Mon. Not. R. Astron. Soc. **382,** 903–914 (2007) doi:10.1111/j.1365-2966.2007.12449.x

Reproducing the assembly of massive galaxies within the hierarchical cosmogony

Fabio Fontanot,[1,2⋆] Pierluigi Monaco,[2,3⋆] Laura Silva[3⋆] and Andrea Grazian[4⋆]

[1]*Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*Dipartimento di Astronomia, Università di Trieste, via Tiepolo 11, 34131 Trieste, Italy*
[3]*INAF–Osservatorio Astronomico di Trieste, via Tiepolo 11, 34131 Trieste, Italy*
[4]*INAF–Osservatorio Astronomico di Roma, via Frascati 33, I-00040 Monteporzio, Italy*

Accepted 2007 September 10. Received 2007 August 22; in original form 2007 June 27

ABSTRACT

In order to gain insight into the physical mechanisms leading to the formation of stars and their assembly in galaxies, we compare the predictions of the MOdel for the Rise of GAlaxies aNd Active nuclei (MORGANA) to the properties of *K*- and 850-μm-selected galaxies (such as number counts, redshift distributions and luminosity functions) by combining MORGANA with the spectrophotometric model GRASIL. We find that it is possible to reproduce the *K*- and 850-μm-band data sets at the same time and with a standard Salpeter initial mass function, and ascribe this success to our improved modelling of cooling in DM haloes. We then predict that massively star-forming discs are common at $z \sim 2$ and dominate the star formation rate, but most of them merge with other galaxies within $\sim$100 Myr. Our preferred model produces an overabundance of bright galaxies at $z < 1$; this overabundance might be connected to the build-up of the diffuse stellar component in galaxy clusters, as suggested by Monaco et al., but a naive implementation of the mechanism suggested in that paper does not produce a sufficient slowdown of the evolution of these objects. Moreover, our model overpredicts the number of 10^{10}–10^{11} $M_{\odot}$ galaxies at $z \sim 1$; this is a common behaviour of theoretical models as shown by Fontana et al.. These findings show that, while the overall build-up of the stellar mass is correctly reproduced by galaxy formation models, the 'downsizing' trend of galaxies is not fully reproduced yet. This hints to some missing feedback mechanism in order to reproduce at the same time the formation of both the massive and the small galaxies.

Key words: galaxies: evolution – galaxies: formation.

1 INTRODUCTION

The Lambda cold dark matter (ΛCDM) cosmology is consistent with a large body of observations of the large-scale Universe (see e.g. Spergel et al. 2006). Then, the predicted hierarchical evolution of dark matter (DM) perturbations subject to gravitational instability provides a standard framework to study the formation and evolution of luminous structures in the Universe. However, while the cosmological framework is fixed with a small uncertainty, several open questions, regarding to the formation and evolution of galaxies, arise from the complex evolution of baryons within the potential wells of the DM haloes.

Galaxy formation is observationally constrained by many multi-wavelength surveys of deep fields (e.g. COMBO17, Wolf, Meisenheimer & Röser 2001; DEEP2, Davis et al. 2003; GOODS, Giavalisco et al. 2004; GEMS, Rix et al. 2004; UKIDSS, Lawrence et al. 2007; COSMOS, Scoville et al. 2007; etc.). To compare with these data sets, many thousands of galaxies must be generated by models. This makes a straightforward numerical approach problematic (it has been attempted, e.g., by Nagamine et al. 2005; Saro et al. 2006; Robertson et al. 2007), so that simpler and quicker models have been developed, based on sets of recipes that address the various processes involved. These 'semi-analytical' models (see e.g. Somerville, Primack & Faber 2001; Granato et al. 2004; Menci et al. 2004; Somerville et al. 2004; Baugh et al. 2005; Kang et al. 2005; Bower et al. 2006; Cattaneo et al. 2006; Croton et al. 2006; De Lucia et al. 2006; Monaco, Fontanot & Taffoni 2007) have been tested against an impressive number of observational constraints. In this decade-long testing process, many specific models have been unsuccessful in reproducing constraints like the high-mass cut-off of the luminosity function (LF) (Benson et al. 2003), the level of α-enhancement in elliptical galaxies (Nagashima et al. 2005;

⋆E-mail: fontanot@mpia-hd.mpg.de (FF); monaco@oats.inaf.it (PM); silva@oats.inaf.it (LS); grazian@mporzio.astro.it (AG)

Thomas et al. 2005), the redshift distribution of *K*-band sources (Cimatti et al. 2002a), the surface density of extremely red objects (EROs) (Cimatti et al. 2002b; Daddi et al. 2002; Smith et al. 2001).

Many of these difficulties have been overcome by later versions of the models, but some of them have required strong assumptions: for instance, a top-heavy initial mass function (IMF) in starbursts is required by Baugh et al. (2005) to reproduce the submillimetre counts. In the most recent models (see i.e. Bower et al. 2006; Croton et al. 2006) the cut-off of the galaxy LF and the bimodality of galactic colours is obtained only if cooling flows in large haloes are quenched by active galactic nucleus (AGN) feedback, i.e. by the energy emerging from massive black holes accreting at a relatively low rate (in the so-called *radio mode*).

A recently highlighted observational trend of galaxy formation is the so-called *downsizing* of galaxies: at variance with the hierarchical trend of DM haloes, more massive galaxies tend to form their stars earlier and in a shorter period than smaller galaxies, which experience more prolonged star formation histories. While this general trend is known to be not incompatible with cosmology once stellar and AGN feedback are properly taken into account (see e.g. Granato et al. 2004; Neisteinn, van den Bosch & Dekel 2006), recent observations show that the details, like for instance the (almost) parallel evolution of the star formation density for galaxies of different mass (Zheng et al. 2007), are still not reproduced.

Massive elliptical galaxies have always been remarkably elusive objects in this regard. The first versions of hierarchical galaxy formation models (see e.g. White 1996) predicted that these galaxies form late ($z \sim 1$) by the merging of already assembled discs, while evidence from stellar populations (see Matteucci 1996, for a review), the tightness of the fundamental plane (Renzini & Ciotti 1993), the evolution of the colour–magnitude relation (Kodama et al. 1998; Blakeslee et al. 2003; Ellis et al. 2006) and the local Mg_2–σ relation (Bernardi et al. 2003) suggested that they formed early ($z > 2$) in a short burst of star formation. Clearly, in a hierarchical universe the age of stars does not need to coincide with the assembly age of the galaxy, defined as the time at which the most massive progenitor has at least half of the final stellar mass. Massive ellipticals could then be assembled late by *dry mergers* of other ellipticals, so as to preserve the oldness of their stellar populations while producing a low assembly redshift (De Lucia et al. 2006). This possibility is severely constrained by the modest (if any) evolution of the high end of the stellar mass function since $z = 1$ (Cimatti, Daddi & Renzini 2006): while models predict a doubling of stellar masses (De Lucia & Blaizot 2007), evidence excludes an evolution larger than $\sim$0.2 dex (as roughly estimated by Monaco et al. 2006; see the references therein). Recently, Monaco et al. (2006; see also Conroy, Weschler & Kravtsov 2007) have proposed that the formation of a diffuse stellar component in galaxy clusters by scattering of stars during dry mergers (see e.g. Murante et al. 2007) may conspire to decrease the expected evolution of the high end of the stellar mass function.

The early formation of massive galaxies and their following (almost passive) evolution are best constrained by deep observations in the submillimetre band, suited to reveal obscured star formation events at high redshift, and in the *K* band, suited to probe stellar masses at lower redshift.

The warm dust present in star-forming clouds absorbs most of the ultraviolet (UV)/blue photons emitted by young stars and reprocesses them to the FIR, becoming the dominant contributor in that band. Moreover, the steeply decreasing shape of the galactic spectral energy distributions (SEDs) from $\sim$100 μm to $\sim$1 mm gives a negative *K*-correction that promotes the observation of starbursts in the submillimetre bands up to $z \sim 5$. Therefore strong starbursts at high redshift are more easily observed in the submillimetre than in the optical. The submillimetre emission is measurable in a few windows, most notably that at 850 μm. Observations with the Submillimetre Common-User Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope have highlighted the presence of a population of high-redshift massive starbursts (see e.g. Smail, Ivison & Blain 1997; Hughes et al. 1998; Chapman et al. 2002; Scott et al. 2002) commonly interpreted as galaxies forming stars at rates of hundreds if not thousands of $M_\odot\ yr^{-1}$. Given the poor angular resolution of SCUBA images, the identification of optical counterparts is difficult, and can be achieved using interferometric images at longer wavelengths. The resulting redshift distribution is thought to peak at $z \sim 2.4$ (Chapman et al. 2003, 2005).

The *K* band is a very good tracer of the stellar mass at $z \lesssim 1.5$ (Gavazzi, Pierini & Boselli 1996), it is almost unaffected by dust extinction, and requires small *K*-corrections that weakly depend on the morphological type, so it is ideal to follow the assembly of the bulk of stellar mass. The local *K*-band LF has been measured with great accuracy by the Two Micron All sky Survey (2MASS) collaboration (Cole et al. 2001; Kochanek et al. 2001). In this paper we will focus on four *K*-band surveys with (photometric or spectroscopic) redshift coverage. (i) The K20 survey (Cimatti et al. 2002a; Pozzetti et al. 2003) is a $K < 20$ (corresponding to $K_{AB} < 21.84$) limited sample, covering 52 arcmin^2, with a very high-redshift completeness (>90 per cent). (ii) The GOODS-MUSIC (Grazian et al. 2006) catalogue is a multicolour sample extracted from the deep and wide survey conducted over the *Chandra Deep Field*-South in the framework of the GOODS project. Here we use the $K_{AB} < 23.5$ limited galaxy sample defined in Fontana et al. (2006). This sample covers 143.2 arcmin^2; 28 per cent of the galaxies have a spectroscopic redshifts, used to train photometric redshift estimates for all other galaxies. (iii) The UKIDSS Ultra Deep Survey (UDS, Dye et al. 2006) galaxy sample is a complete catalogue of $K_{AB} < 22.5$ selected galaxies over 0.6 deg^2; for each object in the sample a photometric estimate of the redshift is provided. (iv) The VIMOS-VLT Deep Survey (VVDS) (Le Fevre et al. 2005) is a spectroscopic survey designed to measure redshift for $\sim 10^5$ sources selected, nearly randomly from a photometric catalogue. Here we consider a *K*-selected sample, with either photometric and spectroscopic redshifts obtained by Pozzetti et al. (2007) combining a $K_{AB} < 22.34$ limited sample (20 per cent redshift completeness) defined over a 442 arcmin^2 area, with a $K_{AB} < 22.84$ limited sample (29 per cent redshift completeness) defined over a 172 arcmin^2 area.

