BHs, as function of the *accretion efficiency*, ε, and the *Eddington factor*, f_{Edd}

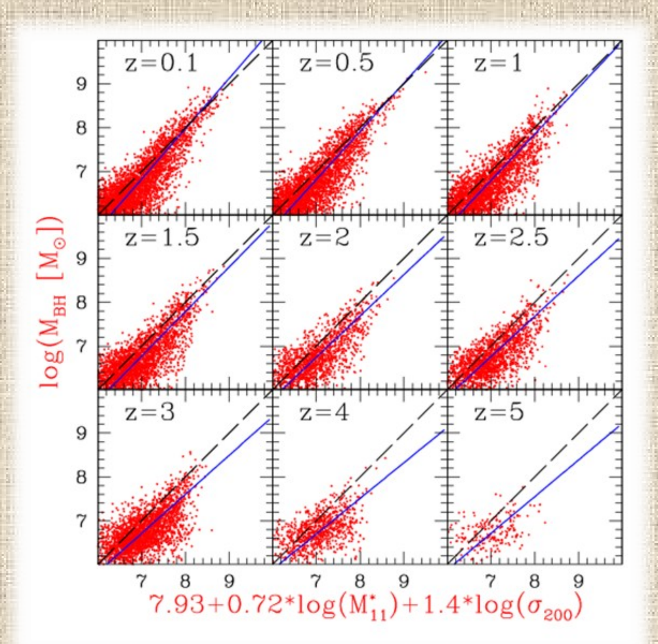
$$f_{Edd}(t) := L_{bol}(t)/L_{Edd}(t)$$

$$d\ln M_{BH}(t) = \frac{dt}{t_{ef}(t)}, \quad t_{ef}(t) = \frac{\varepsilon}{1-\varepsilon}\frac{t_{Edd}}{f_{Edd}(t)}$$

