Durham E-Theses

The nature of early-type galaxies in hierarchical models

Almeida, Cesário Manuel de Deus Lavaredas de

How to cite:

Almeida, Cesário Manuel de Deus Lavaredas de (2008) *The nature of early-type galaxies in hierarchical models*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/2181/

Use policy

The Nature of Early-Type Galaxies

in Hierarchical Models

Cesário Manuel de Deus Lavaredas de Almeida

A Thesis submitted to the University of Durham
in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy

Extragalactic Astronomy & Cosmology Group
Department of Physics
University of Durham
England

July 2008

Early-type galaxies in hierarchical cosmologies

by Cesário Manuel de Deus Lavaredas de Almeida

July 2008

Abstract

In this Thesis we describe the properties of early-type galaxies in the context of hierarchical galaxy formation. We use two variants of the GALFORM model originally introduced by Cole et al.: the Baugh et al. and the Bower et al. models. We test the prescription defined by Cole et al. to calculate the sizes of bulges, by comparing GALFORM predictions with local observational data. We find that the model reproduces successfully several tight correlations observed for early-type galaxies: the relation between velocity dispersion and luminosity, the velocity dispersion–age relation and the Fundamental Plane. However, there is an important disagreement between the models and observations: in the model, the radii of the luminous spheroids are smaller than expected. We analyse how the physical ingredients involved in the calculation of the sizes influence these results.

We explore the physics of massive galaxy formation in the models, by predicting the abundance, properties and clustering of luminous red galaxies (LRGs). Without adjusting any parameters in the two models, we find a good agreement between the GALFORM model and observations. We find that model LRGs are mainly elliptical galaxies, with stellar masses around $2 \times 10^{11} \, h^{-1} M_\odot$ and velocity dispersions of 250 kms^{-1}. The models predict the correlation function of LRGs to be a power law down to small scales, which is in excellent agreement with the observational estimates.

Finally, we predict the abundance, colour and clustering of submillimeter galaxies (SMGs), which are thought to be the progenitors of local massive early-type galaxies. At the wavelengths where these galaxies are detected, 850 μm, the predictions using the standard GALFORM model are inaccurate, hence the time-consuming GRASIL code is used in addition to GALFORM. We develop a new method, based on artificial neural networks, to rapidly generate galaxy spectra from a small set of properties.

Beyond the bend in the road

There may be a well, and there may be a castle,

And there may be just more road.

I don't know and I don't ask.

As long as I'm on the road that's before the bend

I look only at the road before the bend,

Because the road before the bend is all I can see.

It would do me no good to look anywhere else

Or at what I can't see.

Let's pay attention only to where we are.

There's enough beauty in being here and not somewhere else.

If there are people beyond the bend in the road,

Let them worry about what's beyond the bend in the road.

That, for them, is the road.

If we're to arrive there, when we arrive there we'll know.

For now we know only that we're not there.

Here there's just the road before the bend, and before the bend

There's the road without any bend.

Alberto Caeiro (Fernando Pessoa), 1914

CONTENTS

iv

LIST OF FIGURES

LIST OF TABLES

DECLARATION

The work in this thesis is based on research carried out between 2004 and 2008 at the Institute for Computational Cosmology, Department of Physics, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text. Portions of this thesis have been published in the following papers:

- Almeida C., Baugh C.M., Wake D.A., Lacey C.G., Benson A.J., Bower R.G, Pimbblet K., 2008, MNRAS, 386, 2145

- Almeida C., Baugh C.M., Lacey C.G., 2007, MNRAS, 376, 1711

ACKNOWLEDGEMENTS

Chapter 1

Introduction

1.1 The Universe as we know it

It is generally assumed that the Universe is uniform on large scales. This cosmological principle, homogeneity and isotropy, means that if we study the Universe on scales well in excess of ~ 100 Mpc, the direction of observation is not important as any part of the sky will look similar structurally, i.e. no matter were we look in the Universe (at large scales), we should find the same distribution of objects. Furthermore, an important corollary of the cosmological principle is that the laws of physics are universal, along with the physical constants. With the comological principle, the properties and physics of our local surroundings can be extrapolated across the whole Universe. On small scales the picture is definitely different. The homogeneity no longer holds: galaxies and gas are clustered, forming denser regions and voids. There is also anisotropy on small scales: galaxies tend to be aligned in filaments and walls (see for example the results from the Two-degree Field Galaxy Redshift Survey, Colless et al., 2001). It is accepted that the observed structure in the Universe is a result of the growth of the initial small perturbations in the density field due to gravitational instability. The existence of these perturbations is a prediction of inflation theory (Guth, 1981), which says that the Universe must have undergone a period of exponential growth to a size greater than the cosmological horizon just after the Big Bang. During this process, tiny quantum fluctuations arose in the scalar field and were amplified to classical scales by the inflation process itself.

Perhaps the two most important observational discoveries in cosmology, which

are accepted as confirmation of the Big Bang theory, are the detection of the cosmic microwave background (CMB) and the detection of temperature anisotropies in the CMB which arise from primordial fluctuations in the density field. According to the theory, the current background temperature of the Universe should be described by a blackbody spectrum at 2.7 K (Gamow, 1956). Penzias & Wilson (1965) accidentally detected this radiation and correctly associated it with the CMB. In the early 1990s, the *Cosmic Background Explorer* (COBE) satellite confirmed this measurement with greater precision and showed that, in fact, the spectrum of the cosmic microwave background is amazingly similar to a blackbody (Mather et al., 1990). Moreover, COBE was able to detect anisotropies in the temperature of the CMB which are related to fluctuations in density (Smoot et al., 1992). These measurements further validated the Big Bang theory.

The pattern of the fluctuations in the density field, and consequent structure formation, depend on the composition of the Universe. The first evidence of matter that did not emit or reflect radiation, *dark matter*, came from kinematical studies of galaxies in the Coma cluster (Zwicky, 1933). Zwicky found that the luminous mass, using a given mass-to-light ratio (M/L), was much smaller than the dynamical mass calculated using simple Newtonian physics. It was more than 40 years later before further observations added support to the existence of dark matter, through observations of the rotation curves of spiral galaxies which revealed that the rotation speed at large distance from the center did not follow Keplerian laws, but remains constant (e.g. Faber & Gallagher, 1979; Rubin et al., 1980).

The composition of dark matter is still unknown. Several candidates have been proposed, classified as to whether the particles move at relativistic ("hot" dark matter) or non-relativistic velocities ("cold" dark matter) (CDM) at the decoupling time. However, nowadays we know that Hot Dark Matter by itself can not explain the formation of structure on small scales in the Universe (e.g. White, Frenk & Davies, 1983). This is because the fast moving particles (the best candidate is the neutrino) would have smoothed away the initial small scale clumping due to a phenomenon called free streaming (Bond et al., 1980). In a Cold Dark Matter

universe, however, the non-relativistic particles[1] have the potential to collapse and form clumps at high redshift. In this scenario structure formation is hierarchical, with larger objects forming at later times as a consequence of the merging of smaller structures.

CDM is currently the model that has accumulated the most support over the recent years. The most convincing is the anisotropies found in the cosmic microwave background radiation, as we noted previously. More recently, large sky surveys, such as the 2dFGRS (Colless et al., 2001) and the SDSS (Abazajian et al., 2005), have mapped the large scale structure of the Universe using the galaxy distribution, and have provided further support for the model. Comparison with observations also provides a means to constrain the parameters of the CDM model. For example, according to the most recent measurements of the CMB by the high angular resolution Wilkinson Microwave Anisotropy Probe (WMAP) (Hinshaw et al., 2008) combined with large scale structure measurements, the baryonic matter in the Universe represents only 4.6% of the critical density, while the dark energy density is $\Omega_\Lambda = 0.74$. Constraints on the curvature of the Universe point to a value very close to $k = 0$, which states that the Universe is flat, and so its density may be expressed as $\Omega_m + \Omega_\Lambda = 1$ (with the matter density, evidently, given by $\Omega_M = 0.26$).

Lastly, it is important to mention that the computational implementation of the ΛCDM model is straightforward, which, in turn, facilitates the implementation in semi-analytic modelling and, ultimately, provides a direct comparison with observations.

1.2 Creating Universes

The semi-analytic modelling of the formation and evolution of galaxies had its origins with the work of White & Rees (1978) and later by Cole (1991) (see also White & Frenk, 1991; Lacey & Silk, 1991). These authors devised a simple galaxy formation model, where dark matter haloes evolved hierarchically following a Monte Carlo approach based on the Press-Schechter theory (Press & Schechter, 1974; Bond et al.,

---

[1]several particles fit the bill, e.g. WIMPS, axions, neutralinos

1991). Inside these structures, radiative cooling processes allow the condensation of the baryonic matter to take place, creating objects which would constitute the galaxies. The higher density of condensed baryons makes star formation possible and also means that these objects can survive intact when their host dark matter halo merges into a large object. Cole (1991) showed that with such a simple prescription it was possible to reproduce the local observed luminosity functions. In fact, the authors demonstrated that it was also imperative to include feedback effects from supernovae to regulate the efficiency of galaxy formation in haloes of different masses. Further developments were implemented by White & Frenk (1991), with the inclusion of simple recipes to calculate chemical enrichment and synthetic stellar spectra, and a few years later by Kauffmann et al. (1993) and Cole et al. (1994), and calculations of the formation of galaxy discs and bulges (Baugh et al., 1996). Another landmark in the context of semi-analytic modelling was set by the Cole et al. (2000) model, GALFORM. In this model, the authors added a physical calculation of the sizes of galaxies, and it was one of the first models to include a self-consistent calculation of dust attenuation. The model successfully matches the local luminosity functions (arguably, a prerequisite for any model of galaxy formation to satisfy before considering further predictions), the distribution of colours, the ratio of elliptical-to-spiral galaxies and the distribution of disc sizes. In Granato et al. (2000), a comprehensive treatment of the dust effects was added to the GALFORM semi-analytic code, through the GRASIL code (Silva et al., 1998). These authors were able to use this hybrid code to reproduce the local luminosity functions from the ultraviolet (UV) to the far-infrared.

A common problem that semi-analytic models face is that of the overproduction of bright galaxies (e.g. White & Frenk, 1991; Kauffmann et al., 1993). Several mechanisms have been proposed to suppress the formation of luminous galaxies, e.g. the injection of energy into the hot gas halo (Bower et al., 2001; McCarthy et al., 2007), ejection of gas from the halo in a "super wind" (Benson et al., 2003), suppression of the cooling flow due to heating by an active galactic nucleus (AGN) (Croton et al., 2006; Bower et al., 2006), among others. Even though the most successful semi-analytic models to date use either superwind or AGN feedback, a clear picture

has yet to emerge, and only direct comparison between model predictions and observations at high redshift might resolve the issue of which is the best mechanism to control massive galaxy formation.

One of such fruitful model is that of Baugh et al. (2005), which is based on the superwind model (Benson et al., 2003). This model has been tested with success over a wide range of observational data: number counts of submillimeter galaxies and luminosity function of Lyman-break galaxies, while reproducing the correct abundance of galaxies in the local Universe (Baugh et al., 2005), luminosity functions of the mid-infrared population and its evolution (Lacey et al., 2008), metallicity of the intracluster medium (Nagashima et al., 2005a) and elliptical galaxies (Nagashima et al., 2005b), the abundance of Lyman-α emitters and their properties (Le Delliou et al., 2005, 2006), some scaling relations of elliptical galaxies including the fundamental plane (Almeida et al., 2007), and the abundance, properties and clustering of luminous red galaxies (Almeida et al., 2008). Semi-analytic models of galaxy formation which implement an AGN feedback scheme have also been successful. In the Bower et al. (2006) model, a variant of the original GALFORM model, the cooling flows in massive haloes are quenched in at low redshift due to the energy released by the accretion of matter onto a central supermassive black hole. As a consequence, the authors were able to improve the match to the local luminosity functions, reproduce the observed bimodality in the colour distribution of galaxies and mimic the deduced evolution of the stellar mass function. A similar model which implemented AGN feedback was that of De Lucia et al. (2006) (see also Croton et al., 2006). Here the authors show that it is possible to accommodate, in a ΛCDM cosmogony, recent observations that suggest an "anti-hierarchical" formation of massive galaxies at high redshift.

In recent years, most of the semi-analytic models of galaxy formation and evolution are leaving behind the Monte Carlo technique to create realizations of dark matter halo trees (which is based on the Press-Schechter theory), in favour of the N-body trees (for example Croton et al., 2006; De Lucia et al., 2006; Bower et al., 2006). The existence of high-resolution N-body simulations of dark matter allowed the advent of hybrid techniques, where the halo history trees are extracted directly

from the simulation. This procedure has the advantage of being more accurate than the Press-Schechter formalism (see Helly et al., 2003), and provides the spatial distribution of galaxies, allowing a more direct comparison between model predictions and observations (we note that, improved Monte Carlo methods now exist which have been tuned to match the trees from N-body simulations (see Cole et al., 2008; Parkinson et al., 2008)). The *Millennium Simulation* (Springel et al., 2005) is the current state of art in simulations of dark matter clustering. It is a simulation of hierarchical clustering of more than 10 billion dark matter particles in a ΛCDM universe, covering a volume of $(500\,h^{-1}\,\mathrm{Mpc})^3$; the mass resolution of the extracted trees is $1.72 \times 10^{10}\,h^{-1}M_\odot$, which corresponds to 20 particles.

1.3   Early-Type Galaxies

Our understanding of early-type galaxies –elliptical and S0 galaxies– in the context of hierarchical galaxy formation is still surrounded by questions. Observations of these galaxies show that they are the reddest, most massive and luminous stellar systems in the local Universe, even though they represent only one third of the total number of the galaxy population at $z = 0$ (for galaxies brighter than $M_B - 5\log h = -19.5$ mag). Historically, these galaxies were thought to be simple, homogeneous and old objects in the sky (e.g. Baade & Gaposchkin, 1963), with little star formation and insignificant quantities of gas. This classical vision motivated the first theories of early-type formation, the *monolithic collapse* scenario, where spheroidal galaxies were thought to be formed by a single burst of star formation at high redshift (possibly $z > 5$), followed by simple passive evolution (e.g. Eggen et al., 1962; Larson, 1975). The proponents of this theory pointed out its success in predicting some important correlations found for early-type galaxies such as the colour–magnitude relation and fundamental plane (see e.g. Kodama et al., 1998). A competing theory was born with the work of Toomre & Toomre (1972), and later by Negroponte & White (1983). With the help of numerical simulations, these authors showed that the merger of two spiral galaxies could produce an elliptical galaxy, i.e. a galaxy with dynamical and photometric properties similar to early-types.

High redshift surveys, in recent years, made possible the study of these galaxies over a wider fraction of cosmic history. These surveys have shown that interactions and mergers are, indeed, common processes in the past history of spheroids. Furthermore, some authors have detected evidence for recent episodes of star formation, in local early-type galaxies (Barger et al., 1996) and also in high redshift populations (van de Ven et al., 2003), which lead to the growth of stellar mass (see Bell et al., 2004; Faber et al., 2007). These results are incompatible with the picture of monolithic evolution, and favour the *hierarchical* formation scenario, where larger and massive galaxies arise from the merging of smaller objects. This is the type of behaviour expected for structure growth in a CDM cosmogony. However, even though a significant fraction of early-type galaxies exhibit signs of mergers and recent star formation, observations show that the mass involved is very small.

Semi-analytic models of galaxy formation have tried to shed some light on the formation of spheroids. In the framework of hierarchical cosmology, the models have been successful in predicting some properties of the early-type galaxies (see for example Baugh et al., 1996; Kauffmann, 1996; De Lucia et al., 2006). For example, De Lucia et al. (2006) found that, in the models, massive elliptical galaxies are the oldest and most metal rich populations, in accordance with observations (Nelan et al., 2005). Moreover, these authors demonstrated that the recently touted "down-sizing" scenario –massive galaxies display the shortest formation timescales and formed earlier–, can be explained in the context of hierarchical assembly of galaxies. However, there are still some discrepancies between semi-analytic models and observations: the models predict that galaxies in denser environments are older than field galaxies, whereas observational results reveal little difference between both populations (e.g. van Dokkum et al., 2007).

1.4   Motivation for this Thesis

Semi-analytic modelling allowed astronomers to successfully predict many properties of galaxy populations. However, the full picture of the formation and evolution of early-type galaxies is still not complete. In this Thesis we will try to answer

to the following questions: (i) Are the simple physical prescriptions used in the models capable of reproducing the tight correlations observed between certain galaxy properties for early-type galaxies? (ii) Which feedback mechanism is responsible for the suppression of the formation of bright galaxies? and (iii) In the context of the Millennium Simulation, how do we make mock catalogues at wavelengths where the standard GALFORM is inaccurate? Such models will allow new tests of the formation of spheroids (e.g. clustering of submillimeter galaxies).

In the semi-analytic model described by Cole et al. (2000), GALFORM, a simple method was used to calculate the sizes of galactic discs and bulges by assuming conservation of energy and angular momentum. Cole et al. (2000) tested their prescription for calculating the disc scalelengths against the observational estimates of De Jong & Lacey (2000) and found good agreement. However, Cole et al. did not test their method for predicting the sizes of spheroids, which follow a different scheme to that outlined for the discs. This is one of the aims of this Thesis. At first sight, the tight kinematic–photometric relations found empirically for spheroids would appear to pose a challenge to semi-analytic models of galaxy formation due to the heterogeneity of merger histories (e.g. Bell et al., 2004; Bernardi et al., 2003a, 2005).

The bright end of the luminosity function is dominated by early-type galaxies with passively evolving stellar populations. In the context of hierarchical galaxy formation it is known that the models tend to overpredict the number of bright galaxies. Different mechanisms have been proposed to regulate the formation of bright galaxies (e.g. Bower et al., 2001; Benson et al., 2003; Maller & Bullock, 2004; Bower et al., 2006). In fact, there are actually two variants of the GALFORM model proposed by Cole et al. (2000) which use different mechanisms to suppress the formation of bright galaxies (among other differences): the Baugh et al. (2005) model invokes a wind which ejects cold gas from galaxies, limiting the amount of gas available for star formation; in the model proposed by Bower et al. (2006), the cooling of the gas is suppressed due to the heating of the halo gas by the energy released by the accretion of matter on to a black hole. In this Thesis we test both these feedback mechanisms by predicting the properties of a special set of galaxies, the

luminous red galaxies (Eisenstein et al., 2001). These bright, massive galaxies are much rarer and luminous than L_* galaxies, representing a stern challenge for the models. They are especially important from the point of view of understanding and developing the model physics, since their abundance and nature probe the regime where the models are currently most uncertain.