This paper is the third of a series devoted to describe the MOdel for the Rise of GAlaxies aNd Active nuclei (MORGANA). With respect to similar models of galaxy formation, MORGANA presents a different, more sophisticated treatment of the mass and energy flows between galactic phases (cold and hot gas, stars) and components (bulge, disc, halo). In particular, the process of radiative cooling of the shocked gas is treated with a new model (tested against simulations in Viola et al. 2007), while feedback is inserted following the model by Monaco (2004a) and galaxy winds and superwinds are allowed. The model is described in detail in Monaco, Fontanot & Taffoni (2007; hereafter Paper I). The prediction of the properties of the AGN population is presented in Fontanot et al. (2006; hereafter Paper II), and the prediction of the evolution of the stellar mass function, together with those of other similar models, has been compared to the results inferred from the GOODS-MUSIC data in Fontana et al. (2006). As mentioned above, MORGANA has been used by Monaco et al. (2006) to address the lack of evolution of the high end of the stellar mass function and its connection with the building of the diffuse stellar component in galaxy clusters. In this paper we

compare the predictions of MORGANA to the data mentioned above of deep fields in submillimetre (850 μm) and *K* bands to test to what extent the model is able to reproduce the formation and assembly of massive galaxies. To this aim, we have combined MORGANA with the spectrophotometric code GRASIL (Silva et al. 1998) that computes the UV to radio SEDs of model galaxies, including a three-dimensional bulge + disc geometry with a two-phase interstellar medium (ISM), the radiative transfer through the dusty ISM, a realistic dust grain model, and a direct computation of the dust temperature distribution.

The paper is organized as follows. In Section 2.1 we describe the main properties of the MORGANA model. In Section 2.2 we describe how we compute LFs, number counts and redshift distributions interfacing the output of the model with GRASIL. In Section 3 we present our results, a discussion is given in Section 4, and in Section 5 we give our conclusions. Throughout this work we assume, whenever necessary, the concordance cosmological model $\Omega_\Lambda = 0.7, \Omega_\mathrm{m} = 0.3, H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}, \sigma_8 = 0.9$.

2 MODEL

2.1 Galaxy formation model: MORGANA

MORGANA is described in full detail in Paper I, while AGN accretion is described in Paper II. We give here only a brief description, aimed at highlighting the main processes included in the model.

2.1.1 Algorithm

MORGANA follows the typical scheme of semi-analytic models, with some important differences. Each DM halo contains one galaxy for each progenitor[1]; the galaxy associated with the main progenitor is the central galaxy. Baryons in a DM halo are divided into three components, namely a halo, a bulge and a disc. Each component contains three phases, namely cold gas, hot gas and stars. For each component the code follows the evolution of its mass, metal content, thermal energy of the hot phase and kinetic energy of the cold phase.

The main processes included in the model are the following.

(i) The merger trees of DM haloes are obtained using the pinocchio tool (Monaco, Theuns & Taffoni 2002; Taffoni, Monaco & Theuns 2002).

(ii) After a merging of DM haloes, dynamical friction, tidal stripping and tidal shocks on the satellite (the smaller DM halo, with its galaxy at the core) lead to a merger with the central galaxy or to tidal destruction as described by Taffoni et al. (2003). At each merger a fraction of the satellite stars is scattered to the stellar halo component (Monaco et al. 2006; Murante et al. 2007). These stars are not associated to galaxies but to the intracluster light.

(iii) The evolution of the baryonic components is performed by numerically integrating a system of equations for all the mass, energy and metal flows; this allows not to be restricted to linear dynamics.

(iv) The intergalactic medium infalling on a DM halo is shock-heated, as well as the hot halo component of merging satellites (which is given to the main halo) and that of the main halo in case of major merger ($M_\mathrm{sat} > 0.2 M_\mathrm{tot}$). Following Wu, Fabian & Nulsen (2001), shock-heating is implemented by assigning to the infalling gas a specific thermal energy equal to 1.2 times the specific virial energy, $-0.5 U_\mathrm{H}/M_\mathrm{H}$ (where M_H and U_H are the mass and binding energy of the DM halo).

(v) The profile of the hot halo gas is computed at each time-step by solving the equation for hydrostatic equilibrium with a polytrophic equation of state and an assumed polytrophic index $\gamma_\mathrm{p} = 1.2$. No hot gas is present within a cooling radius r_cool, which is set to a vanishingly small value at major mergers.

(vi) The cooling flow is computed by integrating the contribution to radiative cooling of each spherical shell, taking into account the heating from (stellar and AGN) feedback from galaxies. Given the importance of cooling for the results presented in this paper, we give here some detail of the cooling model. If T_{g0} and ρ_{g0} are the temperature and density of the hot halo gas extrapolated to $r = 0$, $\mu_\mathrm{hot}\, m_\mathrm{p}$ its mean molecular weight and $r_\mathrm{s} = r_\mathrm{H}/c_\mathrm{nfw}$ the scale radius of the halo (of radius r_H and concentration c_nfw), then the mass cooling flow $\dot{M}_\mathrm{co,H}$ results:

$$\dot{M}_\mathrm{co,H} = \frac{4\pi r_\mathrm{s}^3 \rho_{g0}}{t_\mathrm{cool,0}} \mathcal{I}\left(\frac{2}{\gamma_\mathrm{p}-1}\right), \tag{1}$$

where the integral $\mathcal{I}(\alpha)$ is defined as $\int_{r_\mathrm{cool}/r_\mathrm{s}}^{c_\mathrm{nfw}} \{1 - a[1 - \ln(1+t)/t]^\alpha t^2\}\,\mathrm{d}t$, with $a = [3T_\mathrm{vir}(\gamma_\mathrm{p}-1)c_\mathrm{nfw}(1+c_\mathrm{nfw})]/\{\gamma_\mathrm{p} T_{g0}\,[(1+c_\mathrm{nfw})\ln(1+c_\mathrm{nfw}) - c_\mathrm{nfw}]\}$ (T_vir being the virial temperature of the halo). The cooling time $t_\mathrm{cool,0}$ is computed using the central density (the density gradient is taken into account by the integral $\mathcal{I}$) and the temperature at r_cool (the temperature gradient is neglected):

$$t_\mathrm{cool,0} = \frac{3kT_\mathrm{g}(r_\mathrm{cool})\mu_\mathrm{hot} m_\mathrm{p}}{2\rho_{g0}(\Lambda_\mathrm{cool} - \Gamma_\mathrm{heat})}. \tag{2}$$

Here Λ_cool is the metal-dependent Sutherland & Dopita (1993) cooling function and the heating term Γ_heat is computed assuming that the energy flow $\dot{E}_\mathrm{hw,H}$ fed back by the galaxy is given to the cooling shell:

$$\Gamma_\mathrm{heat} = \frac{\dot{E}_\mathrm{hw,H}}{4\pi r_\mathrm{s}^3 \mathcal{I}(2/(\gamma_\mathrm{p}-1))}\left(\frac{\mu_\mathrm{hot} m_\mathrm{p}}{\rho_{g0}}\right)^2. \tag{3}$$

Whenever $\Gamma_\mathrm{heat} > \Lambda_\mathrm{cool}$ the cooling flow is quenched. The cooling radius r_cool is treated as a dynamical variable whose evolution takes into account the hot gas injected by the central galaxy ($\dot{M}_\mathrm{hw,H}$):

$$\dot{r}_\mathrm{cool} = \frac{\dot{M}_\mathrm{co,H} - \dot{M}_\mathrm{hw,H}}{4\pi\rho_\mathrm{g}(r_\mathrm{cool}) r_\mathrm{cool}^2}. \tag{4}$$

This equation is valid if pressure is balanced at r_cool, an assumption which clearly does not hold in general. We mimic the pressure force acting on mass shell at r_cool as follows:

$$\dot{r}'_\mathrm{cool} = \dot{r}_\mathrm{cool} - c_\mathrm{s}, \tag{5}$$

where c_s is the sound speed computed at r_cool. This recipe is different from that used in most other semi-analytical models, where a cooling radius is computed by inverting the cooling time as a function of radius. Viola et al. (2007) have compared analytic cooling models to *N*-body + hydro simulations showing that, while the 'classical' cooling model significantly underestimates the amount of cooled mass, the present cooling model gives a very good fit.

(vii) When the hot halo phase is heated by feedback beyond the virial temperature, it can leave the DM halo in a galactic superwind. A similar thing happens to the cold halo gas when it is accelerated by stellar feedback. To compute the time at which the ejected gas falls back into a DM halo, its merger history is scrolled forward in time until the circular velocity is larger than the (sound or kinetic) velocity of the gas at the ejection time.

[1] Each DM halo forms through the merging of many haloes of smaller mass, called progenitors. At each merging the largest halo survives (it retains its identity), the others become substructure of the largest one. The main progenitor is the one that survives all the mergings. The mass resolution of the box used for computing the merger trees sets the smallest progenitor mass, as explained in Section 2.1.2.