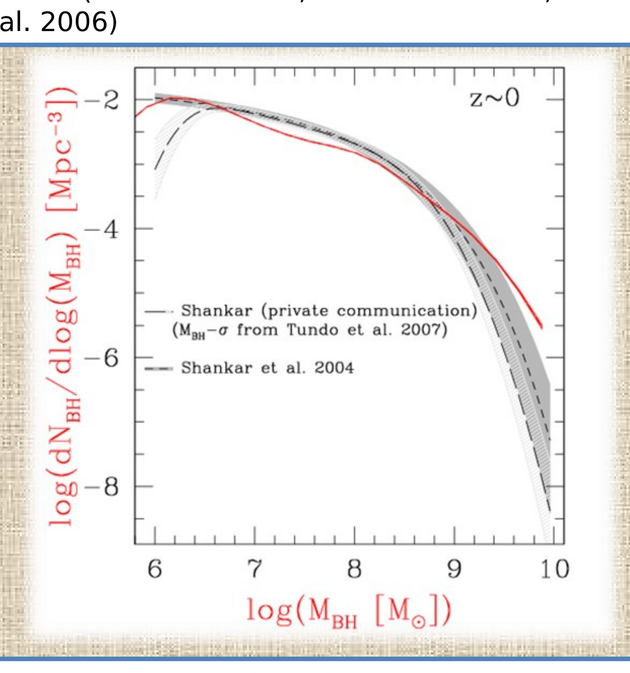
Model vs Observations



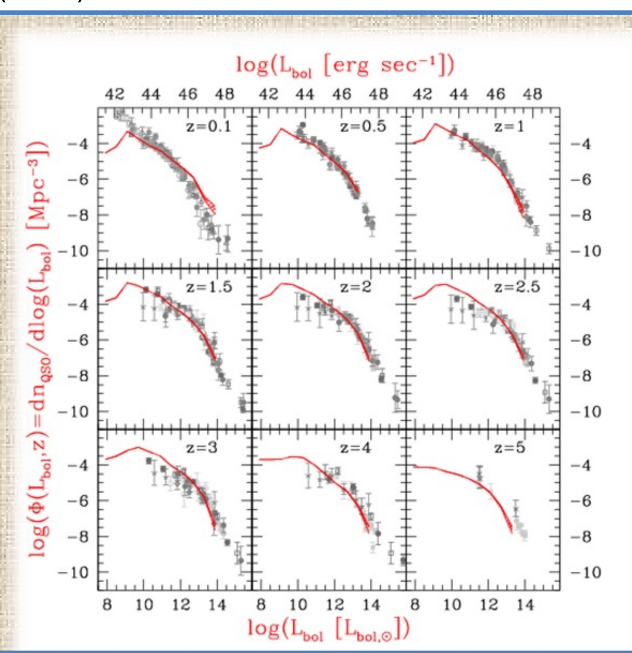
Scaling relations between the masses of the central BHs in the simulated galaxies (red dots and blue lines) with six different properties of their hosts (black dots and black lines): the K- and B-band bulge magnitude (Marconi et al. 2004), the bulge velocity dispersion (Ferrarese&Ford 2005) and mass (Haring&Rix 2004), the circular velocity of the galaxy (Baes et al. 2003) and the virial mass of the DM halo (Ferrarese 2002, Baes et al. 2003 , Shankar et al. 2006)



The **BH fundamental plane** in the redshift range 0.1< z < 5. The galaxy stellar mass is given in units of 1e11 Msun, while the bulge velocity dispersion is in units of 200 km/s. Red dots: model predictions; blue lines: best-fits to the model outputs; dashed black lines: Hopkins et al. (2007)



The **BH mass function** predicted by our model (red) compared with the one obtained by Shankar et al. (2004), and with the new one derived by Shankar et al. (private communication) using the Mbh-sigma relation by Tundo et al. (2007) (dark and light grey)



The **AGN bolometric luminosity functions** predicted by our model (red) when the mass accreted during the *quasar mode* at high redshift has been increased according to:

$$f_{BH} = 0.01\cdot\log\left(\frac{M_{BH}}{10^3 M_\Theta}+1\right)\cdot z \qquad z>1.5, M_{BH}>10^6 M_\Theta$$

$$\Delta M_{BH,Q} = 0.01\cdot m_{cold} \qquad z>6$$

Grey dots show several observated AGN luminosity functions, converted to bolometric ones by Hopkins et al. (2007)

Results

The model is able to reproduce the observed **BH scaling relations** over the whole range of BH masses and galaxy properties probed by observations. The intrinsic scatter in the model is significantly larger than in the data, a mismatch that can be accounted for by adopting the observational selection criteria to obtain a mock BH catalogue with similar characteristics as the observed one

Our model predictions also match the **BH fundamental plane** relation derived by Hopkins et al. (2007), and successfully predict very little evolution of this plane, at least at $z < 3$

The model **BH mass function** is in good agreement with the observed one within the mass range accessible by observations, but in the range 1e7-1e9 Msun, in which the number density of model BHs lies slightly below the observations

The **AGN luminosity function** (LF) is systematically underestimated by assuming that BHs accrete mass with a constant Eddington factor f_{Edd} =1: at high z the model matches the faint-end of the LF but underpredicts the number density of the brightest objects, while the situation is reversed at z~0. Reducing the Eddington ratio alleviates the faint-end mismatch but amplifies the bright-end discrepancy at high redshifts. A significant improvement at low redshifts is obtained when the Eddington-limited growth of the BH is followed by a long quiescent phase with lower Eddington ratios, as suggested by Hopkins et al. (2005). In this case our model is able to match the observed AGN LF in the interval 0.1<z<1. However, our predicted number density of bright AGN is still biased low at z>1

We were not able to eliminate this mismatch by simply modifying the accretion efficiency, the Eddington factor or the BH seed mass (when considered in physically plausible ranges). One possible way out is to increase the mass fraction accreted during the *quasar mode* at high redshift