It is important to be able to extend the predictions of the models so that we can carry out further tests between these scenarios. One area yet to be exploited is the spatial distribution of galaxies selected by their submillimeter emission, for example. To make an accurate calculation of the submillimeter emission, we need to combine GALFORM with a sphectrophotometric code, GRASIL (Silva et al., 1998). Running GRASIL is time-consuming and prohibitively expensive (and, indeed, even infeasible) for a large cosmological volume. Therefore, in the final chapters of the Thesis, we devise a novel application of artificial neural networks to predict the results of running GRASIL. This allows us to rapidly populate a simulation volume with galaxies selected at wavelengths which GALFORM finds challenging to calculate. We show some applications of this technique.

Chapter 2

The Semi-Analytical Model of Galaxy Formation: GALFORM

In this Thesis, we use a semi-analytical galaxy formation model, GALFORM, to predict the formation and evolution of galaxies. A comprehensive overview of the GALFORM model and the philosophy behind semi-analytical modelling can be found in Cole et al. (2000) (see also the review by Baugh, 2006). Important extensions to the model are described in Benson et al. (2002) and Benson et al. (2003). In this Chapter, we give a brief outline of the two most recently published models, Baugh et al. (2005) and Bower et al. (2006), which we will use to predict the properties of galaxies within the framework of the CDM structure formation model.

2.1 Dark Matter Haloes

2.1.1 Halo Merger Trees

Hierarchical galaxy formation models assume that galaxies form inside dark matter haloes, and their evolution is connected to the evolution of the underlying dark matter structure (White & Rees, 1978). Therefore, it is crucial to have information about the abundance, formation history and properties of dark matter haloes. There are two general approaches used to generate dark matter halo histories: (i) using a Monte Carlo method, based on an extended version of the theory introduced by Press & Schechter (1974), or (ii) extracting the formation histories from an N-body simulation.

(i) Monte Carlo dark matter halo trees: In this process, the halo merger history is calculated using a Monte Carlo algorithm, which randomly generates a formation path for the haloes. After defining a halo mass at specified redshift, a random walk is performed in the halo-mass space using the following distribution (Bond et al., 1991; Bower, 1991):

$$f_{12}(M_1, M_2)dM_1 = \frac{1}{\sqrt{2\pi}} \frac{\delta_{c1} - \delta_{c2}}{(\sigma_1^2 - \sigma_2^2)^{3/2}} \exp\left(-\frac{(\delta_{c1} - \delta_{c2})^2}{2(\sigma_1^2 - \sigma_2^2)}\right) \frac{d\sigma_1^2}{dM_1}dM_1, \qquad (2.1)$$

where $f_{12}(M_1, M_2)dM_1$ represents the fraction of mass M_2, at time t_2, which, at time t_1, was in haloes of mass between M_1 and $M_1 + dM_1$; σ_1 and σ_2 are the linear theory rms overdensities in spheres enclosing mass M_1 and M_2; δ_{c1} and δ_{c2} are the critical thresholds on the linear theory overdensity for collapse at times t_1 and t_2. $\sigma(M)$ is calculated from the power spectrum of the density fluctuations, and δ_c is calculated from the cosmological model assumed (Eke et al., 1996). This formula results from an extension of the Press & Schechter (1974) theory (see Lacey & Cole, 1993, for further details).

Also worth mentioning is the fact that, for each halo, we only allow two progenitors, i.e. we follow a binary tree formalism; this can be achieved by considering a sufficiently small timestep (Cole et al., 2000, gives a detailed description of the procedure).

One important assumption of the Monte Carlo approach is that there is no correlation between the formation history of the haloes and their environment. Recently, it was shown that this is not totally true: using a high resolution simulation of dark matter haloes, the Millennium Simulation, Gao et al. (2005) found that there is a dependence of the clustering on halo formation time. Furthermore, other authors showed that there is also a weak correlation between the clustering and halo concentration (Wechsler et al., 2006; Gao & White, 2007; Angulo et al., 2008). Nevertheless, the Monte Carlo method is accurately enough for most of the applications (see Baugh, 2006, for further details). In a recent paper, Parkinson et al. (2008) (see also Cole et al., 2008) presented a modified version of the Monte Carlo algorithm to generate halo merger trees in accordance with the merger trees extracted from the Millennium Simulation, however it does not take into account any environmental effect.

The Baugh et al. model uses this process to create the merger histories of dark matter haloes.

(ii) Dark matter halo trees extracted from a N-body simulation: Traditionally, extracting merger trees from N-body simulations was very challenging. Not only due to the complexity of the process, but also because it consumed a lot of computing power to probe the parameter space and the mass resolution of haloes was limited. Nowadays, observational estimates of cosmological parameters have narrowed down the range of the cosmological parameter space, and the computing power has continually increased. This has made it feasible to run N-body simulations with sufficient volume and mass resolution to allow merger trees to be extracted for representative samples of haloes. The state of art is the Millennium Simulation of the hierarchical clustering of dark matter in a ΛCDM universe (Springel et al., 2005). The simulation covers a volume of $0.125h^{-3}\mathrm{Gpc}^3$, and the mass resolution of the trees extracted from the simulation is $1.72 \times 10^{10}h^{-1}M_\odot$, which corresponds to approximately 20 particles. The Bower et al. model utilizes the halo merger histories from this N-body simulation.

The use of N-body merger trees has some advantages over the Monte Carlo approach. With the N-body approach, environment influences on the formation histories of galaxies are followed. Also, it is simple to model galaxy clustering and, perhaps, even galaxy mergers in this case. Nevertheless, the group finding methods used to define gravitationally bound structures in N-body simulations are not always successful, particularly for small haloes. Furthermore, as mentioned previously, they suffer from limited halo mass resolution. In the case of GALFORM, the typical mass resolution used in the Monte Carlo trees approach (Baugh et al. model) is a factor of three better than that of the trees drawn from the Millennium, because all of the computer memory can be devoted to a single tree.

Helly et al. (2003) compared model predictions obtained using Monte Carlo and N-body merger trees and reached the conclusion that the results are very similar above some luminosity which is set by the mass resolution, with the two prescriptions giving divergent answers for the luminosity function at faint luminosities. For the resolution of the Millennium, the predictions for the luminosity function are robust

down to around three magnitudes fainter than the break in the luminosity function.

2.1.2 Halo Density Profile

We assume that all haloes have spherical symmetry and a "NFW" density profile (Navarro, Frenk, & White, 1997) given by:

$$\rho(r) = \frac{\Delta_{\text{vir}}\rho_{\text{crit}}}{f(a_{\text{NFW}})} \frac{1}{r/r_{\text{vir}}(r/r_{\text{vir}} + a_{\text{NFW}})^2}, \tag{2.2}$$

defined for $r \leq r_{\text{vir}}$; $f(a_{\text{NFW}}) = \ln(1 + 1/a_{\text{NFW}}) - 1/(1 + a_{\text{NFW}})$, and Δ_{vir} is the mean density within the virial radius in units of the critical density, $\rho_{\rho crit}$ –both defined by the adopted cosmology. The free parameter a_{NFW} establishes the inverse of the concentration of the halo, and it was shown to be correlated with the halo mass (e.g. Navarro, Frenk, & White, 1997), although with considerable scatter (Jing, 2000).

2.1.3 Spin Distribution and Halo Rotation Velocity

As dark matter structures evolve, tidal torques will be created due to gravitational inhomogeneities. This process induces the increase of the angular momentum of the halo, and causes it to spin. It is common to quantify the spin by a dimensionless parameter:

$$\lambda_{\text{H}} = \frac{J_{\text{H}}|E_{\text{H}}|^{1/2}}{G M_{\text{H}}^{5/2}}, \tag{2.3}$$

where G is Newton's gravitational constant, J_{H}, M_{H} and E_{H} are the total angular momentum, mass and total energy of the halo. Cole & Lacey (1996) showed that the distribution of the halo spin parameter, drawn from a N-body simulation of dark matter, is well described by a log-normal:

$$P(\lambda_{\text{H}})d\lambda_{\text{H}} = \frac{1}{\sqrt{2\pi}\,\sigma_\lambda} \exp\left(-\frac{(\ln\lambda - \ln\lambda_{\text{med}})^2}{2\sigma_\lambda^2}\right) \frac{d\lambda_{\text{H}}}{\lambda_{\text{H}}}, \tag{2.4}$$

with a median $\lambda_{\text{med}} = 0.039$ and a standard deviation of $\sigma_\lambda = 0.53$; these values show little dependence on either the halo mass or the shape of the power spectrum.

For both the Baugh et al. and Bower et al. models, we randomly assign a spin parameter to each newly formed halo, using the previous equation. In a recent study using the Millennium Simulation data, Bett et al. (2007) showed that extracting the

spin parameter from a N-body simulation is not advisable due to its correlation with the halo definition and the noise in low mass haloes.

The spin parameter is an important quantity in the sense that it is a direct measurement of the angular momentum and energy of the halo, which will determine the angular momentum of the baryonic gas during the cooling process.

In order to calculate the angular momentum of the cooling gas, we need to determine the rotational velocity of the halo, $v_{\rm rot}$. To do so, we start by assuming that the mean rotational velocity is constant with radius, for concentric shells of halo material (see Warren et al., 1992; Cole & Lacey, 1996). This leads to:

$$v_{\rm rot} = A(a_{\rm NFW})\lambda_{\rm H}V_{\rm H}. \tag{2.5}$$

Here $V_{\rm H} = \sqrt{GM/r_{\rm vir}}$ is the circular velocity of the halo, as measured at the virial radius, and the dimensionless coefficient $A(a_{\rm NFW})$ is weakly correlated with $a_{\rm NFW}$ ($A \approx 3.9$ for $a_{\rm NFW} = 0.01$, and $A \approx 4.5$ for $a_{\rm NFW} = 0.3$; see Appendix A of Cole et al. 2000).

2.2 Gas Dynamics

In the GALFORM model we assume that the galactic discs form by cooling of the diffuse gas in the halo. While the halo collapses, the gas –which we designate by "hot" gas– will shock and heat to the virial temperature of the halo, defined by:

$$T_{\rm vir} = \frac{\mu\, m_{\rm H}\, V_{\rm H}^2}{2k_{\rm B}}, \tag{2.6}$$

where $k_{\rm B}$ is the Boltzmann constant, μ is the mean molecular mass and $m_{\rm H}$ is the mass of the hydrogen atom. Motivated by high-resolution hydrodynamical simulations of the formation of galaxy clusters (e.g. Frenk et al., 1999), we assume that the density distribution of the hot gas in the halo is less centrally concentrated than the NFW dark matter density profile. The gas still settles into a spherically symmetric distribution and we use a β-model, which is traditionally adopted for X-ray emitting gas in galaxy clusters (Cavaliere & Fusco-Femiano, 1974), to describe its density profile:

$$\rho_{\rm gas}(r) \propto \frac{1}{r^2 + r_{\rm core}^2}. \tag{2.7}$$

Here, we assume a constant temperature equal to the virial temperature of the halo, for simplicity (Frenk et al., 1999, found that this is a good approximation). In the Bower et al. model, r_{core} is kept fixed at 0.1 of the virial radius. In the Baugh et al. model, the initial value of the core radius is set to $r_{core} = r_{NFW}/3$. We expect this value to change with time, because cooling will convert the hot gas from the central parts of the halo, with the lowest entropy, into galactic gas (the efficiency of the cooling process is determined by the density and the entropy of the gas: the densest gas with the lowest entropy will cool first). Ultimately, haloes at later times have a larger core radius. To simulate this effect, we devised a simple model in which the core radius is adjusted according to the fraction of hot baryons (baryons in the form of hot gas) in the halo (as compared to the universal baryon fraction, Ω_b/Ω_0), in order to give the same density at r_H as if no gas had cooled.

The cooling time for the gas in the halo is given by the ratio between the energy per unit mass of gas, and the rate at which it is radiating energy, i.e.

$$\tau_{cool}(r) = \frac{3}{2} \frac{\bar{\mu} m_p \, k_B \, T_{gas}}{\rho_{gas}(r) \, \Lambda(T_{gas}, Z_{gas})}, \qquad (2.8)$$

where T_{gas} and Z_{gas} are the temperature and metallicity of the gas, respectively; $\Lambda(T_{gas}, Z_{gas})$ is the cooling function, tabulated by Sutherland & Dopita (1993). $\bar{\mu} m_p$ is the mean particle mass. From the cooling time we can define a cooling radius, $r_{cool}(t)$, as the radius at which we have $\tau_{cool} = t$.

As mentioned previously, as the gas cools, it settles into a disc at the center of the halo. We also take into account the time that the newly cooled gas takes to be accreted by the disc –the free-fall time. Similarly, we define a free-fall radius, r_{ff}, as the radius beyond which the material has not had time to fall into the galactic disc, i.e., the radius at which we have $t_{ff} = t$, with:

$$t_{ff}(r) = \int_0^r \left[\int_r^y -\frac{2 \, G \, M(x)}{x^2} dx \right]^{-1/2} dy. \qquad (2.9)$$

To calculate the total gas mass that cools and is accreted onto the disc during a timestep Δt, we estimate the mass of the hot gas enclosed by the spherical shell defined by the radius $r_{min}(t) = \min[r_{cool}(t), r_{ff}(t)]$ and $r_{min}(t + \Delta t) = \min[r_{cool}(t + \Delta t), r_{ff}(t + \Delta t)]$, and equate it to $\dot{M}_{cool}\Delta t$; this defines the cooling rate, $\dot{M}_{cool}$, which

is an important quantity for estimating the star formation rate, metal enrichment and feedback.

It is important to note that we assume that angular momentum is conserved by gas during the cooling process. Consequently, the specific angular momentum of the new gas added to the existing disc is equal to the angular momentum of the gas originally within the spherical shell defined by $r_{min}(t)$ and $r_{min}(t + \Delta t)$. Libeskind et al. (2007) tested this non-trivial assumption, using numerical hydrodynamical simulations of galaxy formation, and found that the conservation of angular momentum is likely to be achieved if the halo gas is only allowed to cool by radiative processes. Furthermore, the authors found that there is a good alignment between the angular momentum of the galaxies and their dark matter host halo (see Libeskind et al., 2007, for further details).

In the Baugh et al. model, the process of cooling of hot gas is locked during the lifetime of the halo, in the sense that the cooling of additional gas, accreted through infalling haloes or gas that is returned to the halo by processes of feedback, is not allowed until the next step in the halo merger tree. In the Bower et al. model, the cooling calculation is improved by allowing the cooling of gas reheated from galactic stellar feedback (see Bower et al., 2006, for further details).

A final note on the cooling process: in the GALFORM model, the cooled gas is accreted by the central galaxy in the halo and not by any satellite galaxies which may share the same halo. Font et al. (2008) have recently considered an alternative model in which the hot gas is not stripped completely from haloes as they fall into a larger halo. This produces little change in the galaxy LF but does have an impact on the colour of the satellite galaxies.

2.3 Disc and Star Formation

As described in the previous section, once the hot gas cools, it settles into a rotationally supported gas disc, for which we can calculate the size and circular velocity, knowing the angular momentum of the infalling material and applying centrifugal equilibrium. The process of star formation then starts. There are important effects

that every semi-analytic model has to consider: the end of massive stars, supernovae (SNe) which will eject energy and metals into the interstellar medium (ISM) and will, eventually, cause some of the disc gas to be expelled into the halo component, and enrichment the ISM with metals, which happens before the gas is expelled into the halo.

2.3.1 Star Formation

In the GALFORM model we assume that the instantaneous star formation rate, ψ, is:

$$\psi = \frac{M_{\text{cold}}}{\tau_*}. \tag{2.10}$$

Here τ_* is the star formation timescale, which is related to the circular velocity of the galaxy disc. Bower et al. follow the same prescription as in Benson et al. (2003), i.e., the star formation timescale scales with the dynamical time of the disc, τ_{disc}:

$$\tau_* = \varepsilon_*^{-1} \tau_{\text{disc}} \left(\frac{V_{\text{disc}}}{200 \, \text{km s}^{-1}} \right)^{\alpha_*}, \tag{2.11}$$

where ε_*, α_* are dimensionless parameters set to 0.0029 and -1.5 respectively, and $\tau_{\text{disc}} = r_{\text{disc}}/V_{\text{disc}}$, with r_{disc} being the half-mass radius of the disc and V_{disc} the circular velocity of the disc, measured at half-mass radius (see Benson et al., 2003, for a detailed description of these parameters). In the Baugh et al. model, a slightly different law is adopted:

$$\tau_* = \tau_{*0} \left(\frac{V_{\text{disc}}}{200 \, \text{km s}^{-1}} \right)^{\alpha_*}, \tag{2.12}$$

with $\tau_{*0} = 8$ Gyr and $\alpha_* = -3$. This changes were driven by the necessity of having more gas available at high redshifts to fuel starbursts (see Baugh et al., 2005, for details).

2.3.2 SNe Feedback

The reheating of cooled gas from winds from hot stars and SNe is modelled by:

$$\dot{M}_{\text{eject}} = \beta \, \psi. \tag{2.13}$$

The efficiency of the feedback process, here defined by β, is defined as:

$$\beta = \left(\frac{V_{\text{disc}}}{V_{\text{hot}}} \right)^{-\alpha_{\text{hot}}}. \tag{2.14}$$

V_{hot} has the dimensions of velocity and α_{hot} is dimensionless; these parameters are set to 300 km s^{-1} and 2 respectively, in the Baugh et al. model, and to 485 km s^{-1} and 3.2, in the Bower et al. model.

2.3.3 Chemical Enrichment and Gas Equations

As mentioned, a correct modelling of the process of star formation needs to take into account the fraction of metals expelled by the high mass stars into the ISM (through SNe and stellar winds). We adopt an instantaneous recycling approximation to model the enrichment of the gas. The change in mass and metals of the gas is given by the following equations:

$$\dot{M}_* = (1 - R)\,\psi, \tag{2.15}$$

$$\dot{M}_{hot} = -\dot{M}_{cool} + \beta\,\psi, \tag{2.16}$$

$$\dot{M}_{cold} = \dot{M}_{cold} - (1 - R + \beta)\,\psi, \tag{2.17}$$

$$\dot{M}_*^Z = (1 - R)\,Z_{cold}\,\psi, \tag{2.18}$$

$$\dot{M}_{hot}^Z = -\dot{M}_{cool}\,Z_{cold} + \beta\,Z_{cold}\,\psi, \tag{2.19}$$

$$\dot{M}_{cold}^Z = \dot{M}_{cool}\,Z_{hot} + [p - (1 + \beta - R)\,Z_{cold}]\,\psi. \tag{2.20}$$

Here, the metallicity of the cold/hot gas is:

$$Z_{cold/hot} = \frac{M^Z_{cold/hot}}{M_{cold/hot}}. \tag{2.21}$$

M^Z is the mass in metals, R is the recycled fraction (the fraction of mass returned to the ISM per unit mass of stars formed), and p is the yield (the mass of metals returned to the ISM per unit mass of stars formed); these parameters are set once we define an initial mass function (IMF). The cooling rate, $\dot{M}_{cool}$, and the hot gas metallicity, Z_{hot}, are assumed to be constant over a given timestep of the merger tree.