(viii) The cooling gas is let infall on the central galaxy on a dynamical time-scale (computed at $r_{\rm cool}$). It is divided between disc and bulge according to the fraction of the disc that lies within the half-mass radius of the bulge.

(ix) The gas infalling on the disc keeps its angular momentum; disc sizes are computed with an extension of the Mo, Mao & White (1998) model that includes the contribution of the bulge to the disc rotation curve.

(x) Disc instabilities and major mergers of galaxies lead to the formation of bulges. We also take into account a possible disc instability driven by feedback. In minor mergers the satellite mass is given to the bulge component of the larger galaxy.

(xi) Star formation and feedback in bulges and discs are inserted following the model of Monaco (2004a). According to that model, the regime of stellar feedback in a galaxy depends mainly on the density and vertical scalelength of the galactic system. In thin systems, like spiral discs with gas surface density $\Sigma_{\rm cold,D}$ and fraction of cold gas $f_{\rm cold,D}$, the time-scale for star formation $t_{*,\rm D}$ is predicted to be

$$t_{\star,\rm D} = 9.1 \left(\frac{\Sigma_{\rm cold,D}}{1\,{\rm M}_\odot\,{\rm pc}^{-2}}\right)^{-0.73} \left(\frac{f_{\rm cold,D}}{0.1}\right)^{0.45} {\rm Gyr}. \quad (6)$$

Due to the correlation of $f_{\rm cold,D}$ and $\Sigma_{\rm cold,D}$ (galaxies with higher gas surface density consume more gas), this relation is compatible with the Schmidt law. Thick systems like star-forming bulges (or mergers) are dominated by transients which are very difficult to model, so the straightforward Schmidt law is used:

$$t_{\star,\rm B} = 4 \left(\frac{\Sigma_{\rm cold,B}}{1\,{\rm M}_\odot\,{\rm pc}^{-2}}\right)^{-0.4} {\rm Gyr}. \quad (7)$$

In both cases, hot gas is ejected to the halo (in a hot galactic wind) at a rate equal to the star formation rate (as predicted by Monaco 2004a), though massive bulges with circular velocity $V_{\rm B} \gtrsim 300\,{\rm km\,s}^{-1}$ are able to bind their hot phase component. The thermal energy of this reheated gas is not scaled to the DM halo circular velocity but to the energy of exploding supernovae (SNe).

(xii) In star-forming bulges cold gas is ejected in a cold galactic wind by kinetic feedback due to the predicted high level of turbulence driven by SNe (this process is described in full detail in Paper II, Section 2.2).

(xiii) Accretion of gas on to massive black holes (starting from small seeds present in all galaxies) is connected to the ability of cold gas to loose angular momentum by some (unspecified) mechanism driven by star formation. This is explained in full detail in Paper II.

(xiv) Because accretion on to black holes is triggered by star formation, AGN feedback can quench cooling flows only after their start. Alternatively, the cooling flow can be quenched when a fiducial energy criterion is met, without letting any gas fall to the galaxy and form stars. This 'forced quenching' procedure has been used in Paper II and will be discussed in a forthcoming paper.

(xv) Metal enrichment is self-consistently modelled in the instantaneous recycling approximation.

2.1.2 Runs

All models are based on the same pinocchio run we introduced in Paper I, a 512^3 realization of a 150 Mpc comoving box ($h = 0.7$). The mass particles is $1.0 \times 10^9\,{\rm M}_\odot$, and the smallest halo we consider contains 50 particles, for a mass of $5.1 \times 10^{10}\,{\rm M}_\odot$. The branches of the DM halo merger trees are tracked starting from a mass of 10 particles, corresponding to $1.0 \times 10^{10}\,{\rm M}_\odot$; this is the mass of the smallest progenitors. For sake of comparison, the corresponding values for the Millennium Simulation (Springel et al. 2005) are 1.2×10^9 for the particle mass and 2.5×10^{10} for the smallest progenitor. We have tested the overall stability of our results by running the model over two other 512^3 boxes of size 200 and 100 Mpc (see Paper I for more details).

The stellar mass of the typical galaxy contained in the smallest DM halo at $z = 0$ is $3 \times 10^8\,{\rm M}_\odot$. In order to estimate the completeness limit for the stellar mass function we consider an higher resolution pinocchio run (512^3 realization of a 100 Mpc comoving box, see also Paper I, appendix B, for a discussion about the numerical stability of the model). We study the typical stellar mass of the galaxies belonging to 2.5–5.1 $\times\ 10^{10}\,{\rm M}_\odot$ DM haloes and we estimate a completeness limit of $4.5 \times 10^8\,{\rm M}_\odot$. For each run we compute the evolution of (up to) 300 trees (i.e. DM haloes at $z = 0$) per logarithmic bin of halo mass of width 0.5 dex. This implies that while all the most massive haloes are considered, smaller haloes are randomly sparse-sampled. To properly reconstruct the statistical properties of galaxies we assign to each tree a weight $w_{\rm tree}$ equal to the inverse of the fraction of selected DM haloes in the mass bin.

A standard parameter choice was presented in Paper I. However, it was noticed there that spiral discs tend to be more compact than in the real Universe. A Schmidt–Kennicutt-like law (equation 6) is used to compute star formation rates, so this results in higher surface densities, shorter star formation time-scales and lower gas fractions at $z = 0$. Forcing DM halo concentrations to lower values allows us to alleviate this problem, while reproducing the zero-point of the Tully–Fisher relation and preserving a good fit of the Schmidt–Kennicutt law. We then scale concentrations so as to take a given value for a $10^{12}\,{\rm M}_\odot$ DM halo at $z = 0$. The value was proposed to be 4 in Paper I; however, we noticed that this choice tends to lower the stellar mass function at the knee, so we use the slightly higher value of 6 as a good compromise. Moreover, following Monaco et al. (2006) we scatter to the halo stellar component 40 per cent of the stellar mass of satellites at each galaxy merger. Finally, we implement AGN feedback by adopting the 'forced quenching' procedure of Paper II; this choice has little influence on the results presented here, and will be discussed in a forthcoming paper. Also, we have checked that the inclusion of quasar-triggered galaxy winds, necessary to reproduce the accretion history of massive black holes (Paper II), does not change the qualitative behaviour of the models.

These changes of parameters do not influence in any way the conclusions drawn in this paper. In fact, we are not proposing a 'best-fitting' model of the galaxy population; on the contrary, we will show that MORGANA is able to reproduce many observables, but its agreement with data breaks as soon as one tries to reproduce some relevant aspects of the 'downsizing' trend, so that a global fine-tuning of the model is not possible. The most important point of such an investigation is to understand the limits of these models and the origin of the disagreement with data. Fine-tuning to fit specific data sets, like the local LFs, will be required in other contexts, for instance to create galaxy mock catalogues.

2.2 SED model: GRASIL

For each galaxy modelled by MORGANA we compute the corresponding UV to radio SED with GRASIL. The details of the code are given in Silva et al. (1998) (and the subsequent updates and improvements in Silva 1999; Granato et al. 2000; Bressan, Silva & Granato 2002; Panuzzo et al. 2003; Vega et al. 2005), while we summarize here the main features. (i) Stars and dust are distributed in a bulge (King profile) + disc (radial and vertical exponential profiles) axisymmetric

Table 1. Adopted values for the GRASIL parameters that are not provided by MORGANA.

Parameters	Values
$t_{\rm esc}$	10^7 yr
$f_{\rm MC}$	0.5
$M_{\rm MC}/r^2_{\rm MC}$	$10^6\,{\rm M}_\odot/(16\,{\rm pc})^2$
$h^*_{\rm d}/r^*_{\rm d}$	0.1
$h^d_{\rm d}/r^d_{\rm d}$	0.1

geometry. (ii) We consider the clumping of both (young) stars and dust, through a two-phase ISM with dense giant molecular clouds (MCs) embedded in a diffuse ('cirrus') phase. (iii) The stars are assumed to be born within the optically thick MCs and gradually to escape from them as they get older on a time-scale $t_{\rm esc}$. This gives rise to an age-selective extinction, with the youngest and most luminous stars suffering larger dust extinction than older ones. (iv) The dust consists of graphite and silicate grains with a distribution of grain sizes, and polycyclic aromatic hydrocarbon (PAH) molecules. In each point of the galaxy and for each grain type the appropriate temperature is computed (either equilibrium T for big grains or probability distribution of temperature for small grains and PAHs). The detailed PAH emission spectrum has been updated in Vega et al. (2005) based on the Li & Draine (2001) model. (v) The radiative transfer of starlight through the dust distribution is computed yielding the emerging SED. The simple stellar population (SSP) library (Bressan, Granato & Silva 1998; Bressan et al. 2002) includes the effect of the dusty envelopes around asymptotic giant branch (AGB) stars, and the radio emission from synchrotron radiation and from ionized gas in H II regions.

The GRASIL parameters that are not provided by MORGANA are summarized in Table 1, and are described in the following. (i) The escape time-scale of young stars for the parent MCs is set to $t_{\rm esc} = 10^7$ yr. This is intermediate between the values found by Silva et al. (1998) to well describe the SED of spirals ($\sim$ a few Myr) and starbursts ($\sim$ a few 10 Myr), and is of the order of the estimated destruction time-scale of MCs by massive stars (e.g. Monaco, 2004b). (ii) The gas mass predicted by MORGANA is subdivided between the dense and diffuse phases, fixing the fraction of gas in MCs $f_{\rm MC}$ to 0.5. The resulting SEDs are not much sensitive to this choice. (iii) The mass of dust is obtained by the gas mass and the dust to gas mass ratio δ which is set to evolve linearly with the metallicity ($\delta = 0.45\,Z$). (iv) The optical depth of MCs, driving their spectrum, depends on the mass and radius set for MCs through $\tau \propto \delta\, M_{\rm MC}/r^2_{\rm MC}$, with $M_{\rm MC} = 10^6\,{\rm M}_\odot$, $r_{\rm MC} = 16$ pc; (iv) the bulge and disc scale radii for stars and gas are given by MORGANA, while the disc scale-heights, h^*_d and $h^d_{\rm d}$ for stars and dust, respectively, are set to 0.1 the corresponding scale radii. In addition, the dust grain size distribution and abundances are set to match the mean Galactic extinction curve and emissivity (as in Silva et al. 1998 and Vega et al. 2005) and are not varied here.

2.3 Interfacing MORGANA with GRASIL

The output of MORGANA consists, for each galaxy, in a time sampling of the main dynamical variables of the model: for each component (halo, bulge, disc) the code issues mass, kinetic energy and metal mass of cold gas; mass, thermal energy and metal mass of hot gas; mass and metals of stars; average (over the time bin) and punctual (at the end of the time bin) star formation rate (these are given only for bulge and disc). For each galaxy it gives also expelled mass and metals, black hole mass, cooling radius, bulge and disc radii and velocities, punctual values of the accretion rate on to the black hole. Star formation histories are reconstructed following all the exchanges of stellar matter between galaxies and galaxy components, so that they refer to the stars contained in the component at a given time and not to the stars formed in that component.

Of this information GRASIL uses the star formation and cold gas metallicity history for bulge and disc, the mass of gas in the two components at the required galactic age, and the stars and gas scale radii at the same age. The time sampling of these quantities is 0.1 Gyr. GRASIL resamples the star formation histories on a much finer time grid to optimally account for the short lifetimes of massive stars. However, the width of the time bin of MORGANA is rather large, so the number of massive stars can be rather inaccurate. We then split the last time bin into a sub-bin of 10 Myr, to which we assign the punctual value of the star formation rate at the end of the bin, and a larger, earlier one of 90 Myr.[2] to which we assign a star formation rate such that the integral in the two sub-bins gives the correct final amount of stars. Using test star formation histories, we have verified that this sampling allows us to reproduce the magnitudes to within $<$0.1 mag; indeed, even the strongest star formation events do not have associated time-scales much smaller than the time bin.

It is worth mentioning that running the spectrophotometric code on model galaxies is the main bottleneck of the computation; it is then necessary to devise strategies to optimize the computations by estimating the minimal number of galaxies needed to have a reliable result.

2.4 Luminosity functions, number counts and redshift distributions

To simulate a deep field, the computation of galaxy SEDs is performed on the box at several output times. As explained in Paper I, the sparse-sampling of trees (Section 2.1.2) results in an oversampling of small satellites with respect to central galaxies of similar stellar mass. We correct for this oversampling by further sparse-sampling the satellites as follows. First, we construct from the results of the run at $z = 0$ an average curve of mass of the central galaxy as a function of DM halo mass; secondly, we randomly sparse-sample the satellites with a probability equal to the ratio between the weights $w_{\rm tree}$ of the tree the satellite belongs to and that of the DM halo whose central galaxy has on average the same mass as the satellite. The inverse of this probability is a new weight, $w_{\rm sat}$. Central galaxies are all selected and assigned a unity weight. If a galaxy is destroyed by mergers (or tides) then its stellar mass at the destruction time is used to compute $w_{\rm sat}$. This procedure gives a roughly constant number of galaxies in logarithmic intervals of mass. Because each galaxy is present in many time bins, the number of selected galaxies is typically very high. To limit this number we introduce a third weight $w_{\rm gal}$ as the inverse of a further sampling factor, equal for all galaxies. The first two samplings (of merger trees and galaxies) are both computed at $z = 0$, but the weights $w_{\rm tree}$ and $w_{\rm sat}$ assigned to the galaxies are used at any redshift. This is done with no loss of generality, as a fair reconstruction of LFs or number counts only requires that the weights are used consistently with the sparse-sampling procedure; in other words, we only need to require that the properties of a sparsely sampled population are

[2] As the age of Universe in the assumed cosmology is 13.47 Gyr, the time bin corresponding to $z = 0$ is smaller than 0.1 Gyr; in this case we make the earlier bin consistently smaller than 90 Myr.

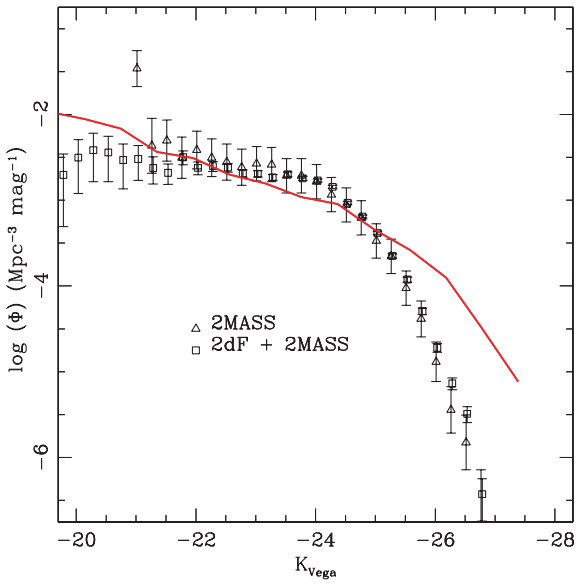

Figure 1. *K*-band LF. Data refer to the observations of Cole et al. (2001) and Kochanek et al. (2001). Solid line refers to MORGANA prediction.

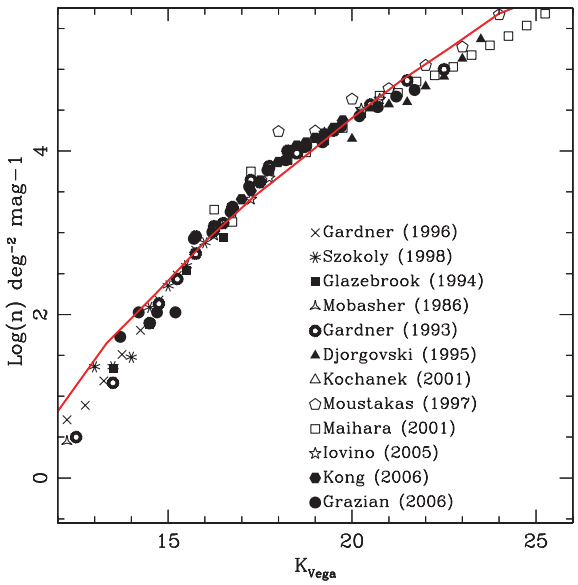

Figure 2. Source number counts in the *K* band. Data refer to observations as listed in the figure. Solid line refers to MORGANA prediction.

weighted by the inverse of the sampling probability, whatever the redshift at which the sampling is performed.

In order to ensure a smooth description of the redshift evolution of the properties of our simulated galaxies, we use the following procedure. At every integration time bin, corresponding to a redshift $z(t)$, at the end of integration we apply the three sampling procedures to all the galaxies in the box and select a subsample. It is worth to notice that the random sparse sampling ensures that a different subsample of galaxies is considered at each $z(t)$, so that galaxies at different redshifts are not always the same galaxies seen at different times. We define a fourth weight $w_{\rm time}$ as the ratio between the width of the time sampling (0.1 Gyr) and the time interval span by $l_{\rm box}$ at $z(t)$. We also compute the angle subtended by a square of comoving side $l_{\rm box}$ at $z(t)$. We then compute the SEDs of model galaxies using GRASIL and we use them to estimate absolute and apparent magnitudes (both in the Vega and AB system) and infrared fluxes. We collect the information in the catalogues.

LFs, number counts, and redshift distributions of magnitude-limited samples are then computed by performing weighted sums over the galaxies using the product of the four weights defined above. LFs at high redshift are computed over the same redshift intervals as in the observations, while number counts and redshift distributions of magnitude-limited samples are computed using galaxies starting form $z \simeq 6$.

3 RESULTS

3.1 *K* band

The resulting *K*-band LF at $z = 0$ (in the Vega system) is compared in Fig. 1 to that obtained using the local 2MASS sample (Jarrett et al. 2000), limited to $K_{\rm Vega} < 13.5$ (Kochanek et al. 2001), and the combination of the 2MASS and 2dF samples (Cole et al. 2001). Being the *K*-band luminosity a good tracer of stellar mass, this comparison is analogous to that of fig. 7 of Paper I, where the stellar mass function was compared to that inferred by data (obtained also with the same 2MASS + 2dF sample). In agreement with that result, the $z = 0$ *K*-band LF shows an overestimate of both the bright (massive) end and the faint (low-mass) end of the LF. In particular, the overestimate of the high tail is much more relevant here, with respect to the estimated mass function. We find that this is in part due to a discrepancy in the adopted M_*/L_K ratios when passing from the LF to the mass function. The average M_*/L_K ratio of the MORGANA + GRASIL model galaxies is $\sim$0.7 $\mathrm{M}_\odot/\mathrm{L}_{K\odot}$, while for instance Cole et al. (2001) adopt an higher average value, $M_*/L_K = 1.32\,\mathrm{M}_\odot/\mathrm{L}_{K\odot}$, typical of very old stellar populations (see e.g. fig. 24 in Maraston 2005). In spite of these discrepancies, the model is able to reproduce correctly the overall normalization of the LF.

Fig. 2 shows the predicted *K*-band number counts compared to data available in the literature (Mobasher, Ellis & Sharples 1986; Gardner, Cowie & Wainscoat 1993; Glazebrook et al. 1994; Djorgovski et al. 1995; Gardner et al. 1996; Moustakas et al. 1997; Szokoly et al. 1998; Kochanek et al. 2001; Maihara et al. 2001; Iovino et al. 2005; Grazian et al. 2006; Kong et al. 2006). The models fit well the data in the range $15 \lesssim K_{\rm Vega} \lesssim 22$. The excess at bright fluxes is a signature of the overestimate of the LF at the bright end, while the excess at faint fluxes is commented below.