2.3.4 Additional Feedback Processes

The two models use different feedback mechanisms to counter the overproduction of bright galaxies, which was a long-standing problem for hierarchical models (see
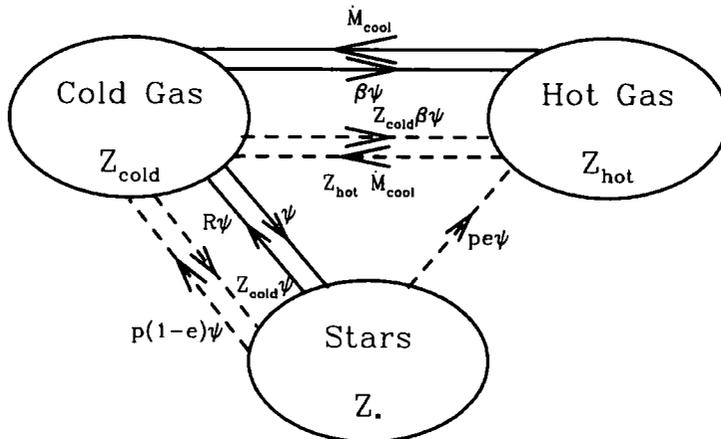
Figure 2.1: Schematic representation of the transfer of mass and metals between three phases: hot gas, cold gas and galactic disc and stars in the disc. Source: Fig. 3, Cole et al. (2000).

Baugh, 2006). Baugh et al. invoke a wind which ejects cold gas from galaxies at a rate which is a multiple of the star formation rate (see Benson et al., 2003). The gas thus ejected is not allowed to recool, even in more massive haloes. This is another "channel" for the energy released by supernovae to couple to the cold gas reservoir available for star formation which operates alongside the feedback mechanism described in Subsection 2.3.2. The *superwind* and the SNe feedback have distinct parameterizations in terms of the star formation rate, and differ in the fate of the reheated gas, as discussed above. The effect of the feedback from *superwind* is modelled using:

$$\dot{M}_{\text{eject}} = \beta_{\text{sw}}\,\psi, \tag{2.22}$$

where the efficiency is given by:

$$\beta_{\text{sw}} = f_{\text{sw}} \min[1, (V_{\text{c}}/V_{\text{sw}})^{-2}], \tag{2.23}$$

with $f_{\text{sw}} = 2$ and $V_{\text{sw}} = 200$ km s^{-1} (see Lacey et al., 2008, for further discussion). There is observational evidence for superwind outflows in the spectra of Lyman-break galaxies and in local starburst galaxies (Adelberger et al., 2003; Wilman et al., 2005).

In the Bower et al. model, the cooling of gas in quasi-static hot gas haloes is

suppressed, effectively cutting off the "fuel supply" for star formation. These are haloes in which the cooling time of the gas exceeds the free fall time within the halo, and the central AGN has sufficient power to quench the cooling flow (this condition will be true if the AGN power, defined as a fraction of the Eddington luminosity, is greater than the cooling luminosity, L_{cool}), by heating the halo gas. We also model the growth of the black hole, due to the accretion from the quiescent inflow of hot gas, using the model described by Malbon et al. (2007): $\dot{M}_{BH} = L_{cool}/(0.2c^2)$.

2.4 The Formation of Spheroids

In our model, the formation of spheroids takes place through two channels: galaxy mergers and instability of galactic discs.

2.4.1 Dynamical Friction

Whenever two haloes merge, we assume that the most massive galaxy becomes the central galaxy of the new halo. All the other pre-existing galaxies, i.e. satellite galaxies, are expected to orbit around the central galaxy, and may eventually merge. This is because satellite galaxies will lose energy and angular momentum due to the interaction with the dark matter component, via dynamical friction. We assume that the satellite galaxies enter the new halo on a random orbit, and merge with the central galaxy on a timescale given by (Chandrasekhar, 1943):

$$\tau_{mrg} = f_{df}\,\Theta_{orbit}\,\tau_{dyn}\,\frac{0.3722}{\ln\Lambda_{Coulomb}}\,\frac{M_H}{M_{sat}}, \qquad (2.24)$$

where M_{sat} is the mass of the satellite galaxy and its halo, $\tau_{dyn} = \pi\,r_{vir}/V_H$ is the dynamical time of the new halo, and the Coulomb logarithm is $\ln\Lambda_{Coulomb} = ln(M_H/M_{sat})$. f_{df} is a free parameter which is used to take into account the uncertainties in the adopted model for the dynamical timescale or that M_{sat} is smaller than the original satellite halo mass due to stripping. We include all the orbital dependencies in the term Θ_{orbit},

$$\Theta_{orbit} = \left(\frac{J}{J_c(E)}\right)^{0.76}\left(\frac{r_c(E)}{r_{vir}}\right)^2. \qquad (2.25)$$

E and J are the initial energy and angular momentum of the satellite galaxy. $r_c(E)$ and $J_c(E)$ are the radius and the angular momentum respectively of a circular orbit with energy E. Tormen (1997), using N-body simulations, found that the distribution of the orbital parameter, Θ_{orbit}, is well described by a lognormal distribution, with $\langle \log \Theta_{orbit} \rangle = -0.14$ and a standard deviation of 0.26.

The implementation of this process in GALFORM is simple: when a new halo is formed, and after defining a central galaxy, we assign a random value of Θ_{orbit} to each of the satellite galaxies, according to the lognormal distribution defined above, and we calculate the merger timescale. Evidently, the satellite galaxy will only merge with the central galaxy if the merger timescale is shorter than the age of the halo; if this condition is not true then, as the halo merges with a larger system, the satellite galaxy will be incorporated into the newly formed halo and the process is repeated.

2.4.2 Galaxy Mergers and Bursts of Star Formation

In GALFORM, the consequences of a galaxy merger are characterized by the ratio, $R = M_{sat}/M_{central}$, of the mass of the accreted satellite galaxy (M_{sat}) to the mass of the primary or central galaxy in the halo ($M_{central}$), onto which the satellite is accreted. The mass ratio R is compared to two thresholds, f_{ellip} and f_{burst}, to establish the severity of the merger (see, for example, Bournaud et al., 2005). These thresholds are model parameters. If $R > f_{ellip} = 0.3$, then the galaxy merger is termed "major". In the case of major mergers, the disc of the primary galaxy is destroyed. All stars are transferred to the spheroid component and any cold gas present participates in a burst of star formation which adds stars to the spheroid. If the mass ratio of satellite to primary falls between the thresholds, i.e. $f_{burst} < R < f_{ellip}$, then the stellar disc of the primary survives and the stars from the accreted satellite are added to the spheroid. In this case, if the primary is also gas rich, that is if cold gas accounts for at least 75% of the total mass in the disc, then we assume that the accretion of the satellite induces an instability which drains the primary disc of cold gas, leading to a burst of star formation in the spheroid. This latter merger event is called "minor" merger.

In our model, a galaxy can move in either direction along the Hubble sequence (Baugh et al., 1996). The accretion of gas from the hot halo and subsequent quiescent star formation leads to a late-type (disc dominated) galaxy. A major merger between two such galaxies produces a descendent galaxy which jumps to the opposite end of the Hubble sequence, becoming an early-type galaxy (bulge dominated). Further accretion of cooling gas allows the galaxy to grow a new disc around its bulge, moving the galaxy back towards the late-type part of the sequence.

In addition to the merger mode of spheroid production, Bower et al. also consider the secular production of bulges from discs which are unstable due to their strong self-gravity (e.g. Efstathiou et al., 1982; Binney & Tremaine et al., 1987; Mo et al., 1998). We use a similar approach to Mo et al. (1998), and define a stability criterion:

$$\epsilon_m = \frac{V_c}{\sqrt{G\, M_{disc}/r_{disc}}}. \tag{2.26}$$

We set $\epsilon_m = 0.8$ for a disc to be unstable. If a given disc satisfies $\epsilon_m < \epsilon_{disc}$, then we assume that the disc is unstable and, consequently, will evolve into a bulge; all the stellar mass present in the disc will be transfered to the galaxy's bulge, and the cold gas (if any) will undergo a burst of star formation. This mode is most important in less massive galaxies.

2.4.3 The Scalelengths of the Disc and the Bulge Components

We assume that discs have an exponential profile, with a half-mass radius given by r_{disc}, and bulges have a $r^{1/4}$ profile in projection, with a half-mass in 3D given by r_{bulge}.

The scalelength of the disc is determined by the angular momentum of the cooled halo gas, as mentioned previously. With a simple calculation we can show that, if we assume that the gas and dark matter halo are described by an isothermal sphere, the mean specific angular momentum of the cool gas is: $j_{cool} = \pi/8\, r_{cool}\, V_{rot}^{gas} = \sqrt{2}\, \lambda_H\, r_{cool}\, V_H$ (see Appendix A of Cole et al., 2000, for details). The mean specific angular momentum of the disc is given by: $j_{disc} = 2\, h_D\, V_H$, for an exponential profile and ignoring the self-gravity. These two equations lead to $r_{disc} = 1.68\, h_D =$

$1.19\,\lambda_{\rm H}\,r_{\rm cool}$, which show the correlation between the disc radius and the halo spin parameter and radius of the cooling material. GALFORM uses an improved version of this simple calculation, in the sense that we use non-isothermal halo and gas profiles, an accurate distribution of halo spin parameters, and we take into account its self-gravity and the contraction of the dark matter halo due to the baryonic component.

Spheroids are formed in galaxy mergers or through disc instabilities as outlined in the previous Subsection. The size of the spheroid resulting from a galaxy merger, $r_{\rm m}$, is determined by applying the conservation of energy and the virial theorem (see Section 4.4.2 of Cole et al., 2000):

$$\frac{(M_1 + M_2)^2}{r_{\rm m}} = \frac{M_1^2}{r_1} + \frac{M_2^2}{r_2} + \frac{f_{\rm orbit}}{c}\frac{M_1\,M_2}{r_1 + r_2}, \tag{2.27}$$

where $M_{\rm i}$ represents the total mass (stellar, cold gas and dark matter) of one of the merging objects, within $r_{\rm i}$, and the form factor c and the parameter $f_{\rm orbit}$ are related to the self-binding energy and orbital energy by

$$E_{\rm bind} = -c\frac{G\,M_{\rm i}^2}{r_{\rm i}}, \tag{2.28}$$

$$E_{\rm orbit} = -\frac{f_{\rm orbit}}{2}\frac{G\,M_1\,M_2}{r_1 + r_2}. \tag{2.29}$$

For simplicity, we adopt $c = 0.5$ and $f_{\rm orbit} = 1$. In the Chapter 3, we explore the impact of varying $f_{\rm orbit}$ on the scaling relations of early-type galaxies.

Once the scale lengths of the disc and bulge components have been calculated as outlined above, we next take into account the self-gravity of the baryons and the contraction of the dark matter halo in response to the gravity of the condensed baryons. New radii are computed for the disc and bulge by applying an adiabatic contraction of the disc, bulge and dark matter components (Blumenthal et al., 1986; Jesseit et al., 2002). In summary, as gas cools and settles in the centre of the dark matter halo, it provides an additional gravitational force on the dark matter component, thereby increasing the density of dark matter in the central part of the halo. This, in turn, changes the gravitational force on the baryons due to the underlying dark matter. The degree of contraction is computed by knowing that, in a slowly varying potential, the action integral, $\oint p_{\rm i}dq_{\rm i}$, is an adiabatic invariant for each particle of mass i, where $p_{\rm i}$ is the conjugate momentum of the coordinate $q_{\rm i}$ (Barnes

& White, 1984; Blumenthal et al., 1986; Jesseit et al., 2002). The assumption of spherical symmetry and circular orbits, simplifies the action integral to the conservation of angular momentum in spherical shells, $rM(r)$. In the case of the disc, the total specific angular momentum is conserved. The bulge and dark matter halo are not rotationally supported. Nevertheless, it is useful to define an equivalent circular velocity using the velocity dispersion of each of these components, and, using this, to define a quantity which we refer to as a *pseudo-angular momentum*. For the bulge the pseudo-angular momentum is given by: $j_{\text{bulge}} = r_{\text{bulge}} V_c(r_{\text{bulge}})$. This quantity is conserved during the adiabatic contraction. A similar quantity is conserved for the dark matter (see Appendix C of Cole et al., 2000, for a detailed calculation and equations).

In the case of the secular growth of spheroids (in the Bower et al. model), we again apply the conditions of virial equilibrium and energy conservation as in Eq. 2.27:

$$\frac{c_{\text{B}}\,(M_{\text{disc}} + M_{\text{bulge}})^2}{r_{\text{new}}} = \frac{c_{\text{D}}\,M_{\text{disc}}^2}{r_{\text{disc}}} + \frac{c_{\text{B}}\,M_{\text{bulge}}^2}{r_{\text{bulge}}} + f_{\text{int}}\,\frac{M_{\text{disc}}\,M_{\text{bulge}}}{r_{\text{disc}} + r_{\text{bulge}}}. \tag{2.30}$$

Here, the factors $c_{\text{D}} = 0.49$ and $c_{\text{B}} = 0.45$ were adopted; the term that represents the gravitational interaction between the disc and bulge, f_{int} is set to 2. After the calculation of the radius of the new bulge, we readjust adiabatically the spheroid and halo terms to reach the new equilibrium.

2.5 Spectral Synthesis and Dust Extinction

To compute the spectral and photometric properties of the stellar population we use the model of Bruzual & Charlot (2003). This model predicts the spectral energy distribution (SED) for a single stellar population, $l_\lambda(Z, t)$, with a given metallicity, Z, and age, t. Once, this quantity is defined, the SED can be obtained by convolving l_λ with the star formation of the galaxy:

$$L_\lambda(t) = \int_0^t l_\lambda[t - x, Z(x)]\,\psi(x)\,dx, \tag{2.31}$$

where $Z(x)$ and $\psi(x)$ are the metallicity and formation rate of stars formed at $t = x$. An important point is that $l_\lambda(Z, t)$ depends on the initial mass function chosen. Both models, Baugh et al. and Bower et al., adopt a standard solar neighbourhood IMF,

the Kennicutt (1983) IMF, for quiescent star formation (i.e. the star formation in discs):

$$\frac{dN}{d\ln m} \propto m^{-x},$$

(2.32)

with $x = 0.4$ for $m < 1$ $M_\odot$. and $x = 1.5$ for $m > 1$ $M_\odot$; the recycled fraction and yield are then: $R = 0.41$ and $p = 0.02$. Bower et al. also use this IMF in starbursts, whereas Baugh et al.invoke a top heavy IMF (a power law with $x = 0$). This is the primary ingredient responsible for this model's successful reproduction of the number counts of sub-mm galaxies. This choice, has the effect of increasing the recycled fraction by a factor of 2.2, to $R = 0.91$, and increases the yield by a factor of 6.5, to $p = 0.15$. The choice of a top-heavy IMF in starbursts is controversial, but has been tested successfully against the metal content of the intra-cluster medium (Nagashima et al., 2005a) and the metallicity of elliptical galaxies (Nagashima et al., 2005b). In both models, a small fraction of star formation forms brown dwarfs, Υ, which, evidently, contribute with mass but no light to the stellar population (i.e. they will not contribute to the SED); it is defined as

$$\Upsilon = \frac{\text{mass in visible stars + brown dwarfs}}{\text{mass in visible stars}}.$$

(2.33)

According to the definition, brown dwarfs reduce the luminosities by a factor of $\sim 1/\Upsilon$. Moreover, both models adopt $\Upsilon = 1$.

Dust in the ISM will attenuate the luminosity of the galaxy, due to the absorption and scattering of starlight by dust grains. We calculate the effect of dust extinction using the models of Ferrara et al. (1999), which considers a three-dimensional smooth distribution of dust and stars in the galactic disc (in fact, it assumes that the dust has the same radial scalength, h_R, as the stars, but, generally, a different scaleheight, h_z). The amount of dust is specified by τ_{V0}, the central V-band optical depth. Cole et al. (2000) showed that for GALFORM this quantity can be calculated from:

$$\tau_{V0} = 0.043 \left(\frac{M_{\text{cold}}/(2\pi h_R^2)}{M_\odot \, \text{pc}^{-2}} \right) \left(\frac{Z_{\text{cold}}}{0.02} \right),$$

(2.34)

where Z_{cold} is the total mass of metals in the cold gas. As for the extinction curve, we adopt a Milky Way extinction curve as our standard choice. This simple calculation of the dust effect will be designated the *shortcut* calculation from hereafter.

2.5.1   Using the GRASIL Model

We also have the possibility of calculating the SED using a state of the art sphectrophotometric model, GRASIL (Granato et al., 2000). This model has been successful in reproducing the SEDs of quiescent and starbursts galaxies (Bressan et al., 2002), from the far-UV to radio. GRASIL has several advantages over the simple prescription described above, in the sense that it includes a comprehensive and self-consistent calculation of the radiative transfer of light through dust. The model uses a population synthesis model based on the Padova stellar evolution and Kurucz model atmospheres (Bressan et al., 1998; Silva et al., 1998), to calculate the SED of the stellar population. Its main characteristics are: (i) it is able to predict the SED for the galactic disc and bulge, separately; (ii) the cold gas and dust are treated independently, where some fraction of the cold gas is dense molecular clouds, embedded in the dust –the diffuse (cirrus) component; (iii) the dust is modelled as a mixture of graphite, silicate grains and Polycyclic Aromatic Hydrocarbon (PAH) molecules, with a distribution of sizes chosen to reproduce the emissivity and dust extinction curve of the local ISM; (iv) it incorporates radio emission from ionized HII regions and synchrotron emission from supernova remnants (see Granato et al., 2000; Baugh et al., 2005; Lacey et al., 2008, for further details).

In Chapter 3 and 4, we will use the simple prescription described at the beginning of this section, to determine the SEDs of model galaxies and its derived photometric quantities. This is because, for the luminosities that we are interested (r, g, and i optical bands), the dust emission and absorption are not dominant, and the Cole et al. (2000) prescription is sufficiently accurate.

In Chapter 5 we use GRASIL instead because our main objective is to reproduce, as accurately as possible, the entire spectral energy distribution of selected galaxies.