Number counts do not strongly constrain the model unless redshift distributions of magnitude-limited samples are available. We then compare our results with redshift distributions obtained from the K20 (Fig. 3, Vega system), GOODS-MUSIC (Fig. 4, AB magnitudes) and VVDS (Fig. 5, AB magnitudes) catalogues. For sake of clarity we show both the differential and the cumulative distributions. We compute the error on the cumulative distributions using a bootstrap technique based on 1000 mock redshift catalogues drawn using the observed redshift distribution. The agreement of MORGANA prediction with K20 observations (relative to $K_{\rm AB} < 21.84$) is good: we predict a total number of 435 objects against 480 ± 22 observed. The position of the peak of the distribution and the long tail of galaxies at $z > 1.5$ are both recovered; we relate the disagreement in the cumulative distribution at low redshift to the known cluster at $z \sim 0.7$; some excess at $z \sim 0.5$ can be again connected to the excess of bright galaxies in the *K*-band LF. The comparison with the deeper but less wide GOODS-MUSIC sample (Fig. 4) confirms the good agreement at brighter limit magnitudes (301 predicted versus 262 observed galaxies with $K_{\rm AB} < 20$; 722 predicted versus

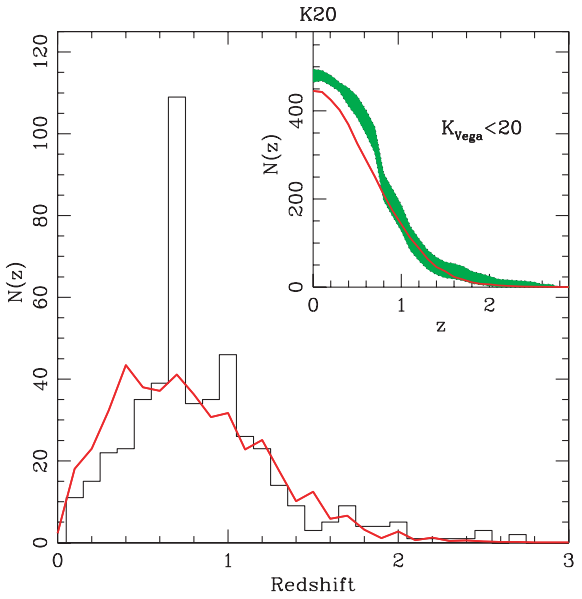

Figure 3. Redshift source distribution compared to K20 sample (histogram, Cimatti et al. 2002a). Solid line refers to MORGANA prediction.

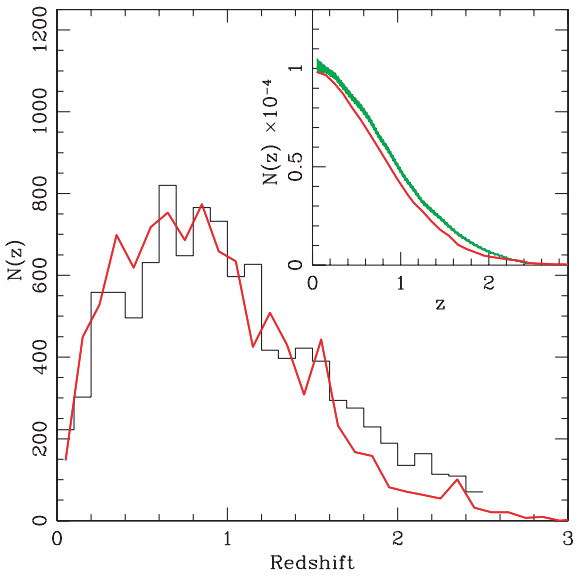

Figure 5. Redshift source distribution compared to VVDS sample (histogram, Pozzetti et al. 2007). Solid line refers to MORGANA prediction.

623 observed galaxies with $K_{AB} < 21$). At fainter limit magnitudes the comparison highlights an excess of model galaxies (1687 predicted versus 1258 observed galaxies with $K_{AB} < 22$; 3572 predicted versus 2603 observed galaxies with $K_{AB} < 23$). This excess is due not only to the small local excess noticeable in the $z = 0$ LF (Fig. 1), but to a generalized excess of faint sources at all redshifts. Fontana et al. (2006) showed that the overprediction of small galaxies ($M \sim 10^{10}\,M_{\odot}$) at $z \sim 1$ is a common problem of galaxy formation models; here we see that the problem is probably present since high redshift. Finally, we also compare MORGANA predictions with the larger VVDS sample (Fig. 5, we combined model predictions for the two subsamples separately in the same way as observational data): we obtain again a good agreement with observations (9822 predicted objects against 10 160 observed); except for a mild underestimate, mostly due to the smaller number of $z \sim 1$–2 objects with respect to observations.

A clearer view is obtained by considering the evolution of the K-band LF with redshift (Fig. 6). We compare our model with the K20 (Pozzetti et al. 2003; Vega system) and UDS (Cirasuolo et al. 2007; AB magnitudes) data; the K20 estimate is based on spectroscopic redshifts but the sampled area is small, the UDS sample covers a larger area of the sky but is based on photometric redshifts. The overall agreement between MORGANA and the two data sets is very good. We can notice some overestimate of the faint end at $z \sim 0.5$, which is not as visible as in Fontana et al. (2006). Also, the lower redshift bins of the wider UDS sample show that the excess at the bright end builds up at $z \lesssim 1$.

From the analysis of Figs 1–6 we conclude that MORGANA is able to reproduce the assembly of the bulk of the stellar mass, which is mostly in place already at $z \sim 1$–2; however, the biggest model galaxies continue growing in mass at $z < 1$, though this growth is partially compensated for by the loss of stars to the halo by

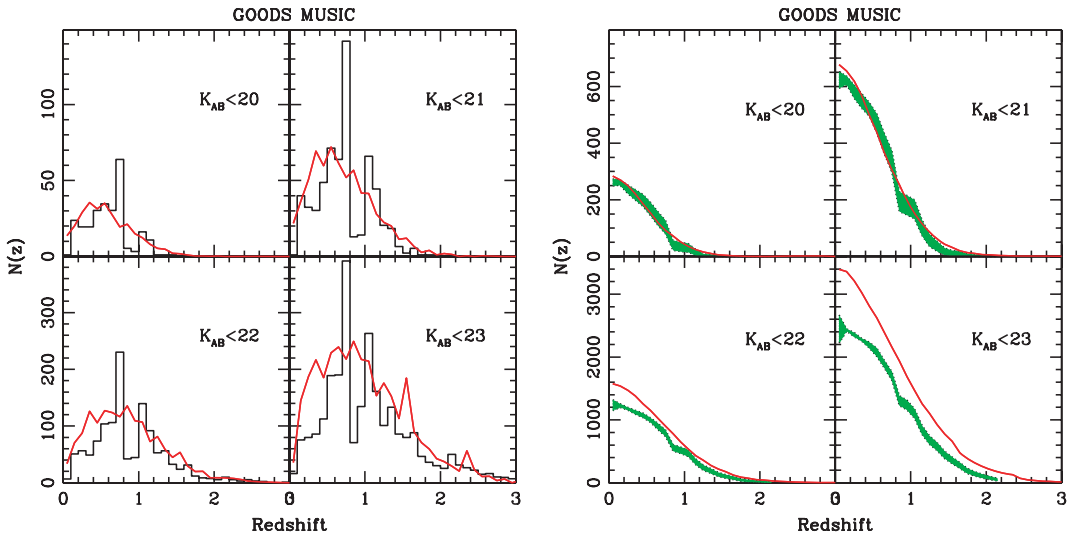

Figure 4. Redshift source distribution compared to GOODS-MUSIC sample (histograms, Fontana et al. 2006). Left-hand panel: redshift distributions. Right-hand panel: cumulative distributions. Solid line refers to MORGANA predictions.

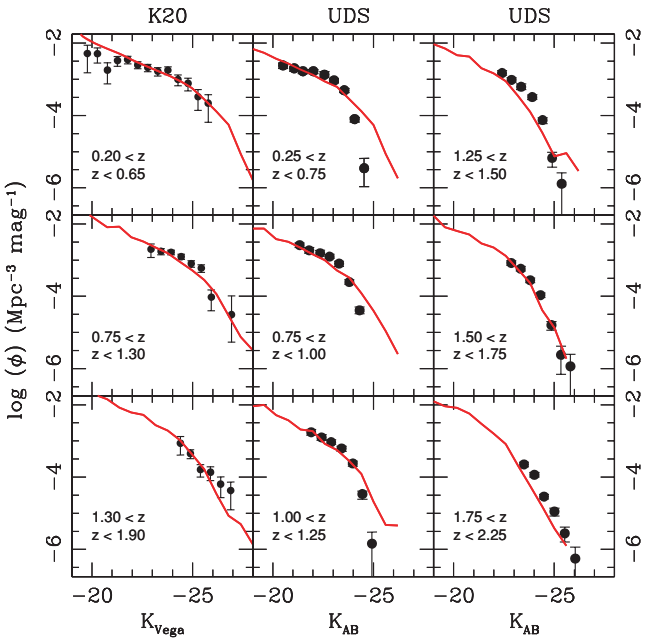

Figure 6. LF redshift evolution. Left-hand panels: K20 sample, data points taken from Pozzetti et al. (2003). Middle and right-hand panels: UDS sample, data points taken from Cirasuolo et al. (2007). Solid line refers to MORGANA predictions.