Chapter 3

The Structural and Photometric Properties of Early-Type Galaxies in Hierarchical Models

3.1 Introduction

Remarkably tight correlations exist between the structural and photometric properties of galaxies. Across the Hubble sequence there is a strong dependence of luminosity on either the rotation speed of galactic discs (Tully & Fisher, 1977) or the velocity dispersion of spheroids (Faber & Jackson, 1976). Other scaling relations observed for early-type galaxies include those between colour and magnitude (Sandage & Visvanathan, 1978a,b), colour and velocity dispersion (Bernardi et al., 2005), radius and luminosity (Sandage & Perelmuter, 1990), radius and surface brightness (Kormendy, 1977) and between morphology and density (Dressler, 1980). Some of these correlations can be combined into a "fundamental plane" which connects the effective radii, velocity dispersions and luminosities of ellipticals (Djorgovski & Davis, 1987; Dressler et al., 1987; Bernardi et al., 2003a).

The existence of these scaling relations and their tightness encode clues about the formation and evolution of elliptical galaxies. For example, the existence of a fundamental plane can be understood by applying the virial theorem to a gravitationally bound stellar spheroid in dynamical equilibrium, after making the assumption that ellipticals of different sizes have the same structure (homology) and a constant mass-

to-light ratio. The deviation of the observed fundamental plane from this prediction can therefore be driven by variations in the mass-to-light ratio across the early-type population or by a non-uniformity of the structure of ellipticals, referred to as structural non-homology, or a combination of these two effects (Ciotti et al., 1996; Bertin et al., 2002; Trujillo et al., 2004a). At first sight, the small scatter around the observed correlations would appear to pose a challenge to hierarchical galaxy formation models, since the variety of merger histories in the models would lead one to expect to a corresponding scatter in the properties of early-type galaxies.

In this chapter, we use GALFORM to model the properties of elliptical galaxies; we focus on the predictions of the model introduced by Baugh et al. (2005). The model predicts the star formation and merger histories of galaxies (for a review of this class of model, see Baugh, 2006). As discussed in Chapter 2, two channels are considered for the formation of spheroids: galaxy mergers or secular evolution of the disc. The first attempts to track the disc and bulge of a galaxy separately simply recorded the mass and luminosity in each component (Baugh, Cole & Frenk 1996; Kauffmann 1996). The models have now progressed to a state where detailed predictions can be produced for the structural properties of galaxies in addition to their stellar populations. Cole et al. (2000) introduced a model for the sizes of the disc and spheroid components of galaxies: the size of a galactic disc is calculated by assuming conservation of the angular momentum of the gas as it cools and collapses in the halo; the size of the spheroid is derived by applying conservation of binding and orbital energy, and by applying the virial theorem to the merging galaxies. The Cole et al. scheme also takes into account the gravitational force of the dark matter and the reaction of the dark matter halo to the presence of the baryons (see Chapter 2 for further details).

Cole et al. tested their model for the sizes of galactic discs against the observed distribution of disc scale lengths estimated by De Jong & Lacey (2000), and verified that the predictions of their fiducial model were in excellent agreement with the observations. However, the authors did not test their prescription for predicting the size of galactic spheroids. This is the focus of this chapter.

In similar studies, Hatton et al. (2003) used a scheme analogous to that outlined

by Cole et al. to compute the sizes of spheroids in the GALICS model. However, these authors adopted a less realistic model for the scale size of galactic discs. In common with many semi-analytical models, they assumed that the scale size of a disc is related to the virial radius of the host dark matter halo by $r_D = \lambda R_{200}/2$, where λ is the dimensionless spin parameter for the dark matter halo, which quantifies its angular momentum, and R_{200} is the halo virial radius. This ignores the self-gravity of the baryons and the contraction they produce in the central regions of the dark matter halo. Several papers have considered the origin of the fundamental plane and the role of gas-rich and gas-poor mergers using numerical simulations, which follow the dark matter and baryons (Kobayashi, 2005; Dekel & Cox, 2006; Robertson et al., 2006; Boylan-Kolchin et al., 2006).

We first compare our predictions to the sample of early-type galaxies drawn from the Sloan Digital Sky Survey (SDSS) by Bernardi et al. (2005): the selection criteria are described in Section 3.2 and the comparisons between our model predictions and the data are given in Section 3.3. In Section 3.4, we explore the sensitivity of our model predictions to various physical ingredients of the models. The evolution with redshift of the model predictions for the scaling relations is presented in Section 3.5. Our conclusions are given in Section 3.6.

3.2 The selection of an early-type sample

We first compare GALFORM model predictions against the scaling relations and statistics of the sample of early-type galaxies constructed from the SDSS DR2 (the Second Data Release) by Bernardi et al. (2003a, 2005) (n.b., this sample was provided by the authors). These authors measured relations between luminosity and various properties of early-type galaxies such as velocity dispersion, effective radius, effective mass, effective density and surface brightness (Bernardi et al., 2003b, 2005). The sample was also used to measure the luminosity function of early-type galaxies (Bernardi et al., 2003b), the fundamental plane (Bernardi et al., 2003c) and the colour-magnitude/colour-velocity dispersion relations (Bernardi et al., 2003d, 2005).

For a complete description of the construction of the early-type sample from the

SDSS, we refer the reader to the above papers. Below we give a summary of the selection criteria applied by Bernardi et al. (2005) (hereafter Ber05). Galaxies are included in the sample if they have:

1. Redshift $z \leq 0.3$, with a median $z_{med} = 0.13$.

2. Apparent r-band Petrosian magnitudes in the range $14.5 < r_{petro} < 17.75$ mag.

3. `eclass` < 0. The `eclass` value is a classification of the spectral type of a galaxy derived from a principal component analysis decomposition of its spectrum (Connolly & Szalay, 1999). Ber05 chose negative values of `eclass` as this corresponds to spectra in which absorption lines dominate, characteristic of early-type galaxies.

4. `fracDev`> 0.8, computed using the $r-$band image. The value of `fracDev` is an indicator of morphology. It is calculated in two steps. First, the best fit de Vaucouleurs and exponential profiles to the galaxy image are found. Second, using the scale lengths of the best fit profiles found in step one, the best fit linear combination of the disc and bulge profiles is derived. The contribution of the de Vaucouleurs profile to this linear combination is the value of `fracDev`.

We attempt to reproduce these selection criteria by imposing the following conditions on GALFORM galaxies:

1. We generate a population of galaxies at an output redshift of $z = 0.13$, the median redshift of the Ber05 sample. In our comparisons, we consider only Ber05 galaxies that lie within the redshift interval $0.11 < z \leq 0.15$, close to the median redshift. This additional selection in redshift reduces the size of the observational sample by a factor of ~ 4 to ~ 11000 objects.

2. We use total magnitudes to select a sample of model galaxies. We apply the same apparent magnitude limits which are used for the data. This is a reasonable approach as the difference between total and Petrosian magnitudes is typically smaller than 0.2 mag (Graham et al., 2005). We have also computed Petrosian magnitudes for our model galaxies and find that using the Petrosian

magnitudes in place of total magnitudes does not make a significant difference to our results.

3. At present, GALFORM does not produce spectra with absorption line features. Therefore we cannot directly calculate a value for the spectral parameter `eclass`. Instead, we use the $g - r$ colour which is more readily predicted for model galaxies. In Fig. 3.1, we use the SDSS DR4 to show that there is a good correlation between $g - r$ colour and `eclass`. We retain galaxies with $g - r \geq 0.8$. The Figure shows that more than 95% of the galaxies with a negative value for `eclass` are selected by this colour cut.

4. We compute the value of `fracdev` in the a similar way as was done for the SDSS galaxies: we assume that the model bulges follow a de Vaucouleurs profile and the discs an exponential profile; these profiles describe the distribution of stellar mass, and so are independent of the passband. We adopt a cut on `fracdev` > 0.8. In Section 3.3.1, we explore the impact on our predictions of replacing the cut in the value of `fracdev` with a simple cut on the bulge-to-total luminosity ratio of the model galaxies.

5. Due to limitations of the SDSS data we also set a surface brightness threshold, $\mu_e < 24.5$ mag arcsec^{-2}.

3.3 Results

In this section, we compare the predictions of the Baugh et al. (2005) model with observational data for early-type galaxies derived from the SDSS sample of Ber05.

3.3.1 The luminosity function of early-type galaxies

The luminosity function is perhaps the most fundamental statistical description of the galaxy population. Later on, we perform fits to the fundamental plane of early-type galaxies in our model. The results for this fit are sensitive to the abundance of galaxies as a function of luminosity, so it is imperative that the model reproduces the observed luminosity function closely.
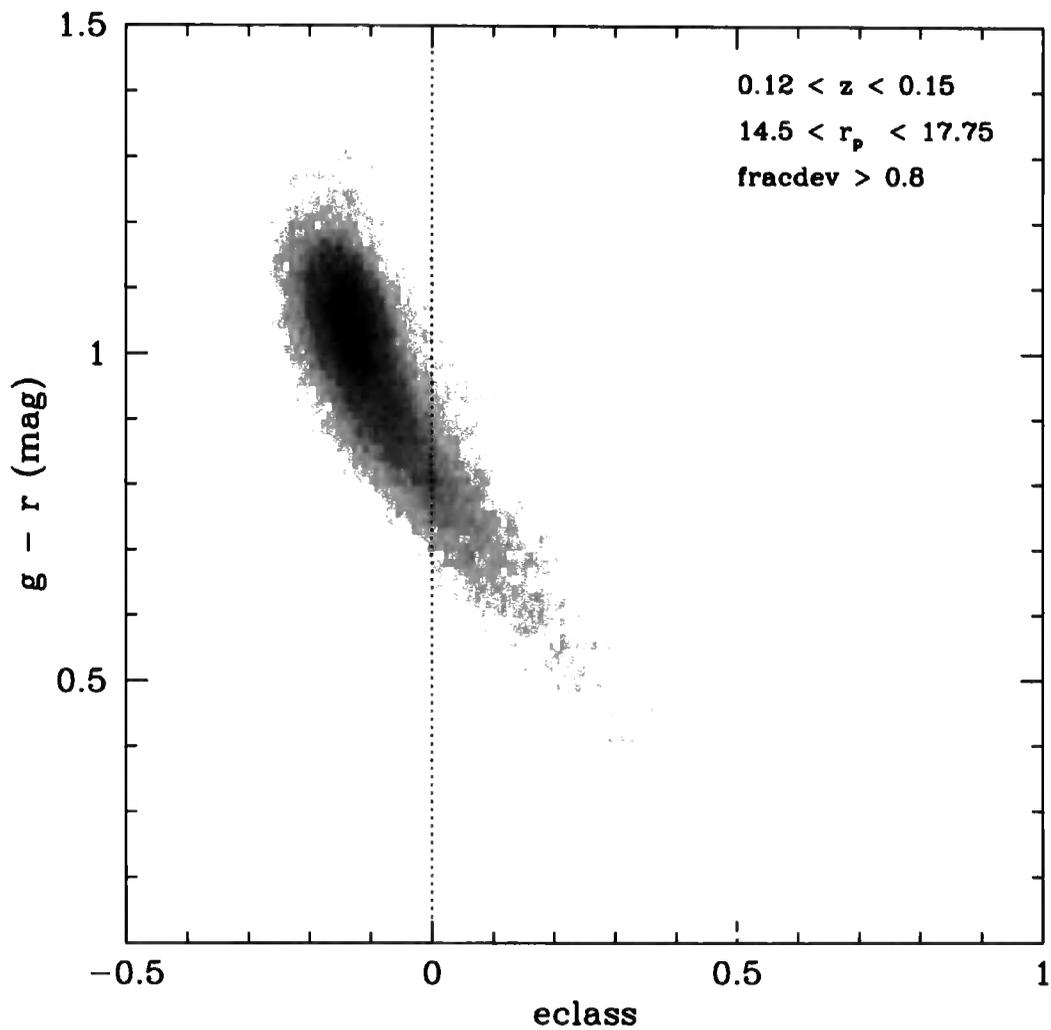
Figure 3.1: The relation between $g - r$ colour and `eclass`, for a sample of galaxies selected from SDSS DR4 in the redshift range $0.12 < z < 0.15$ and with Petrosian magnitudes in the interval $14.5 < r_{petro} < 17.75$ mag. The dotted line represents the selection applied to the observational data, `eclass` < 0, by Bernardi et al. (2005): more than 95% of these galaxies have $g - r > 0.8$.
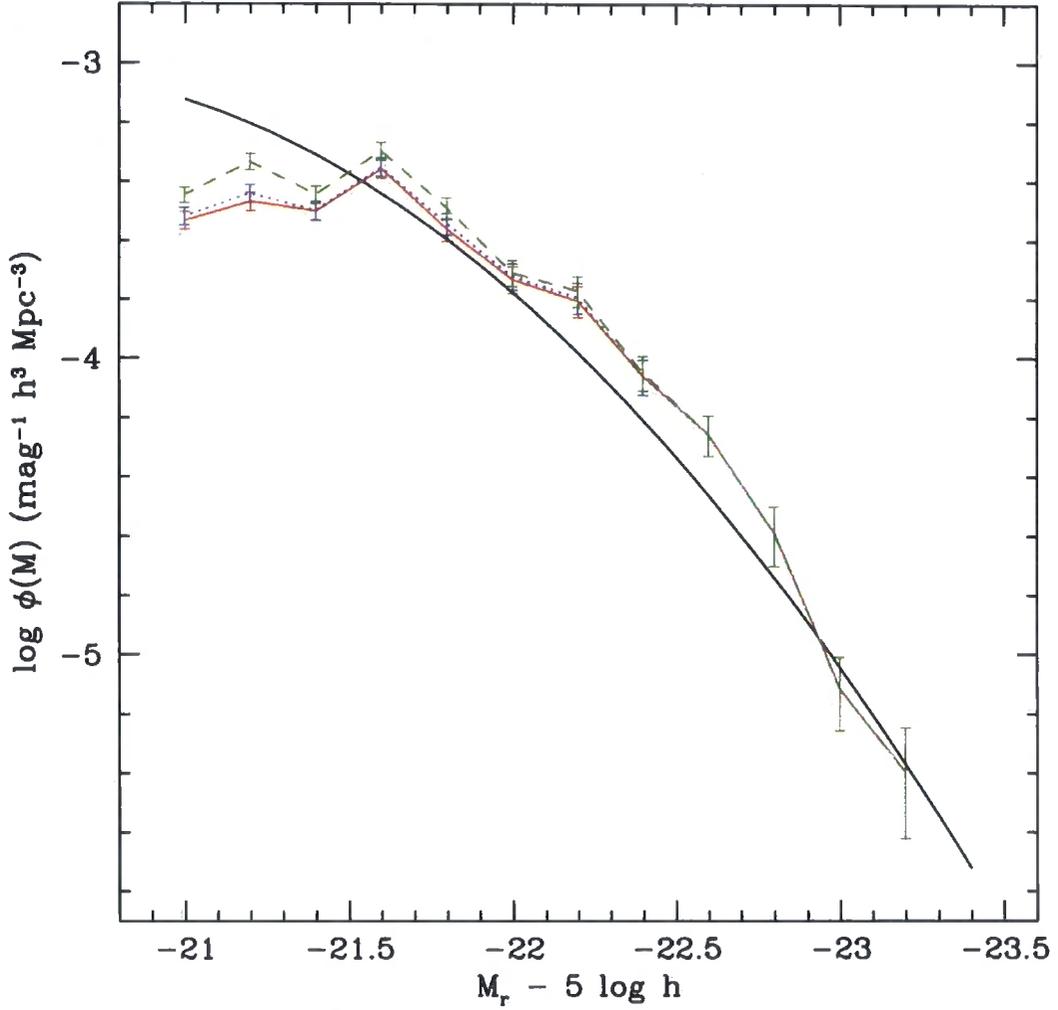
Figure 3.2: The luminosity function of early-type galaxies at $z = 0.13$. The solid black line shows a fit to the luminosity function of the SDSS sample of Bernardi et al. (2005).The results for the GALFORM sample are plotted using coloured lines. The solid red line shows our standard early-type galaxy selection, as outlined in Section 3.2. The errorbars show Poisson errors due to the finite number of galaxies simulated. The other lines show how the luminosity function varies when, instead of using `fracdev` > 0.8, the bulge-to-total r-band luminosity ratio is used; the dashed green line shows the results for $B/T_r > 0.5$ and the dotted blue line for $B/T_r > 0.8$.

In Fig. 3.2, we compare the predictions of the GALFORM model for the r-band luminosity function of early-type galaxies with the estimate from the SDSS sample of Ber05. The luminosity function of SDSS early-types is well described by a Gaussian form: $\phi(M)\,dM = \phi_*/\sqrt{2\pi\sigma^2}\exp\{-(M - M_* + Qz_i)^2/(2\sigma^2)\}$, where $(\phi_*, M_*, \sigma, Q) = (1.99 \times 10^{-3}\mathrm{Mpc}^{-3}, -21.15, 0.841, 0.85)$ respectively (note, Ber05 assume $h = 0.7$). The model predictions are in reasonably good agreement with the luminosity function estimated from the data. We note that, in the this and the next Chapter, we do not use any statistical test of goodness of fit. The main reasons are: (i) our parameters are physical and not from some function that has been designed to represent an observation e.g. the Schechter function; (ii) the parameters affect more than one prediction, so really we are comparing the model against a range of observations; (iii) this brings up the issue of how to combine the errors from different datasets (which data set is more important?).

We also show, in Fig. 3.2, the effect of changing the criteria used to select early-type galaxies in the model. In our standard selection, the primary indicator of morphology is `fracdev` (see Section 3.2). We have also explored using the bulge-to-total luminosity ratio in the r-band, B/T_r, in place of `fracdev` (Baugh et al., 1996). The results for cuts of $B/T_r > 0.8$ and $B/T_r > 0.5$ are shown by the dotted blue and dashed green lines respectively. The luminosity function derived using $B/T_r > 0.8$ is remarkably similar to the one obtained using `fracdev` > 0.8 (shown by the solid red line).

In our subsequent comparisons with the SDSS sample of early-types, we assign each model galaxy a weight which depends upon luminosity, such that the luminosity function of early-types in the model is forced to reproduce exactly the luminosity function of the data.

3.3.2 The Faber-Jackson and σ-age relations

The Faber-Jackson (hereafter FJ) relation was one of the first scaling relations to be discovered for early-type galaxies (Faber & Jackson, 1976). This relation indicates that luminosity is a strong function of velocity dispersion, σ, with brighter early-types displaying larger velocity dispersions. Observational studies suggest that this

relation is approximately given by $L \propto \sigma^4$: Forbes & Ponman (1999), using a local sample of early-type galaxies, found $L_B \propto \sigma^{3.9}$ in the B-band, while Pahre et al. (1998) reported $L_K \propto \sigma^{4.1}$ in the K-band. These results also indicate that the FJ relation is essentially independent of wavelength. In the case of SDSS early-type galaxies, Bernardi et al. (2003b) confirmed these earlier results, finding $L_r \propto \sigma^{3.91}$ in the r-band at $z = 0$, with no significant differences in slope apparent in the g, i or z-bands.