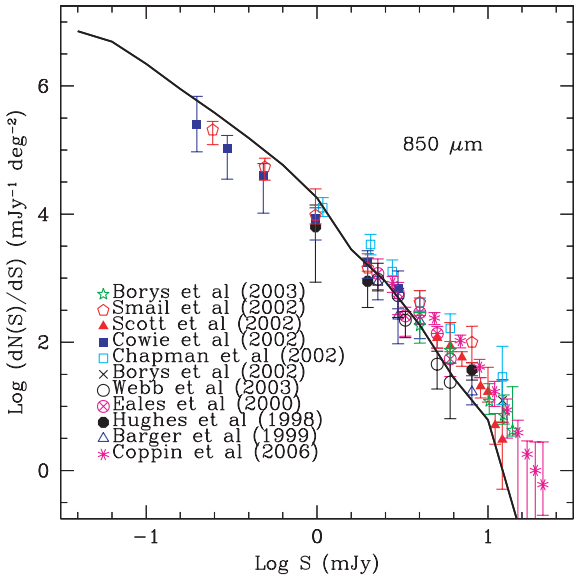

Figure 7. 850-μm source number counts. Data refer to observations as listed in the figure. Solid line refers to MORGANA prediction.

gravitational scattering. Moreover, the progressive build-up of smaller objects is not correctly reproduced; small objects tend to be too many at $z \sim 1$ (and too old at $z \sim 0$) in the model. These findings are in agreement with the analysis presented in Fontana et al. (2006).

3.2 850 μm band

We now test whether the stars visible at $z < 2$ in the K band were formed in much smaller chunks that were later assembled together,

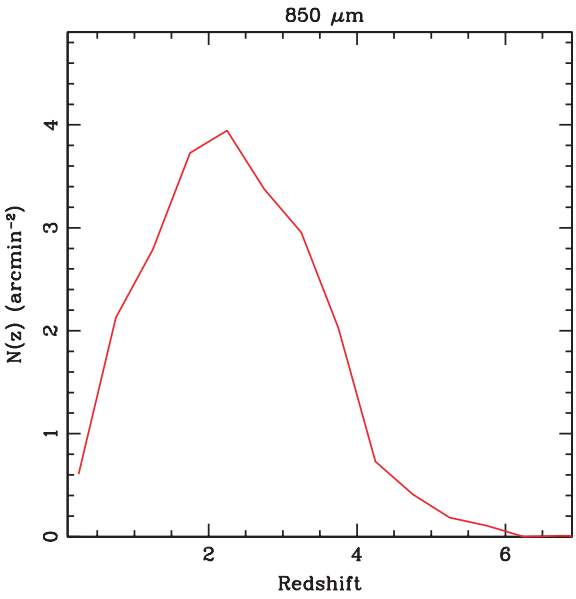

Figure 8. Predicted redshift distribution for 850-μm sources with $S >$ 0.2 mJy.

or in big massive starbursts at $z \gtrsim 2$. This is best tested by comparing the model to number counts at 850 μm, where massive starbursts at $1 \lesssim z \lesssim 5$ are easily observed. Fig. 7 shows our prediction compared to SCUBA observations (Hughes et al. 1998; Barger, Cowie & Sanders 1999; Eales et al. 2000; Borys et al. 2002; Chapman et al. 2002; Cowie, Barger & Kneib 2002; Scott et al. 2002; Smail et al. 2002; Webb et al. 2002; Borys et al. 2003; Coppin et al. 2006; Scott et al. 2006). In order to avoid contamination due to low-redshift sources, we consider here a subsample with $z > 0.5$. We stress that this prediction is done assuming a standard Salpeter IMF. Fig. 8 gives the redshift distribution of the objects with flux >0.2 mJy, which corresponds to the faintest point in Fig. 7. It is peaked at $z \sim 2.0$, in qualitatively agreement with the observations of Chapman et al. (2003, 2005) of the brightest SCUBA sources (redshift distribution peaked at $z \sim 2.4$ with a quartile range of $1.9 < z < 2.8$). The model reproduces well the data, with a possible modest overestimate (underestimate) at faint (bright) fluxes. However, a deeper analysis of the predictions shows that the population of the brightest objects (>5 mJy) at $z > 2$ (Chapman et al. 2005; Aretxaga et al. 2007) is missing: we are able to reproduce the bulk of the heavily star-forming population at $z \sim 2$, but not the most luminous objects.

4 DISCUSSION

Submillimetre counts have long been an elusive piece of evidence to fit for semi-analytic models of galaxy formation (e.g. Guiderdoni et al. 1998; Devriendt & Guiderdoni 2000; Granato et al. 2000; but see Baugh et al. 2005). Conversely, in the long tuning and debugging process of the MORGANA code, we have never found any difficulty in fitting the bulk of number counts using a conservative choice for the IMF. The difference can be ascribed to the different cooling model that MORGANA is implementing. In a forthcoming paper (Viola et al., in preparation) we compare analytic cooling models to N-body + hydro simulations in the simplest case of an isolated halo with its hot gas component initially in hydrostatic equilibrium, and no feedback from star formation or AGN. The classical model of cooling, based on the computation of a cooling radius (White & Frank 1991) is

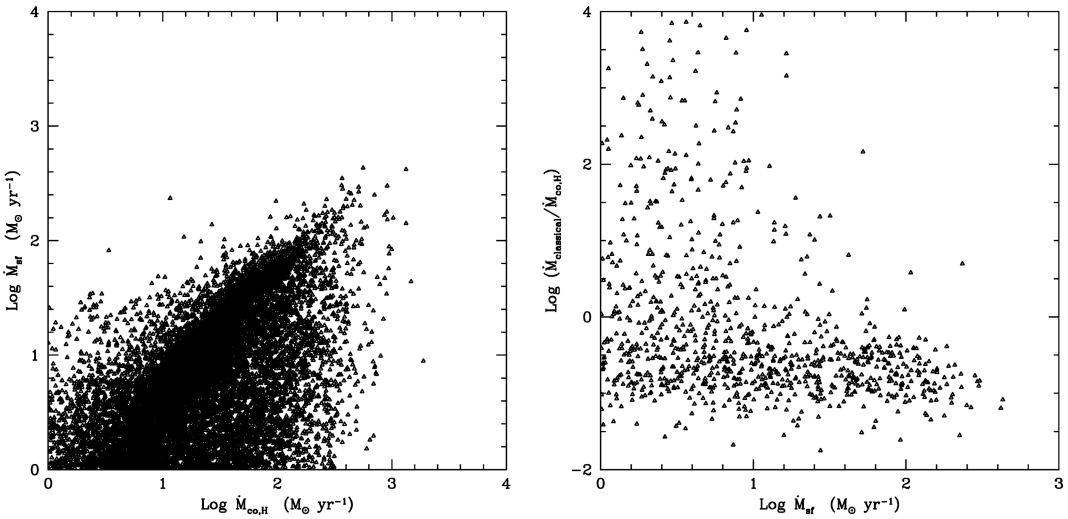

Figure 9. Left-hand panel: star formation rate (bulge + disc) versus cooling rate for central galaxies at $z = 2.5$. Right-hand panel: ratio of the cooling flows as predicted by the classical model and as used by MORGANA versus total star formation rate for the same galaxies as above. In the right-hand panel many points, especially at low star formation, lie out of the plot because the classical model predicts that all the gas has cooled, while MORGANA retains some gas through feedback and cosmological infall.

found to underestimate the amount of cooled gas, especially in the first stages of cooling, while the model implemented in MORGANA and briefly described in Section 2.1 gives a very good fit. All the analytic models of cooling tend to give similar results at later times. The use of a better cooling model accelerates the accumulation of cold gas in the haloes, giving rise to stronger starbursts. The typical mass of the haloes that contain the biggest starbursts is high enough to ensure the formation of a hot halo component through shocks (see e.g. Keres et al. 2005), so that the results based on hot gas haloes in hydrostatic equilibrium apply to this case.

To support this interpretation, we run again the model, computing at each halo major merger the gas profile (as described in Section 2.1) and from it the cooling time $t_{\rm cool}(r)$ as a function of radius. The classical cooling radius $r_{\rm C}(t)$ is the inverse of the function $t_{\rm cool}(r)$. The cooling flow of the classical cooling model is then computed as $\dot{M}_{\rm classical} = 4\pi r_{\rm C}^2 \rho_{\rm g}(r_{\rm C})\, {\rm d}r_{\rm C}/{\rm d}t$. It is not possible to use the classical cooling flow directly in the model, because the cooling model includes the injection of energy by feedback, which would then be absent. Besides, a generalization of the model to include classical cooling would require deep changes in the code and a recalibration of the parameters to reproduce local observables, which is well beyond the interest of the paper. We then simply compare the classical cooling flow with that actually used in the MORGANA model. First, in Fig. 9, left-hand panel, we show MORGANA cooling flow versus star formation for all the central galaxies present in the box at $z = 2.5$. The stronger starbursts are associated with the stronger cooling flows.[3] Secondly, we show for the same haloes (right-hand panel) the ratio of classical and MORGANA cooling flows versus star formation. For massive starbursts ($\dot{M}_{\rm sf} > 100\,{\rm M}_\odot\,{\rm yr}^{-1}$) the classical model predicts cooling flows a factor of 5–10 lower, which would presumably result in correspondingly lower star formation rates. The relation changes at smaller star formation rates, where feedback has been able to suppress the MORGANA cooling flow in many cases.[4] The success of MORGANA should then be ascribed to the different modelling of cooling.

With our model we can also investigate the physical conditions of SCUBA sources. In Fig. 10(a) we show the star formation function of galaxies at $z = 2.5$, divided into bulges (mergers) and discs. In this case, we use the punctual value of the star formation rate computed at the end of the time bin. Interestingly, mergers dominate only at the highest star formation rates, $\gtrsim 200\,{\rm M}_\odot\,{\rm yr}^{-1}$. This implies that most of our starbursts are triggered by cooling/infall more than by mergers, in line with the findings of violently star-forming discs at $z \sim 2$ (Genzel et al. 2006). However, this does not imply that massive starbursts are associated with spiral galaxies. In Fig. 10(b) the star formation rate is computed as the total amount of stars formed in the 0.1 Gyr time bin and eventually found (at the end of the time bin) in a bulge or disc component, divided by the width of the time bin. This time bulges clearly dominate the star formation function. This shows that such massively star-forming discs merge into bulges (disc instabilities are not relevant at this redshift) in less than 0.1 Gyr. Such short time-scales guarantee high levels of α-enhancement, so the stellar populations formed in these starbursts will be typical of early-type galaxies. We conclude that the cooling/infall domination of high-redshift starbursts does not change the conclusion that these events are responsible for the formation of the stars found today in elliptical galaxies.