Fig. 3.3 shows the velocity dispersion-magnitude relation predicted by GALFORM. The one-dimensional velocity dispersion is calculated using $\sigma_{1D} = (1.1/\sqrt{3})V_{c,\text{bulge}}$, where $V_{c,\text{bulge}}$ is the effective circular velocity of the bulge, which is assumed to be isotropic (Frenk et al., 1988; Cole et al., 1994). The factor of 1.1 is an empirical correction which Cole et al. employed to map data for elliptical galaxies onto the Tully-Fisher relation for spiral galaxies. Fig. 3.3 shows that retaining this factor gives good agreement with the observed FJ relation. We find no change in the predictions if we consider, instead, the effective circular velocity of the disc and bulge combined. This is a consequence of our selection which ensures that the model galaxies we consider are bulge dominated. We find reasonably good agreement between model predictions and the FJ relation observed for the Ber05 sample, albeit with a shallower slope, $L_r \propto \sigma^{3.2\pm0.1}$ (note, we plot σ on the y-axis). Whilst the slope of the predicted FJ relation is formally at odds with that measured by Ber05, it is clear from Fig. 3.3 that the velocity dispersion of the model galaxies would change by relatively little even in the case of perfect agreement between the observed and predicted slopes, given the limited range of magnitudes plotted. We find little dependence of the slope of the FJ relation on passband, in agreement with observations.

The evolution of the FJ relation with redshift is plotted in Fig. 3.4. Here, we have chosen output redshifts in GALFORM to match the median redshifts of the Ber05 redshift subsamples: for SDSS galaxies with $z < 0.08$ we use $z_{\text{med}} = 0.06$ for the model galaxies, and for SDSS galaxies with $z > 0.18$ we set $z_{\text{med}} = 0.20$. Little evolution is observed in the zero-point of the FJ relation with redshift, a trend which is reproduced by the model predictions. The shift to brighter magnitudes with increasing redshift is simply a reflection of the fixed apparent magnitude limit of the
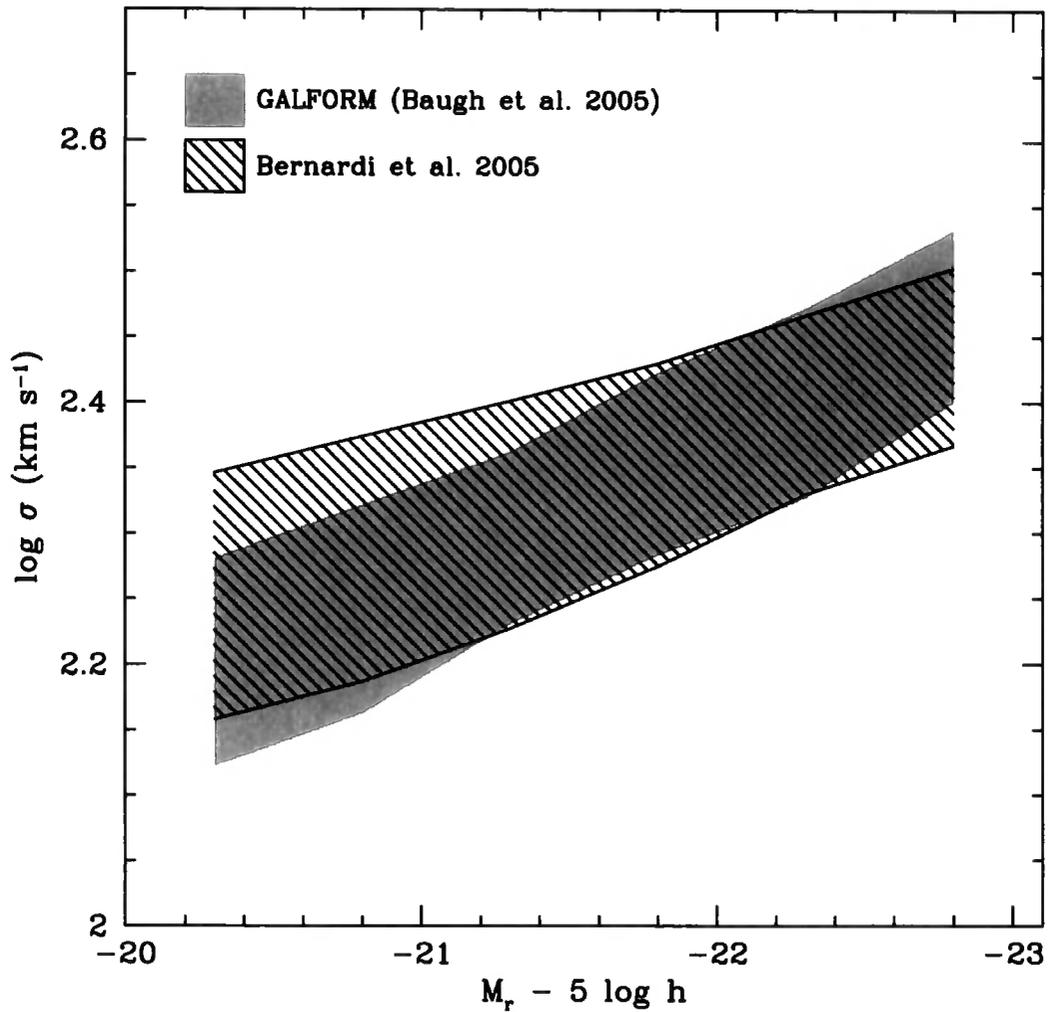
Figure 3.3: The Faber-Jackson relation between luminosity and velocity dispersion. The GALFORM prediction is shown in grey and the hatched shaded region shows the relation for the Ber05 sample. The shaded and hatched regions connect the 10 and 90 percentiles. The one-dimensional velocity dispersion was calculated using $\sigma_{1D} = (1.1/\sqrt{3})V_{c,\text{bulge}}$, as explained in the text.
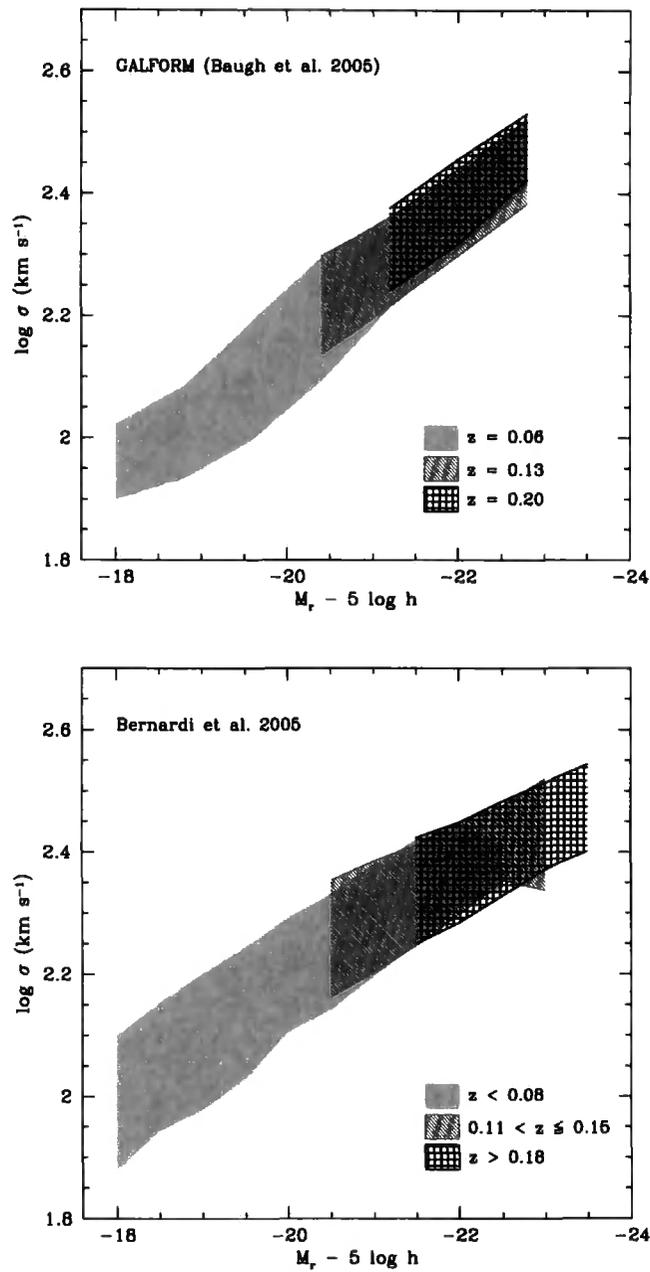
Figure 3.4: The evolution of the Faber-Jackson relation for early-type galaxies. The upper panel shows the model predictions at redshifts $z = 0.06$, $z = 0.13$ and $z = 0.20$, which are the median redshifts of the observational subsamples shown in the lower panel. Shaded regions connect the 10 and 90 percentiles of the distributions.

SDSS. As we shall see in Section 3.5, this absence of evolution is actually expected at these redshifts due to the cancellation of two different evolutionary effects.

Finally, in Fig. 3.5, we plot the luminosity-weighted age of the stellar population, computed in the r-band, as a function of the velocity dispersion. Some authors have argued that a correlation exists between these quantities, which has implications for the scatter in the FJ relation (Forbes & Ponman, 1999; Nelan et al., 2005). Fig. 3.5 reveals that velocity dispersion and luminosity-weighted age are indeed correlated in the model, with galaxies which have larger velocity dispersions also displaying older stellar populations. A linear fit to the model predictions shows that Age $\propto \sigma^{0.58\pm0.02}$, which is in excellent agreement with recent determination by Nelan et al. (2005), who found Age $\propto \sigma^{0.59\pm0.13}$. Furthermore, we verify that the inclusion of AGN feedback, as implemented by Bower et al. (2006), does not change this relation substantially, giving Age $\propto \sigma^{0.51\pm0.03}$. At first sight, these predictions seem to contradict those presented, for a different observational selection, by Nagashima et al. (2005b). However, it is important to note that, at least in the case of the model galaxies, the slope and scatter of the Age $- \sigma$ relation are very sensitive to the selection criteria applied.

3.3.3 Radius–Luminosity Relation

Another component of the fundamental plane of early-type galaxies is the relation between radius and luminosity; galaxies with larger radii are more luminous. This was originally of interest for use in distance scale measurements and cosmological tests (Sandage & Perelmuter, 1990). Different studies indicate that this relation varies slightly with wavelength. For example, Schade et al. (1997) determined $L_B \propto r_e^{1.33}$ in the B-band and Pahre et al. (1998) found $L_K \propto r_e^{1.76}$ in the K-band. For SDSS early-type galaxies, Bernardi et al. (2003b) reported $L_g \propto r_e^{1.50}$ in the g-band and $L \propto r_e^{1.58}$ in the r, i and z-bands, which is consistent with the variation of the slope of this relation with wavelength suggested by the earlier determinations.

We compare the predicted radius–luminosity relation with the Ber05 data in Fig. 3.6. The effective radius plotted here, r_e, is the projected bulge half-light radius of the de Vaucouleurs law, which is related to the half-mass radius in 3D, r_b,
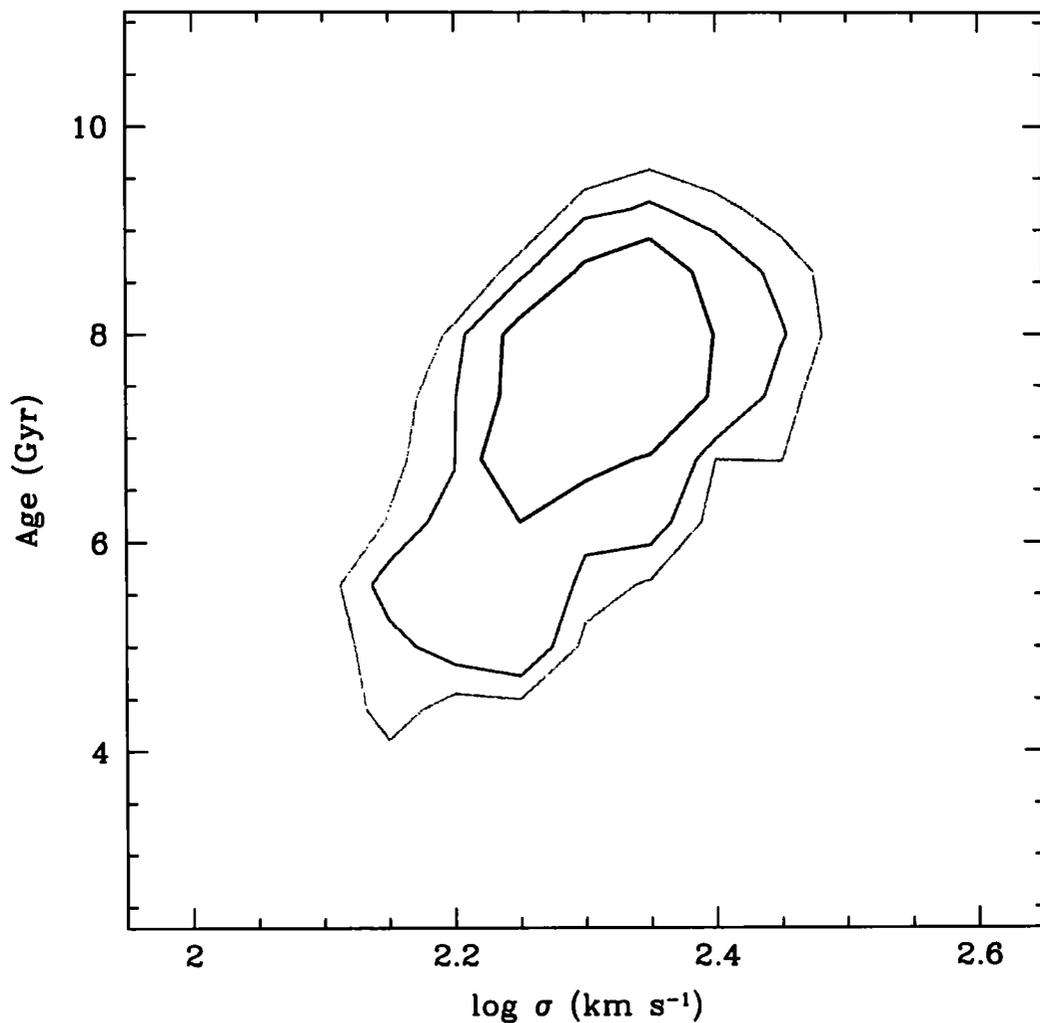
Figure 3.5: The r-band luminosity-weighted age of the stellar population as a function of the velocity dispersion predicted by GALFORM. The model galaxies are at redshift $z = 0.13$ and are selected in a similar way to the observational sample of Ber05. The contours enclose 35%, 72%, 89% and 99% of the number density respectively, moving from black to grey.
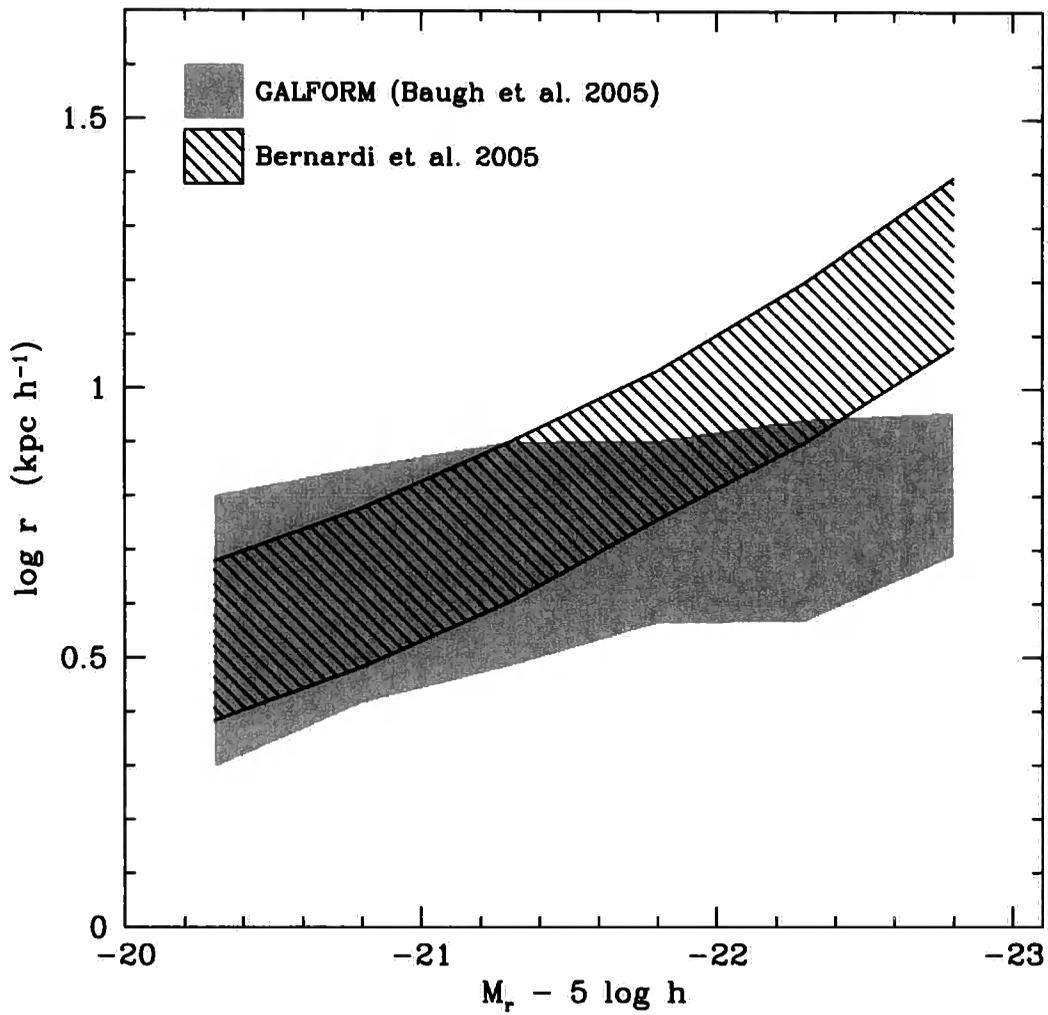
Figure 3.6: The relation between half-light radius and r-band magnitude for early-type galaxies. Again, the predictions for GALFORM galaxies are shown in gray and the hatched shaded distribution represents the SDSS sample; in both cases the shading shows the 10 to 90 percentile range.
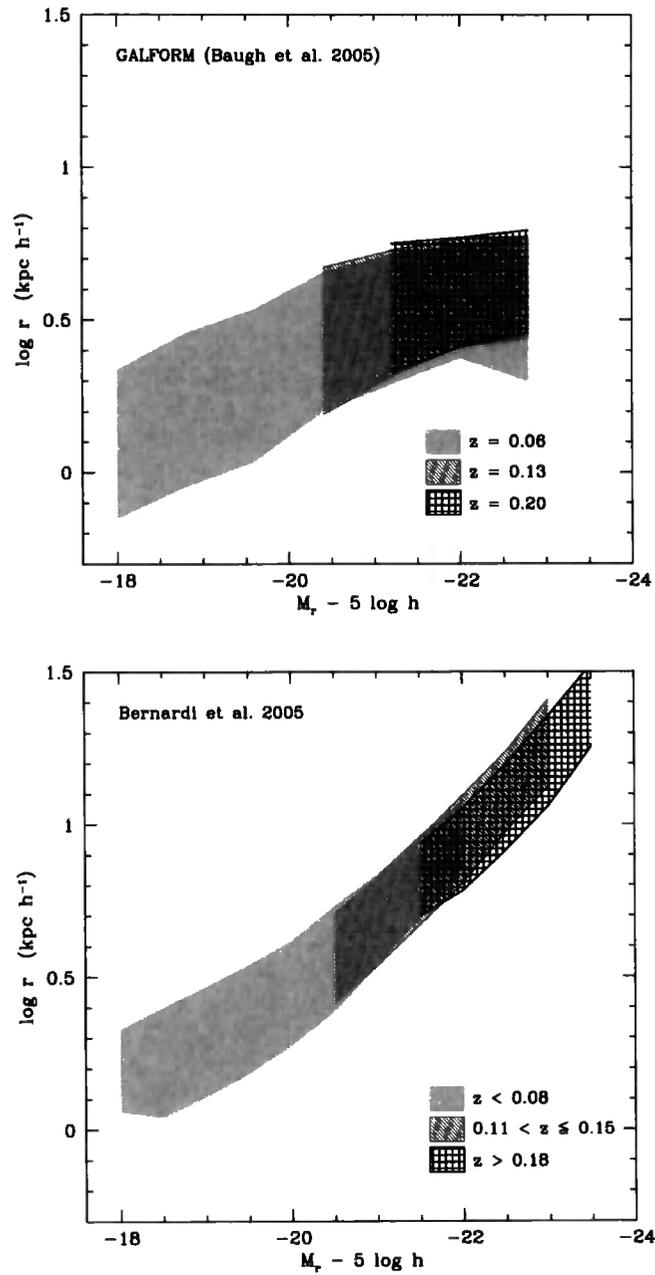
Figure 3.7: The evolution with redshift of the radius – r-band magnitude relation. The upper panel shows GALFORM galaxies at the median redshifts of the observational subsamples: $z = 0.06$, $z = 0.13$ and $z = 0.20$. The lower panel shows the results for the Bernardi et al. (2005) sample divided into volume-limited bins, as indicated by the legend.

by: $r_e = r_b/1.35$. The model predictions do not change if we compute a composite half-mass (half-light) radius from the disc and bulge components of the galaxy, as the model galaxies we consider are bulge dominated. The slope of the predicted radius–luminosity relation is flatter than is observed. The agreement between the model predictions and the observations is best at fainter magnitudes; the brightest early-type galaxies are predicted to be around a factor of three smaller in effective radius than is observed. Fig. 3.6 and Fig. 3.3 suggest that in the model, the high luminosity early-type galaxies have a pseudo-angular momentum which is lower than would be inferred from the data (see Chapter 2 for the definition of the pseudo-angular momentum of a spheroid).

The evolution of the radius–luminosity relation with redshift is plotted in Fig. 3.7. We find no clear change in the slope of the radius–luminosity relation over this redshift interval, in agreement with the results of Bernardi et al. (2003b). We shall return to this point in section 3.5.

3.3.4 Effective Mass

We can define an effective dynamical mass, M_{dyn}, by setting $M_{dyn} \equiv r_e \sigma^2/G$. This quantity differs from the true mass, M, because the definition of M_{dyn} ignores any rotational support and the flattening of galaxies. The difference between the two masses can be quantified by a correction term, ξ: $M = \xi M_{dyn}$. For a galaxy with T-type E0, there is no flattening or rotational support and so $\xi = 1$. In contrast, for the case of an E6 galaxy, the true mass is almost twice as large as the dynamical mass, with $\xi \sim 1.9$ (see Bender et al., 1992, for details).

In Fig. 3.8 we compare our prediction for the relation between dynamical mass and luminosity with the observed result for the Bernardi et al. (2005) sample. The figure reveals reasonable agreement between the model and the observations for fainter galaxies. Brighter galaxies, in the model, have a somewhat lower dynamical mass than observed. As we noted when discussing the radius–luminosity relation, high luminosity galaxies in the model display a lower specific pseudo-angular momentum than is observed, which translates into a smaller dynamical mass.
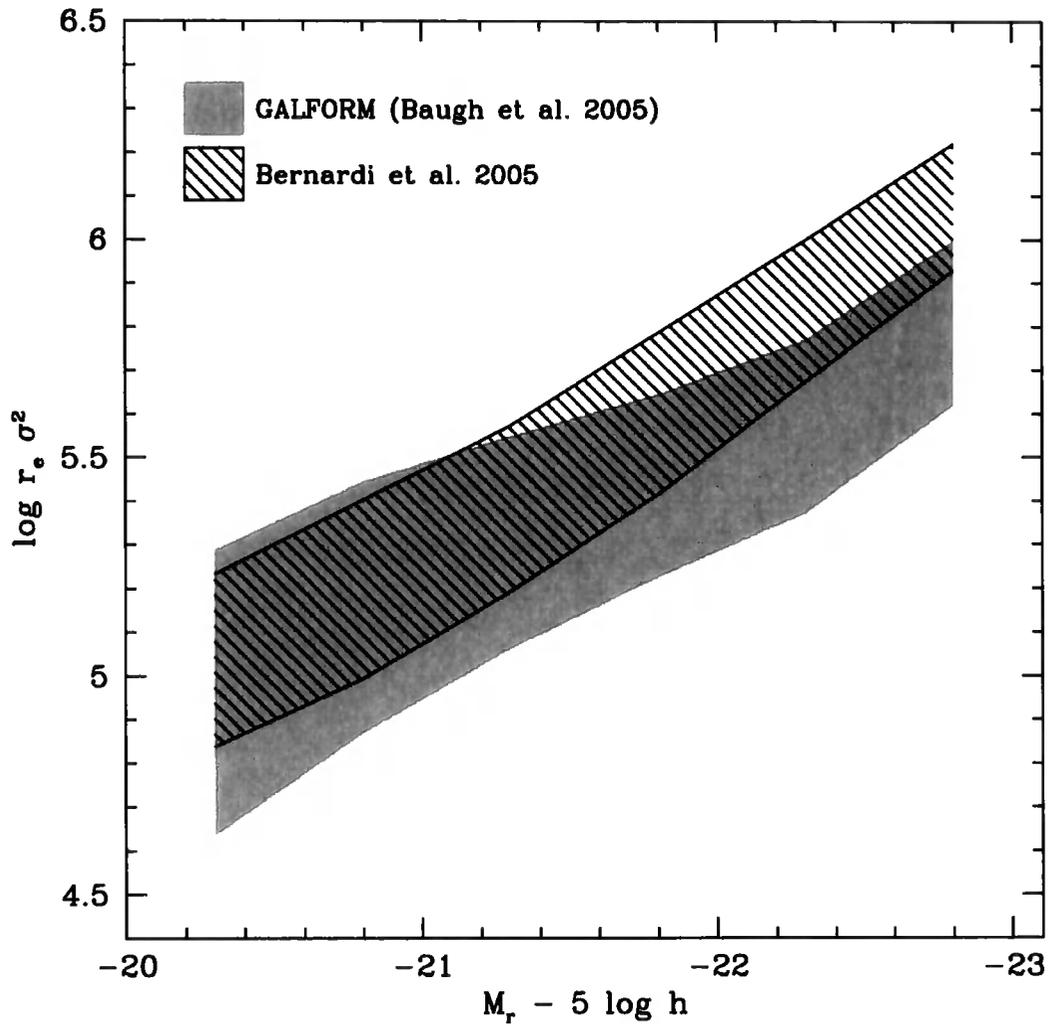
Figure 3.8: The relation between dynamical mass and luminosity. The GAL-
FORM data is represented in gray and the dark hatched shaded region represents
the Bernardi et al. (2005) sample. The shading connects the 10 and 90 percentile
values.

3.3.5 Fundamental Plane

Observational studies indicate that early-type galaxies show tight correlations between their kinematic and photometric properties (Djorgovski & Davis, 1987; Dressler et al., 1987). The remarkably small scatter about the so-called fundamental plane, connecting the effective radius, velocity dispersion and surface brightness of early types, encodes information about the formation and evolution of these galaxies.

The existence of a fundamental plane is expected if a stellar system obeys the virial theorem, which connects the kinetic and potential energies. The assumption of virial equilibrium gives a relation between the three-dimensional velocity dispersion, σ_{3D} and the "gravitational" radius, r_g, assuming that the system is gravitationally bound:

$$\sigma_{3D}^2 = \frac{GM}{r_g}. \tag{3.1}$$

This equation can be rewritten in terms of the central one dimensional velocity dispersion, σ_{1D}, and the effective radius of the galaxy, r_e,

$$\sigma_{1D}^2 = \frac{GM}{\psi_r \psi_v r_e}, \tag{3.2}$$

where we have defined structural constants such that

$$\psi_v \equiv \frac{\sigma_{3D}^2}{\sigma_{1D}^2}, \ \psi_r \equiv \frac{r_g}{r_e}, \tag{3.3}$$

based on the assumption that the population is homologous.

The mean surface brightness within half-mass radius of a galaxy is $I_e \equiv L/2\pi r_e^2$, where L is the total luminosity of the galaxy and the mean surface density is given by $\Sigma_e \equiv M/2\pi r_e^2$. The ratio of the surface brightness to the surface density is equal to the mean mass-to-light ratio of the galaxy, within r_e: $\Sigma_e/I_e = M/L$. Using these definitions, Eq. 3.2 can be rearranged to give an expression for the fundamental plane,

$$r_e = \frac{\psi_r \psi_v}{2\pi G} \frac{\sigma_{1D}^2}{I_e(M/L)}, \tag{3.4}$$

$$\log r_e = 2\log\sigma + 0.4\mu_e + \log(\psi_r\psi_v) - \log(M/L) + \gamma, \tag{3.5}$$

where γ is a constant whose value depends upon G and the choice of units.

The observed plane is slightly different from the form predicted in Eq. 3.5, which follows by applying the virial theorem to a purely stellar galaxy without any dark matter and assuming an homologous population. For example, Jørgensen et al. (1996) found $\log r_e = 1.24 \log \sigma + 0.328 \mu_e + \gamma'$, while Bernardi et al. (2003c) obtained $\log r_e = (1.49 \pm 0.05) \log \sigma + (0.30 \pm 0.01) \mu_e - (8.78 \pm 0.02)$.

The discrepancy between the theoretical prediction outlined above and the observational results is known as the tilt of the fundamental plane. Trujillo et al. (2004a) argued that this tilt is due to a combination of effects: structural nonhomology, which means a change in the surface brightness profile of the early-types with luminosity, and a variation in the mass-to-light ratio of the stellar populations with galaxy luminosity.

The intrinsic thickness or scatter in the fundamental plane poses another challenge, and its interpretation is far from clear. Forbes et al. (1998) showed that the scatter was mainly due to the age of the stellar population. However, Pahre et al. (1999) demonstrated that the position of the galaxy relative to the FP could not be entirely due to age or metallicity effects.

To determine the location of the fundamental plane, we consider an orthogonal fit to the plane given by:

$$\log r_{\rm e} = a \, \log \, \sigma + b \, \mu_{\rm e} + c, \tag{3.6}$$

and determine the values of the coefficients a, b and c by minimizing the quantity

$$\delta = \sum_{i=1}^{N} \frac{(\log r_{{\rm e},i} - a \, \log \, \sigma_{,i} - b \, \mu_{{\rm e},i} - c)^2}{1 + a^2 + b^2}. \tag{3.7}$$

Following this procedure, we obtain a fundamental plane for GALFORM given by $\log r_e = (1.94 \pm 0.01) \log \sigma + (0.19 \pm 0.01) \mu_e - (7.54 \pm 0.03)$ in the r-band.

In Fig. 3.9, we plot the fundamental plane derived from model galaxies, along with the data from Bernardi et al. (2005) in the same projection of the plane. Fig. 3.9 reveals reasonable agreement between the fundamental plane predicted by GALFORM and the observational data: we can reproduce not only the tilt, but also the scatter associated with the plane. In the g-band we calculate: $\log r_e = (2.12 \pm 0.02) \log \sigma + (0.19 \pm 0.01) \mu_e - (7.92 \pm 0.07)$; and similar results in the i
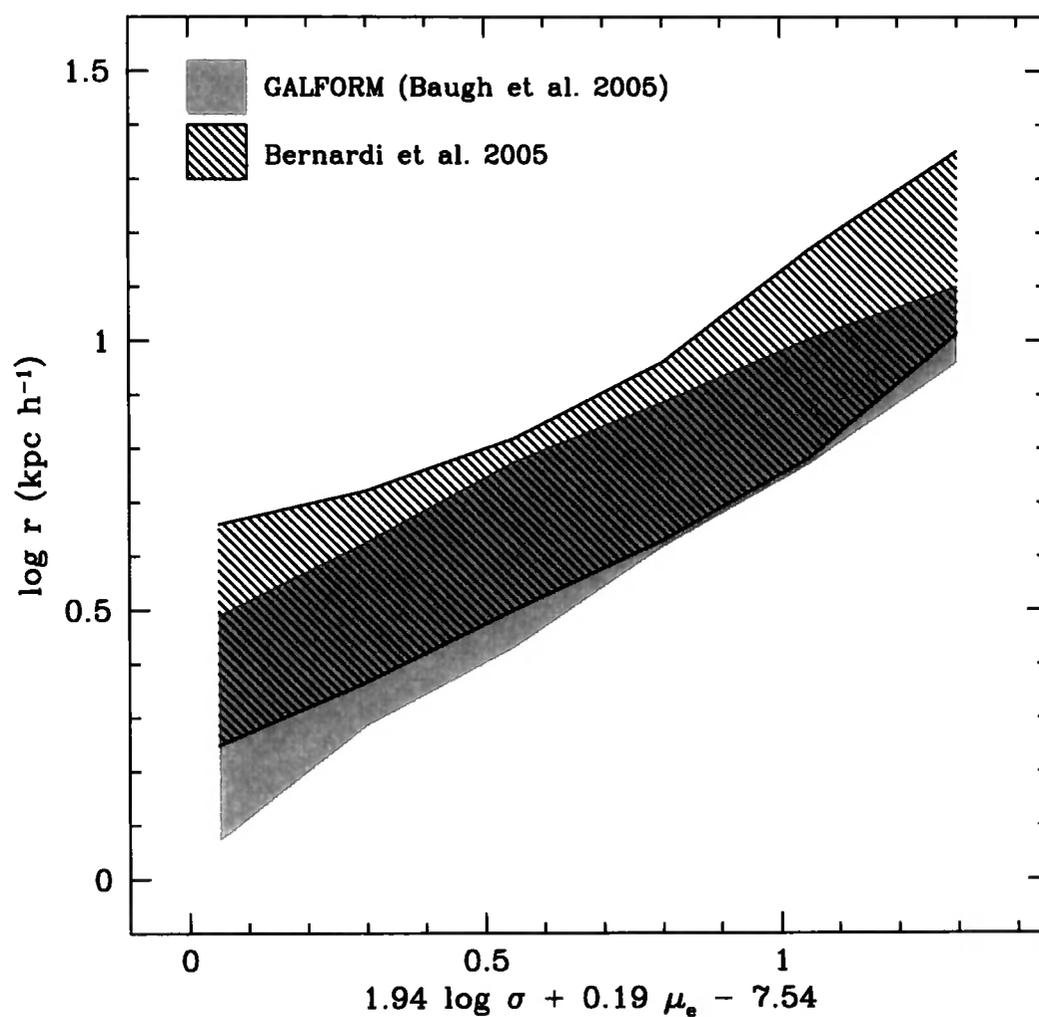
Figure 3.9: The fundamental plane for GALFORM early-type galaxies (gray shading) compared with the observational data from Ber05 (hatched shading). The shading denotes the 10 to 90 percentile interval.

and z-bands. This reveals that the slope of the FP is independent of wavelength, analogous to the results found by Bernardi et al. (2003c).

We plot the fundamental plane at different redshifts in Fig. 3.10. The radius, velocity dispersion, surface brightness and mass-to-light ratios of the model galaxies all evolve with time, so one might expect to see some evolution in the fundamental plane itself, unless the changes in these quantities occur in such a way as to cancel out any evolution in the locus of the plane. Ber05 report evolution in the position of the fundamental plane which corresponds to a change in the mean galaxy surface brightness of $\Delta\mu_e \approx -2z$. We find no clear evidence for evolution in the model predictions over the same redshift interval. In Section 3.5 we show the predictions for the scaling relations over a wider baseline in redshift.