The overall agreement between model and data is in line with recent results by De Lucia et al. (2006), Bower et al. (2006), Croton et al. (2006) and Kitzbichler & White (2006). However, there are three main points of disagreement between our model and the data. First, the evolution of the most massive galaxies at $z < 1$ is too strong, resulting in an excess of bright galaxies in the K band. It is easy to absorb this excess at $z = 0$ by tuning the parameters, but this would be done at the expense of a poor fit of the stellar mass function and K-band LF at $z \sim 1$. As demonstrated by

[3] A few massive starburst are associated with no cooling flow and then lie outside this relation; these are mergers where feedback has already quenched cooling, or where cooling has already deposited most mass in the galaxy.

[4] Many points at low star formation lie out of this plot because the classical model predicts no cooling, while in MORGANA gas is still present due to feedback and cosmological infall.

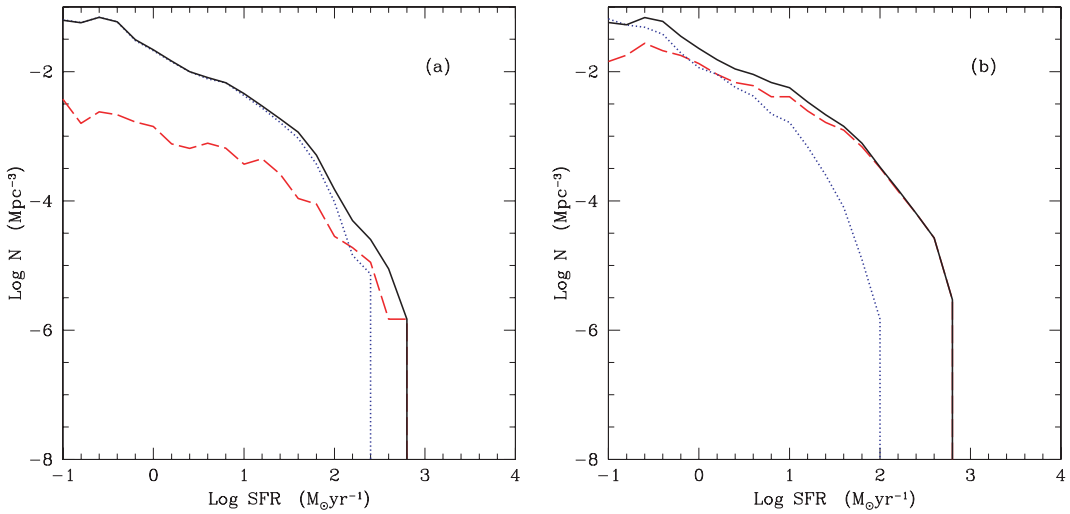

Figure 10. Star formation rate function at $z = 2.5$ for discs (blue dotted line), bulges/mergers (red dashed line) and total (black continuous line). The star formation in panel (a) is the punctual value at the end of the integration over the time bin and that in (b) is the average in the bin (of width 0.1 Gyr) relative to the stars contained in bulges and discs at the end of the integration.

Monaco et al. (2006), this evolution is driven by galaxy mergers, so no feedback recipe can solve it. Our implementation of the suggestion of Monaco et al. (2006) to scatter 40 per cent stars to the diffuse component of galaxy clusters at each merger goes in the right-hand side direction but is largely insufficient to suppress this evolution. A much higher scattered fraction, like ∼80 per cent of the stellar mass of the satellite, would give a better results (as argued also by Conroy et al. 2007; Renzini 2007). However, such an extreme value applied to all mergers and all redshifts does not allow to reproduce the ∼1 per cent fraction of diffuse stars in the Milky Way together with the ∼10 per cent fraction in Virgo and the ∼40 per cent fraction in massive clusters. We conclude that, while the mechanism is promising and deserves further attention, an easy and straightforward implementation does not work; this mechanism should be selectively efficient in the most massive DM haloes at low redshift.

Secondly, the model is not able to reproduce the most luminous SCUBA galaxies at redshift $z \sim 2$, which correspond to the strongest starbursts. This also hints to an insufficient 'downsizing' in our predicted galaxy population. In the most massive galaxies the fraction formed at high redshift in episode of intense star formation is still too small, and the fraction of stars accreted at later times is still too large, with respect to the observed trends. Given the success in reproducing the bulk of the SCUBA population, a detailed investigation on the predictions of the cooling/infall model in the more massive haloes is promising. Also the choice of a less conservative IMF, i.e. the Kroupa IMF, can help to solve the problem, thanks to the larger number of massive stars predicted at each stellar generation.

Thirdly, we confirm the finding of Fontana et al. (2006) that our model produces too many faint objects at $z \sim 1$; in other words, our smaller galaxies are too old. Coupled with the excessive evolution of the most massive galaxies, this finding shows that, while reproducing the average build-up of galaxies, the observed trend of less massive galaxies being on average younger is not properly reproduced by MORGANA. Fontana et al. (2006) have shown very convincingly that this problem is shared by many galaxy formation codes, either numerical or semi-analytic, so our results can be considered as the typical outcome of the ΛCDM cosmogony, given the present understanding of the physics of galaxy formation. Fig. 11 shows a comparison of the evolution of the stellar mass function as inferred by the GOODS-MUSIC sample and as predicted by MORGANA. Clearly, the 'downsizing' trend of more massive galaxies evolving very slowly at $z < 1$, while the population of less massive galaxies builds up, is not reproduced. While the evolution of the high-mass end is driven by mergers, a delay in the build-up of faint galaxies should be caused by some source of feedback.

5 CONCLUSIONS

This paper is the third of a series devoted to presenting MORGANA. We have demonstrated that the model is able to follow the build-up of the bulk of stars, more precisely to reproduce the early assembly and late, almost-passive evolution of massive galaxies. To this aim we have combined MORGANA with GRASIL and compared predictions with observation in the submillimetre (the 850-μm channel), especially sensitive to the strongest and most obscured episodes of star formation, and in the near-infrared (especially in the K band), most sensitive to the stellar mass. Overall consistency between model and observations has been obtained using SCUBA counts at 850 μm, number counts in the K band, redshift distribution of K-limited galaxy samples, and redshift evolution of the K-band LF. The importance of this result is strengthened by the use of a standard Salpeter IMF (e.g. a very conservative choice) along the whole redshift interval. We ascribe this agreement mostly to our improved model for cooling/infall, that correctly reproduces the results controlled numerical experiments.

We predict that most star formation at high redshift is not stimulated by starbursts but is due to the strong cooling/infalling flows that take place at early times. This does not imply that such discs are close analogues of local spiral galaxies, as their gas surface density is very high, their star formation rate is typical of starburst galaxies and most of them are expected to merge within less than 0.1 Gyr. We then predict that gas-rich discs, characterized by star formation rates up to 100–200 $M_\odot$ yr^{-1}, should be very abundant at $z \sim 2$, in line with the observations of Genzel et al. (2006).

Despite these successes, our model does not reproduce the 'downsizing' trend of a modest evolution of the most massive galaxies accompanied by a build-up of small galaxies at $z \lesssim 1$. We propose

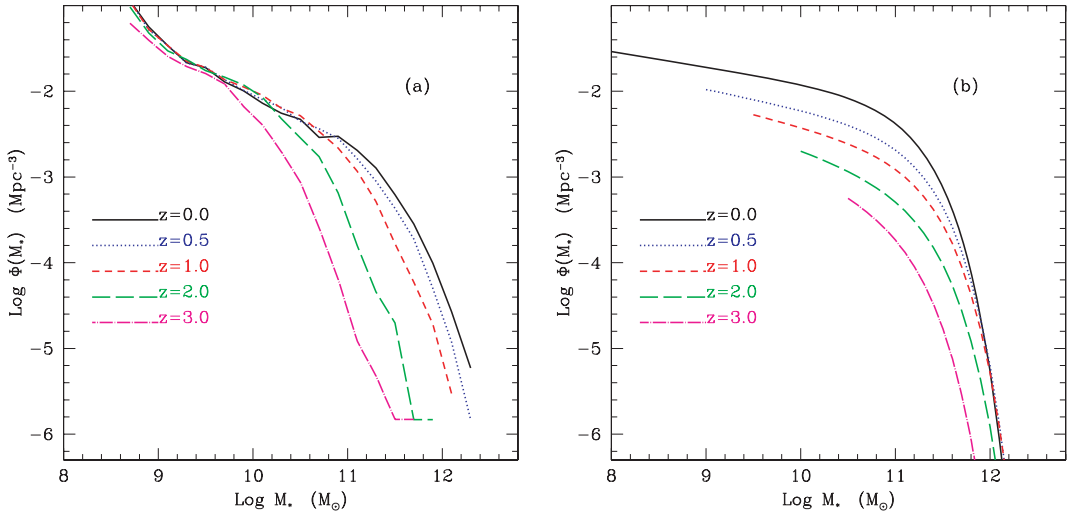

Figure 11. Evolution of the stellar mass function according to MORGANA (panel a) and GOODS-MUSIC (panel b). The latter mass functions are drawn roughly in the same mass interval where data are available.

that a solution of this discrepancy requires at least two mechanisms, because the evolution of the bright end is driven by inevitable galaxy mergers, while star formation in less massive galaxies is sensitive to feedback. The solution proposed by Monaco et al. (2006) for slowing the evolution of massive galaxies requires an implementation of scattering of stars to the diffuse component that is strongly dependent on DM halo mass and/or redshift. On the other hand, a solution of the overabundance of $10^{11}\,M_{\odot}$ galaxies at $z \sim 1$, common to most galaxy formation models (Fontana et al. 2006), calls for some unknown source of feedback.