3.4 The dependence of the structural properties of ellipticals on the physical ingredients of the model

Our calculation of the sizes of galactic spheroids contains several steps and is sensitive to some of the physical ingredients of the galaxy formation model more than others. The beauty of semi-analytical modelling is that we can switch off or vary particular assumptions or processes to isolate their impact on the model predictions. Such a study is only possible to a very limited extent in fully numerical simulations of galaxy formation. Moreover, the high speed of the semi-analytical calculations compared with a fully numerical simulation allows us to examine many different variants in a short time. In this section, we seek to establish the sensitivity of our model predictions for the structural and photometric properties of spheroids to the composition of the model. For this purpose, we study the model predictions at $z = 0$ and consider bulge dominated galaxies, i.e. those with a bulge-to-total luminosity ratio in the r-band of $B/T > 0.8$. The results of this section are presented in Figs. 3.11, 3.12, 3.13 and 3.14, which look, respectively, at how deviations from the fundamental plane correlate with various galaxy properties, the Faber-Jackson
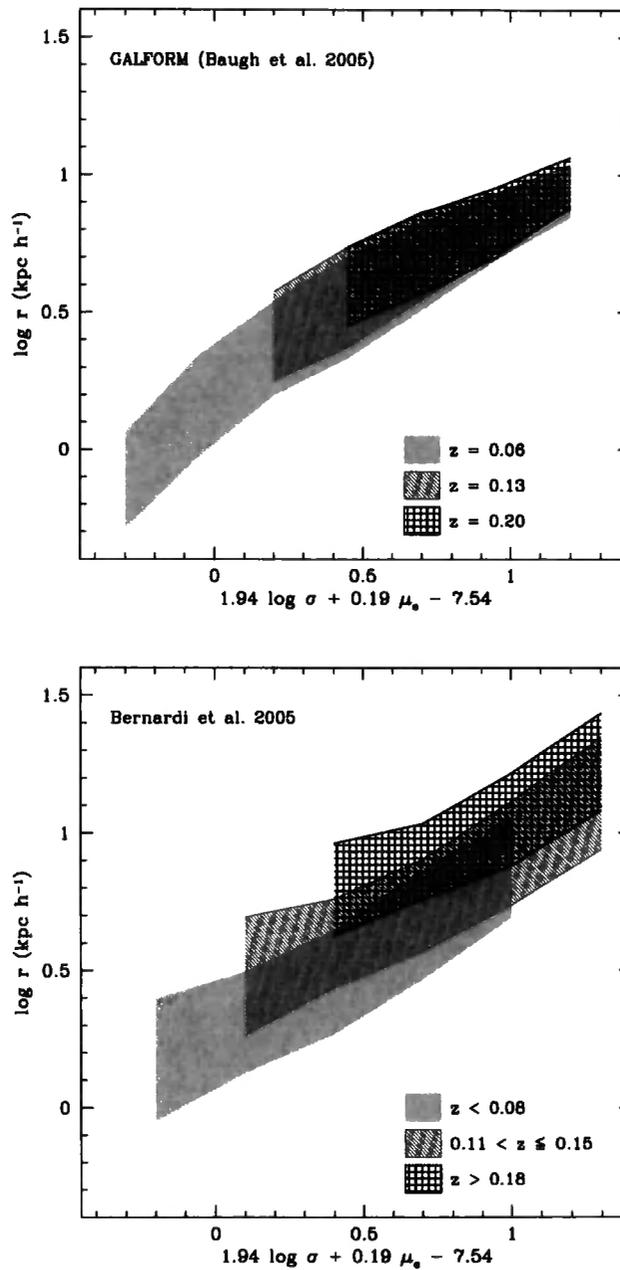
Figure 3.10: The evolution of the fundamental plane with redshift. The upper panel shows model predictions for redshifts $z = 0.06$, $z = 0.13$ and $z = 0.20$; these are the median redshifts of the observational samples plotted in the lower panel. Again, the shaded regions show the 10 to 90 percentile ranges of the distributions.

relation between velocity dispersion and luminosity, the radius–luminosity relation and the fundamental plane.

3.4.1   The deviation from the fundamental plane

As we noted in the previous section, there is some controversy in the literature regarding the source of the dispersion around the fundamental plane, with some authors arguing that the scatter could be due to a number of causes, such as variations in the formation times of galaxies, metallicity trends in stellar populations, or differences in the dark matter content of galaxies.

Fig. 3.11 shows how the deviation from the fundamental plane correlates with various galaxy properties. Here, the quantity Δ represents the offset from the $z = 0$ fundamental plane predicted in the r-band, after applying the selection criteria to match the Ber05 SDSS sample: $\Delta = \log r_b - (1.94 \log \sigma + 0.19 \mu_e - 7.54)$. Fig. 3.11(a) shows that the deviation is correlated with r-band absolute magnitude for galaxies fainter than L_*. This result shows that magnitude-limiting a sample might bias the determination of the fundamental plane. Surprisingly, luminous early-type galaxies ($M_r - 5 \log h < -20$) exhibit no correlation with deviation from the fundamental plane, which is in agreement with the results of Bernardi et al. (2003b). For the remaining panels in Fig. 3.11, we only select galaxies with $M_r - 5 \log h < -19.5$ (i.e. brighter than one magnitude faintwards of M_*), in order to make our results comparable to observations. Fig. 3.11(b) reveals a strong anticorrelation between the deviation and the r-band luminosity-weighted age of the galaxy, in the sense that galaxies which lie above the fundamental plane are younger. A linear fit to the distribution reveals $\Delta = -(0.11 \pm 0.03) \, \text{Age} + (0.64 \pm 0.09)$. This strong correlation indicates that the age of the stellar population plays an important role in determining the position of the galaxy in the fundamental plane space (see Forbes et al., 1998; Pahre et al., 1999). We find that the fundamental plane offset is also anticorrelated with g-r colour (Fig. 3.11(c)). As noted by Bernardi et al. (2003c), this is due to the correlation between colour and velocity dispersion (see also Bernardi et al. 2005). Interestingly, we see in Fig. 3.11(d) that the total stellar mass is anticorrelated with the deviation from the fundamental plane, for galaxies brighter than $M_r - 5 \log h <$
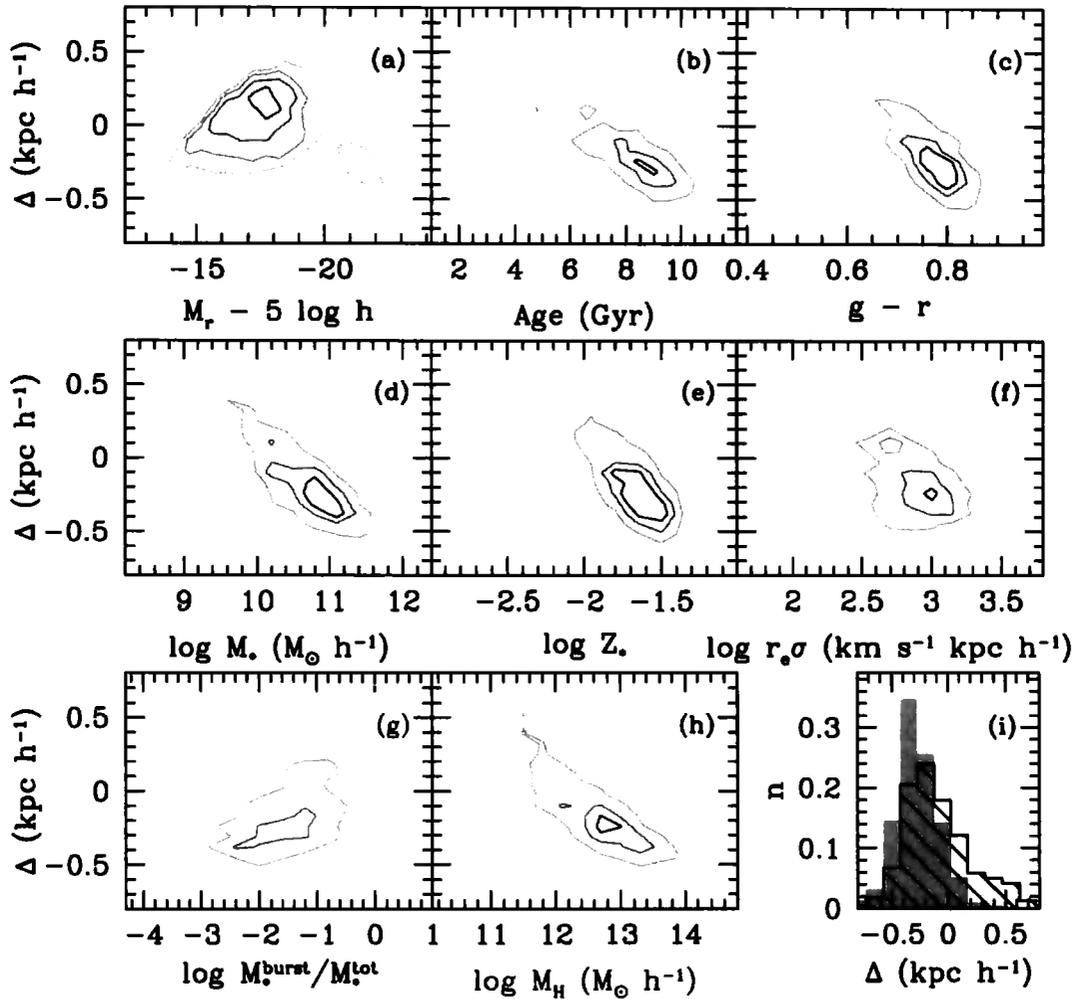
Figure 3.11: The dependence of the deviation (defined as $\Delta = \log r_e - (1.94 \log \sigma + 0.19 \mu_e - 7.54)$) from the fundamental plane in the model on various galaxy properties: (a) the r-band absolute magnitude; then, for $M_r - 5 \log h < -19.5$ galaxies (b) the r-band luminosity-weighted age, (c) the $g - r$ colour, (d) the total stellar mass, (e) the stellar metallicity, (f) the pseudo-specific bulge angular momentum, (g) the ratio between the mass of the stars formed in the last burst, M_*^{burst}, and the total stellar mass, M_*^{tot}, at the present day, (h) the halo mass; and (i) whether or not the galaxy is a central galaxy or a satellite. The contours are indicative of the density of model galaxies. In panel (i), the hatched histogram represents the deviation from the fundamental plane for the central galaxies, and the grey histogram shows the distribution for the satellite galaxies.

-19.5. When all the early-type galaxies are included, the distribution reveals a different picture: similar to the trend seen in panel (a), we find that faint galaxies show a stronger deviation from the fundamental plane. Fig. 3.11(e) shows that the absolute metallicity of the stellar population is anticorrelated with a deviation from the fundamental plane: metal-rich galaxies are to be found predominately below the mean fundamental plane relation. The relation between the pseudo-specific angular momentum, $r_e\sigma$, of the bulge and the FP offset is plotted in Fig. 3.11(f), revealing a weak anticorrelation between these quantities. There is little correlation between the deviation from the fundamental plane and the fraction of the total stellar mass formed in the last burst of star formation triggered by a galaxy merger (Fig. 3.11(g)), which shows that the presence of gas in galaxy mergers does not change significantly the fundamental plane relation. This seems to contradict the recent results of Robertson et al. (2006). However, if we also consider faint galaxies, which tend to have mergers containing a larger fraction of gas, then we find an anticorrelation between $M_*^{\mathrm{burst}}/M_*^{\mathrm{tot}}$ and the deviation from the FP, along the lines of that seen by Robertson et al. The relation between the deviation from the FP and halo mass is shown in Fig. 3.11(h). We find an anticorrelation between these two quantities, such that galaxies which lie in more massive haloes are found below the main FP relation, i.e. cluster galaxies should lie below the mean fundamental plane. In Fig. 3.11(i), we show that the distribution of the FP offset for central galaxies resembles that predicted for satellite early-types.

3.4.2   The physics of the model and the scaling relations

In this section, we examine how the predictions of the Baugh et al. model change if one ingredient at a time is varied. These variant models are not necessarily acceptable galaxy formation models, because they may not give as good a match to the local data used to calibrate the model parameters as was the case for the fiducial Baugh et al. model. We also show the predictions of the Bower et al. model, as a further example of a variant model. In this case, many ingredients have been changed from the ones used in the Baugh et al. model, as explained at length in Chapter 2.
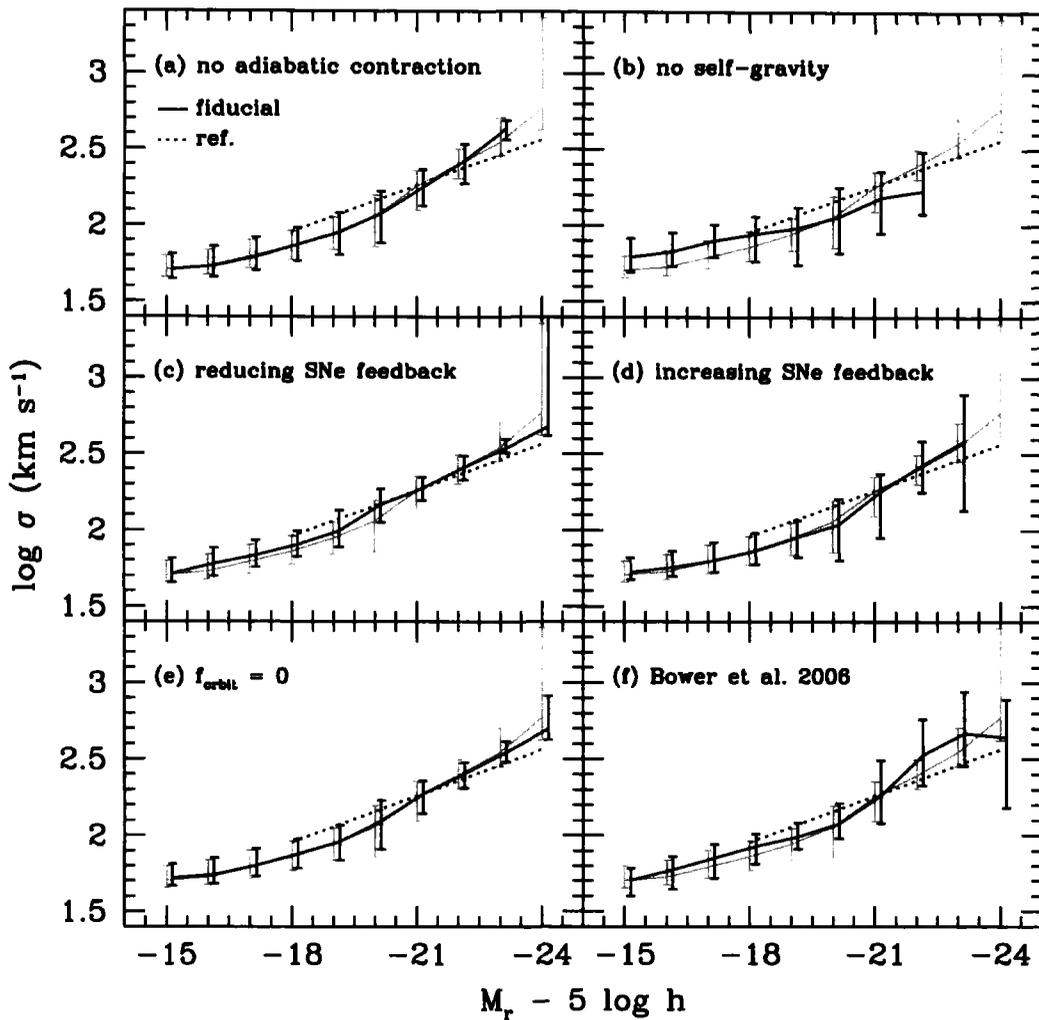
Figure 3.12: The sensitivity of the model predictions for the Faber-Jackson relation to: (a) switching off adiabatic contraction, (b) omitting the self-gravity of the baryons, (c) reducing the strength of supernova feedback, (d) increasing the strength of SNe feedback, (e) the omission of the orbital energy from the calculation of the size of the merger remnant and (f) using the Bower et al. (2006) model with AGN feedback. In each panel, the grey line shows the median prediction from the reference model (Baugh et al., 2005), at $z = 0$. The black solid line shows the median for the variant model. The errorbars indicate the 10 to 90 percentile of the predictions. The dotted line in each panel shows the observed relation for SDSS early-type galaxies for reference (Bernardi et al., 2003b).

The first ingredient we test is the adiabatic contraction model used to take into account the gravitational pull of the baryons on the dark matter. To recap, the condensation of baryons at the centre of the dark matter halo provides an additional gravitational force on the dark matter which causes it to move inwards, thereby increasing the density of dark matter in the central part of the halo. This in turn alters the gravitational force on the baryons due to the dark matter. The degree of contraction is computed by exploiting the fact that, in a slowly varying potential, the action integral, $\oint p_i dq_i$, is an adiabatic invariant for each particle of mass i, where p_i is the conjugate momentum of the coordinate q_i (Barnes & White, 1984; Blumenthal et al., 1986; Jesseit et al., 2002). If we assume spherical symmetry and circular orbits, the action integral simplifies to the conservation of angular momentum in spherical shells, $rM(r)$. The adiabatic contraction of the dark matter leads to a more centrally peaked halo density. The main consequence of switching off the adiabatic contraction of the dark matter halo is that the half-mass radius of the spheroid increases (Fig. 3.13(a)). The radii of bright galaxies increase by a larger factor than those of faint galaxies, leading to a steepening of the radius–luminosity relation. The slope of this relation is in much better agreement with the observed slope on omitting adiabatic contraction, although the model galaxies are too large overall (both spheroids and discs).

Next we ignore the self-gravity of the baryons when computing the size and effective rotation speed of the disc and bulge. This also means that there is no adiabatic contraction. The rotation curve of the galaxy in this case is set purely by the dark matter, which is assumed to have an NFW density profile (Navarro, Frenk, & White, 1997). The consequences of this change are a flattening in the velocity dispersion-luminosity relation (Fig. 3.12(b)), with brighter galaxies displaying a lower velocity dispersion, and a uniform increase in the radius of the spheroid (Fig. 3.13(b)). In combination, these changes result in a different projection of the fundamental plane which looks flatter in the projection which best fits the predictions of the fiducial model.