ACKNOWLEDGMENTS

We thank Stefano Cristiani, Andrea Cimatti, Gianluigi Granato, Adriano Fontana, Alvio Renzini, Carlos Frenk, Cedric Lacey, Rachel Somerville and Eric Bell for many enlightening discussions; we also thank Lucia Pozzetti for providing the VVDS redshift distribution. PM thanks the ICC of Durham for its hospitality. Calculations were carried out both at the 'Centro Interuniversitario del Nord-Est per il Calcolo Elettronico' (CINECA, Bologna) with CPU time assigned under University of Trieste/CINECA grants, and at the PIA cluster of the Max-Planck-Institut für Astronomie at the Rechenzentrum Garching.

REFERENCES

Aretxaga I. et al., 2007, MNRAS, 379, 1571
Barger A. J., Cowie L. L., Sanders D. B., 1999, ApJ, 518, L5
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Bernardi M. et al., 2003, AJ, 125, 1866
Blakeslee J. P. et al., 2003, ApJ, 596, L143
Borys C., Chapman S. C., Halpern M., Scott D., 2002, MNRAS, 330, L63
Borys C., Chapman S. C., Halpern M., Scott D., 2003, MNRAS, 344, 385
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bressan A., Granato G. L., Silva L., 1998, A&A, 332, 135
Bressan A., Silva L., Granato G. L., 2002, A&A, 392, 377
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Chapman S. C., Scott D., Borys C., Fahlman G. G., 2002, MNRAS, 330, 92
Chapman S. C., Blain A. W., Ivison R. J., Smail I. R., 2003, Nat, 422, 695
Chapman S. C., Blain A. W., Smail I. R., Ivison R. J., 2005, ApJ, 622, 772
Cimatti A. et al., 2002a, A&A, 381, 68
Cimatti A. et al., 2002b, A&A, 391, 1
Cimatti A., Daddi E., Renzini A., 2006, A&A, 453, L29
Cirasuolo M. et al., 2007, MNRAS, 380, 585
Cole S., Lacey C. G., Baugh C. M., Frenck C. S., 2000, MNRAS, 319, 168
Cole S. et al., 2001, MNRAS, 326, 255
Conroy C., Weschler R. H., Kravtsov A. V., 2007, preprint (astro-ph/0703374)
Coppin K. et al., 2006, MNRAS, 372, 1621
Cowie L. L., Barger A. J., Kneib J.-P., 2002, AJ, 123, 2197
Croton D. J. et al., 2006, MNRAS, 365, 11
Daddi E. et al., 2002, A&A, 384, L1
Davis M. et al., 2003, Proc. SPIE, 4834, 161
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Devriendt J. E. G., Guiderdoni B., 2000, A&A, 363, 851
Djorgovski S. et al., 1995, ApJ, 438, L13
Dye S. et al., 2006, MNRAS, 372, 1227
Eales S., Lilly S., Webb T., Dunne L., Gear W., Clements D., Yun M., 2000, AJ, 120, 2244
Ellis S. C., Jones L. R., Donovan D., Ebeling H., Khosroshahi H. G., 2006, MNRAS, 368, 769
Fontana A. et al., 2006, A&A, 459, 745
Fontanot F., Monaco P., Cristiani S., Tozzi P., 2006, MNRAS, 373, 1173 (Paper II)
Gardner J. P., Cowie L. L., Wainscoat R. J., 1993, ApJ 415, 9
Gardner J. P., Sharples R. M., Carrasco B. E., Frenk C. S., 1996, MNRAS, 282, 1
Gavazzi G., Pierini D., Boselli A., 1996, A&A, 312, 397
Genzel R. et al., 2006, Nat, 442, 786
Giavalisco M. et al., 2004, ApJ, 600, L103
Glazebrook K., Peacock J. A., Collins C. A., Miller L., 1993, MNRAS, 266, 65
Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, ApJ, 542, 710
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580

Grazian A. et al., 2006, A&A, 449, 951
Guiderdoni B., Hivon E., Bouchet F. R., Maffei B., 1998, MNRAS, 295, 877
Hughes D. H. et al., 1998, Nat, 394, 241
Iovino A. et al., 2005, A&A, 442, 423
Jarrett T.-H., Chester T., Cutri R., Schneider S., Rosenberg J., Huchra J.-P., Mader J., 2000, AJ, 120, 298
Kang X., Jing Y. P., Mo H. J., Börner G., 2005, ApJ, 631 21
Kennicutt R. C., 1989, ApJ, 344, 685
Keres D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kitzbichler M. G., White S. D. M., 2006, MNRAS, 366, 858
Kochanek C. S. et al., 2001, ApJ, 560, 566
Kodama T., Arimoto N., Barger A. J., Aragón-Salamanca A., 1998, A&A, 334, 99
Kong X. et al., 2006, ApJ, 638, 72
Lawrence A. et al., 2007, MNRAS, 379, 1599
Le Fevre O. et al., 2005, A&A, 439, 877
Li A., Draine B. T., 2001, ApJ, 554, 778
Maihara T. et al., 2001, PASJ, 53, 25
Maraston C., 2005, MNRAS, 362, 799
Matteucci F., 1996, ASP Conf. Ser. Vol. 98, From Stars to Galaxies: The Impact of Stellar Physics on Galaxy Evolution. Astron. Soc. Pac., San Francisco, p. 529
Menci N., Cavaliere A., Fontana A., Giallongo E., Poli F., Vittorini V., 2004, ApJ, 604, 12
Mo H. J., Mao S., 2000, MNRAS, 318, 163
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Mobasher B., Ellis R. S., Sharples R. M., 1986, MNRAS, 223, 11
Moustakas L. A., Davis M., Graham J. R., Silk J., Peterson B. A., Yoshii Y., 1997, ApJ, 475, 445
Monaco P., 2004a, MNRAS, 352, 181
Monaco P., 2004b, MNRAS, 354, 151
Monaco P., Fontanot F., 2005, MNRAS, 359, 283
Monaco P., Theuns T., Taffoni G., 2002, MNRAS, 331, 587
Monaco P., Murante G., Borgani S., Fontanot F., 2006, ApJ, 652, L89
Monaco P., Fontanot F., Taffoni G., 2007, MNRAS, 375, 1189 (Paper I)
Murante G., Giovalli M., Gerhard O., Arnaboldi M., Borgani S., Dolag K., 2007, MNRAS, 377, 2
Nagamine K., Cen R., Hernquist L., Ostriker J. P., Springel V., 2005, ApJ, 618, 23
Nagashima M., Lacey C. G., Okamoto T., Baugh C. M., Frenk C. S., Cole S., 2005, MNRAS, 363, L31
Neisteinn E., van den Bosch F., Dekel A., 2006, MNRAS, 372, 933
Panuzzo P., Bressan A., Granato G. L., Silva L., Danese L., 2003, A&A, 409, 99
Pozzetti L. et al., 2003, A&A, 402, 837
Pozzetti L. et al., 2007, preprint (astro-ph/0704.1600)
Refregier A., 2003, ARA&A, 41, 645
Renzini A., Ciotti L., 1993, ApJ, 416, L49
Renzini A., 2007, preprint (astro-ph/0702148)
Rix H.-W., Bardeen M., Beckwith S. V. W. et al., 2004, ApJS, 152, 163
Robertson B., Li Y., Cox T. J., Hernquist L., Hopkins P. F., 2007, ApJ, 667, 60
Saro A., Borgani S., Tornatore L., Dolag K., Murante G., Biviano A., Calura F., Charlot S., 2006, MNRAS, 373, 397
Scott S. E. et al., 2002, MNRAS, 331, 817
Scott S. E., Dunlop J. S., Serjeant S., 2006, MNRAS, 370, 1057
Scoville N. et al., 2007, ApJS, 172, 1
Silva L., 1999, PhD thesis, SZSSA, Trieste
Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103
Silva L., De Zotti G., Granato G. L., Maiolino R., Danese L., 2005, MNRAS, 357, 1295
Smail I., Ivison R., Blain A., 1997, ApJ, 490, L5
Smail I., Ivison R., Blain A., Kneib J.-P., 2002, MNRAS, 331, 495
Smith G. P. et al., 2002, MNRAS, 330, 1
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Somerville R. S. et al., 2004, ApJL, 600, L135
Spergel D. N. et al., 2007, ApJS, 170, 377
Springel V. et al., 2005, Nat, 435, 629
Sutherland R., Dopita M., 1993, ApJS, 88, 253
Szokoly G. P., Subbarao M. U., Connolly A. J., Mobasher B., 1998, ApJ, 492, 452
Taffoni G., Monaco P., Theuns T., 2002, MNRAS, 333, 623
Taffoni G., Mayer L., Colpi M., Governato F., 2003, MNRAS, 341, 434
Thomas D., Maraston C., Bender R., de Oliveira C. M., 2005, ApJ, 621, 673
Umemura M., 2001, ApJ, 560, L29
Vega O., Silva L., Panuzzo P., Bressan A., Granato G. L., Chavez M., 2005, MNRAS, 364, 1286
Viola M., Borgani S., Monaco P., Murantee G., Tornatore L., 2007, MNRAS, in press
Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559
Webb T. et al., 2003, ApJ, 587, 41
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, Cosmology and Large Scale Structure. Elsevier Scientific, Amsterdam, p. 349
Wolf C., Meisenheimer K., Röser H.-J., 2001, A&A, 365, 660
Wu K. K. S., Fabian A. C., Nulsen P. E. J., 2001, MNRAS, 318, 889
Zheng X. Z., Bell E. F., Papovich C., Wolf C., Meisenheimer K., Rix H.-W., Rieke G. H., Somerville R., 2007, ApJ, 661, L41

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