Feedback, the regulation of the star formation rate due to the reheating and ejection of cooled gas following the injection of energy into the interstellar medium
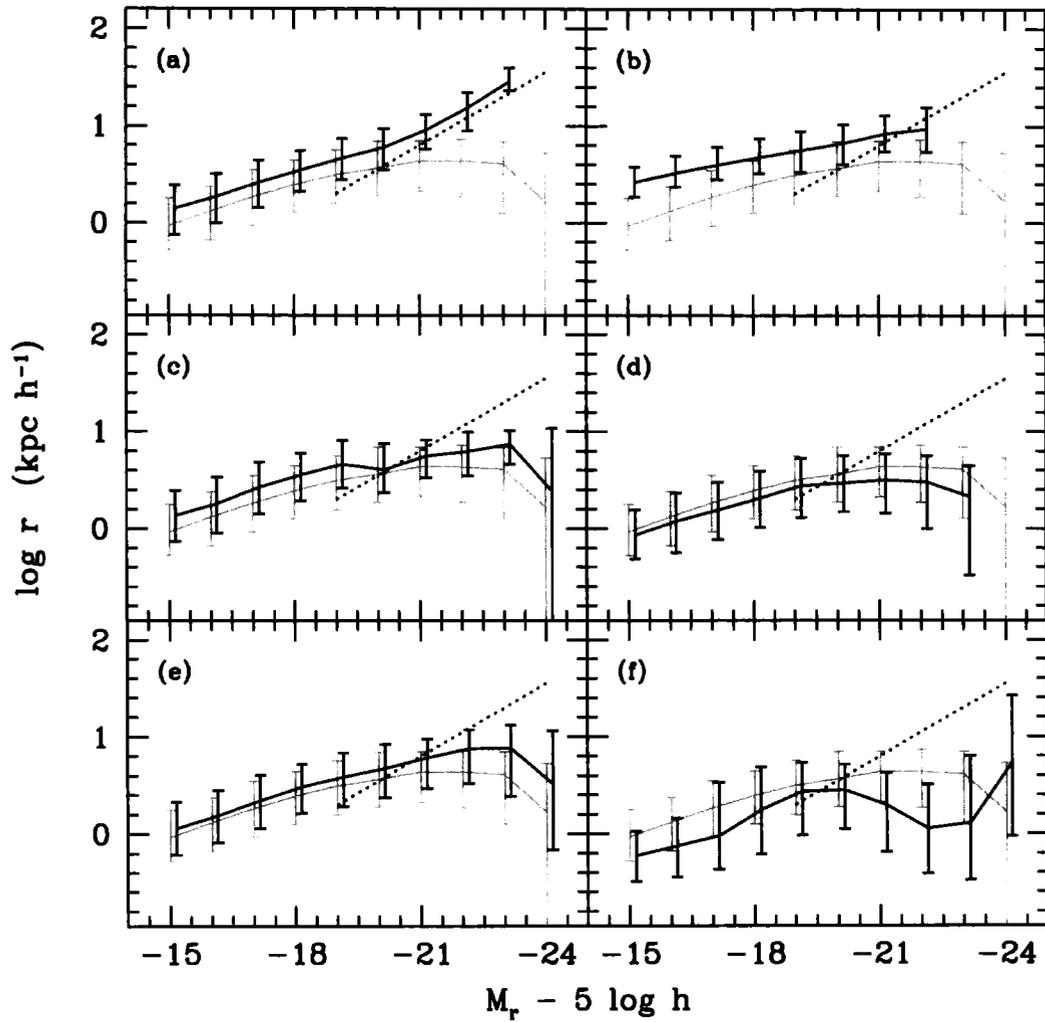
Figure 3.13: The sensitivity of the relation between radius and luminosity to: (a) switching off adiabatic contraction, (b) omitting the self-gravity of the baryons, (c) reducing the strength of supernova feedback, (d) increasing the strength of SNe feedback, (e) the omission of the orbital energy from the calculation of the size of the merger remnant and (f) using the Bower et al. (2006) model with AGN feedback. In each panel, the grey line shows the median prediction from the reference model (Baugh et al., 2005), at $z = 0$. The black solid line shows the median for the variant model. The errorbars indicate the 10 to 90 percentile range of the predictions. The dotted line in each panel shows the observed relation for SDSS early-type galaxies for reference (Bernardi et al., 2003b).

Figure 3.14: The sensitivity of the fundamental plane to: (a) switching off adiabatic contraction, (b) omitting the self-gravity of the baryons, (c) reducing the strength of supernova feedback, (d) increasing the strength of SNe feedback, (e) the omission of the orbital energy from the calculation of the size of the merger remnant and (f) using the Bower et al. (2006) model with AGN feedback. In each panel, the grey line shows the median prediction from the reference model (Baugh et al., 2005), at $z = 0$. The black solid line shows the median for the variant model. The errorbars indicate the 10 to 90 percentile range of the predictions. The dotted line in each panel shows the observed relation for SDSS early-type galaxies for reference (Bernardi et al., 2005).

by supernova explosions, plays an important role in setting the sizes of disc galaxies (Cole et al., 2000). The strength of SNe feedback is quantified in the model by the parameter β (for details see Cole et al., 2000): $\beta = (V_{hot}/V_{disc})^{\alpha_{hot}}$, where V_{hot} and α_{hot} are parameters and V_{disc} is the rotation speed of the disc at the half mass radius. The mass of cold gas which is reheated is given by $\dot{M}_{reheat} = \beta\psi$, where ψ is the star formation rate. In the Baugh et al. model, the values adopted for these parameters are: $\alpha_{hot} = 2$ and $V_{hot} = 300\,\mathrm{km\,s^{-1}}$. We show the impact of reducing (by setting $v_{hot} = 100\,\mathrm{km\,s^{-1}}$) and increasing (by setting $v_{hot} = 600\,\mathrm{km\,s^{-1}}$) the strength of supernova feedback in Figs. 3.12, 3.13 and 3.14-(c)-(d). Cole et al. demonstrated that increasing the strength of supernova feedback results in gas cooling to form stars in larger haloes, which leads to larger discs. Conversely, reducing the feedback allows gas to cool and form stars in smaller haloes resulting in smaller discs. These trends are reproduced in Figs. 3.13(c) and (d). There is little change in velocity dispersion on changing the strength of the supernova feedback. The shift in the zero-point of the radius–luminosity relation produces a change in the location of the fundamental plane (see Fig. 3.14-(c)-(d)).

In the fiducial GALFORM model, spheroids are the end products of galaxy mergers. As we explained in Chapter 2, the radius of the merger remnant is determined by conserving the binding energy of the individual galaxies involved in the merger and their relative orbital energy. The contribution of the orbital energy to the energy budget is parameterized by f_{orbit}: the standard choice is to set $f_{orbit} = 1$ and to include the full orbital energy in the calculation of the remnant size. In Figs. 3.12, 3.13 and 3.14-(e), we show the effect of removing the contribution of the orbital energy from the calculation of the radius of the spheroid produced by mergers, i.e. we set $f_{orbit} = 0$. Perhaps surprisingly, this change results in an imperceptibly small change in the radius of the spheroid, except in the case of the brightest galaxies.

Finally, we consider the model of Bower et al. (2006), who implemented an AGN feedback scheme into GALFORM, in which cooling flows are quenched in massive haloes at low redshift. As a result of this change to the cooling model in GALFORM, Bower et al. (2006) were able to produce improved matches to the local B and K-band luminosity functions, the observed bimodality of colour distribution and

the inferred evolution of the stellar mass function. In Figs. 3.12, 3.13 and 3.14-(f) we plot the scaling relations for Bower et al. (2006) model. Though the model performs quite well in reproducing the local fundamental plane of early-type galaxies and the Faber-Jackson relation, the radius-luminosity relation for bright galaxies is substantially different from both the observations and from the predictions of the Baugh et al. (2005) model: luminous galaxies in the Bower et al. model are up a factor of three smaller in radius than in the Baugh et al. model.

We also considered a variant of the Baugh et al. model in which the Kennicutt IMF was used in starbursts, in place of the top-heavy IMF. This produces scaling relations for early type galaxies which look very similar to those presented for the Bower et al. model in Figs. 3.12–3.14, with the main change being a shift in the predicted radius-luminosity relation. When the IMF is changed, the yield and recycled gas fraction are also changed accordingly, which affect the rate at which gas cools and alter the star formation timescale. This suggests that the primary difference in the predictions of the Baugh et al. and Bower et al. models is due to the choice of the IMF used in starbursts, in spite of the other differences between the models outlined in Chapter 2.

The results in this section suggest that the scaling relations of early-type galaxies are essentially insensitive to variations in some of the model parameters. However, this should not be interpreted as implying that these observations are of limited value in constraining the models. It should be remembered that our starting point is a model of galaxy formation which has already successfully passed a range of comparisons with observed galaxy properties.

3.5 The evolution of scaling relations

We now present the GALFORM predictions for the evolution of the structural and photometric properties of early-type galaxies with redshift. In this section, we consider the evolution over a much wider baseline in redshift than we addressed in the previous section. Furthermore, in order to get a clear picture of the nature of the evolution, we relax some of the selection criteria which we applied to the model

output in previous sections, where the goal was to mimic the Ber05 sample selection as closely as possible. The only selection we apply in this section is that the bulge must account for at least 80% of the total luminosity in the rest-frame B-band.

It is important to be able to disentangle changes in the typical stellar populations of early-type galaxies with redshift from evolution in their structural properties. Hence, we first examine the predicted evolution in the mass-to-light ratio of early-type galaxies in Fig. 3.15. The stellar populations of early-type galaxies at $z = 1$ in GALFORM have lower mass-to-light ratios by a factor of ≈ 3 compared with the early-types at $z = 0$. This result is in agreement with the change in mass-to-light ratio inferred from observations by van de Ven et al. (2003), $d(\log M/L)/dz = -0.47\pm0.11$ (see also van Dokkum et al., 2007; van de Wel et al., 2005; Treu et al., 2005). We find a weak dependence of mass-to-light ratio on total stellar mass. However this is somewhat lower than what is seen observationally (see Wuyts et al., 2004; van de Wel et al., 2005). Note that when we restrict our attention to bright galaxies (i.e. those with $M_B - 5\log h < -19.5$), the prediction for the median mass-to-light ratio steepens considerably, bringing the model predictions into much better agreement with the observational estimates.

The evolution with redshift of the Faber-Jackson relation is shown in Fig. 3.16. We show the correlation between velocity dispersion and the B-band magnitude in the rest-frame for the local universe ($z = 0$), $z = 0.5$ and $z = 1$. Fig. 3.16 shows that the model predicts differential evolution in velocity dispersion with rest-frame luminosity; at brighter luminosities, the velocity dispersion drops by up to a factor of ≈ 3 between $z = 0$ and $z = 1$, whereas for fainter luminosities, the change in velocity dispersion is much more modest. These results are similar to the ones found in observational studies (cf. di Serego Alighieri et al., 2005; Reda et al., 2006). We note that the scatter around the FJ relation seems to increase slightly with redshift. Furthermore, it is clear that the slope of the relation for faint early-type galaxies is shallower than that of the bright-end, which resembles the results for faint galaxies found by some authors (e.g. Matković et al., 2006; Davies et al., 1983).

Fig. 3.17 shows how the relation between radius and luminosity varies with redshift. As we saw in Section 3.3.3, the model predicts that the brightest early-types
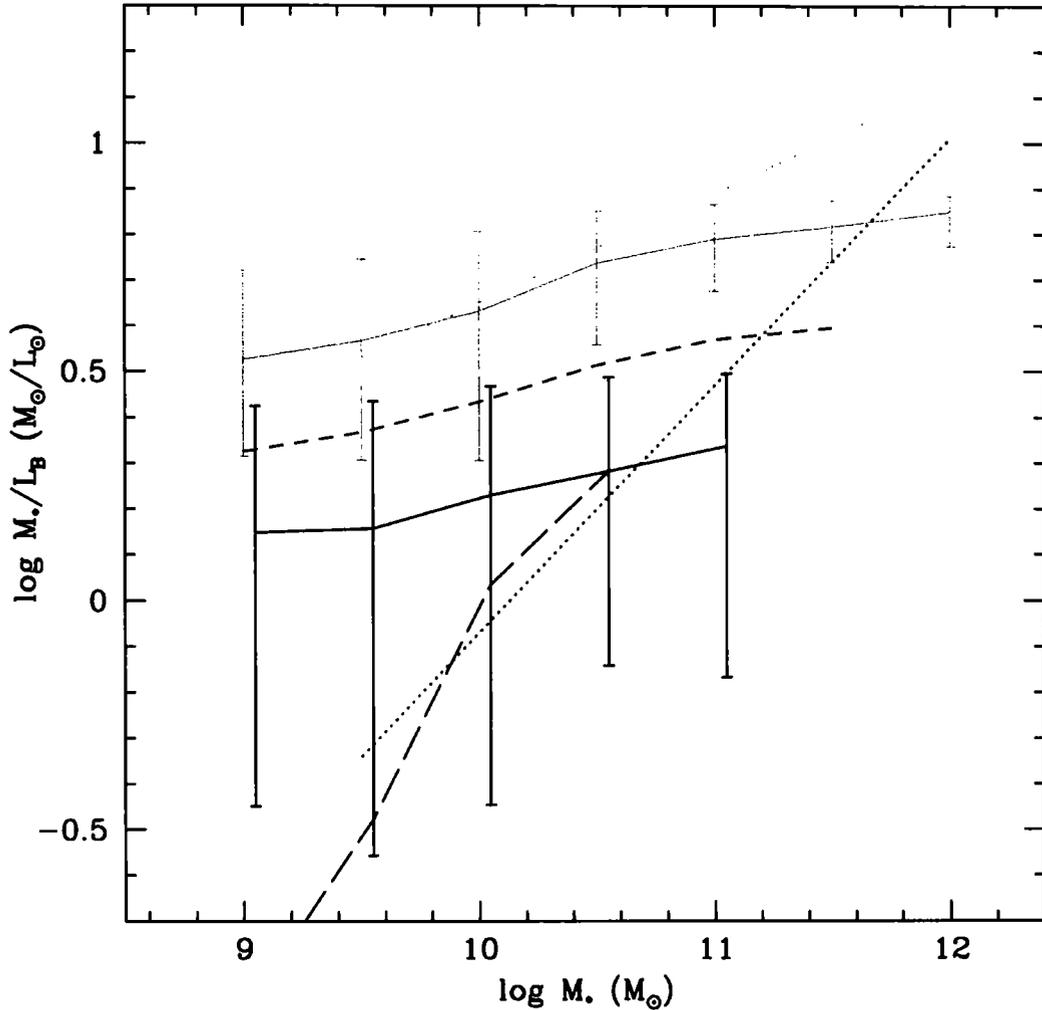
Figure 3.15: The predicted evolution with redshift of the mass-to-light ratio in the rest-frame B-band, plotted against stellar mass. The grey line shows the prediction for $z = 0$, the short dashed line for $z = 0.5$, the solid black line for $z = 1$. The long dashed line shows the prediction for the median mass-to-light ratio at $z = 1$, when only considering galaxies with $M_B - 5 \log h < -19.5$. The dotted lines show the relations found by Jørgensen et al. (2006) for the Coma cluster (grey), $z \approx 0$, and a high redshift sample (black), $z \approx 1$.
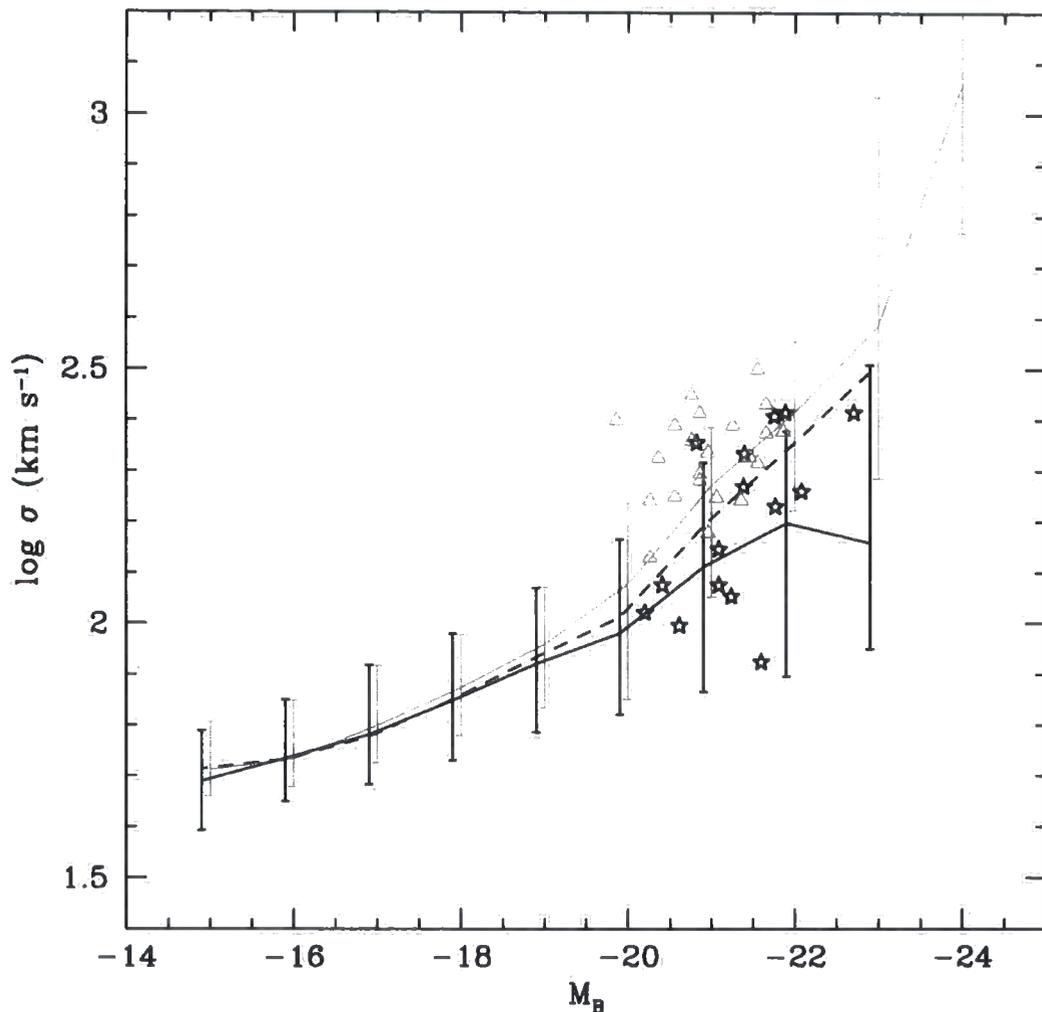
Figure 3.16: The predicted evolution with redshift of the relation between velocity dispersion and rest-frame B-band luminosity. The grey solid line shows the Faber-Jackson relation in the local universe ($z = 0$), while the short-dashed line shows it at $z = 0.5$ and the black solid line at $z = 1$. The errorbars show the 10 to 90 percentile range of the distribution. The gray triangles show local data from Reda et al. (2006) and the stars represent the sample of early-type galaxies in the K20 survey by di Serego Alighieri et al. (2005), with $z \approx 1$.

are too small. Nevertheless, it is interesting to look at the predictions at different redshifts. The primary agent behind the shift in the predictions is the passive evolution of the stellar populations in the elliptical galaxies. Also, at a fixed stellar mass, the galaxy radii decrease with redshift (Coenda et al 2008; in preparation). As redshift increases, we see the bulk of the stars in ellipticals when they were younger and hence brighter. There is no significant trend in the size of the scatter in this relation with redshift.

Trujillo et al. (2004b) estimated the evolution of the radius–luminosity relation of bright galaxies up to $z \sim 3$ (see also Bouwens et al., 2004). These authors found that early-type galaxies, as defined by a high value of Sérsic index, show a size evolution proportional to $(1 + z)^{-1.01 \pm 0.08}$; this is comparable to the amount of evolution we predict in Fig. 3.17 for all ellipticals.

The evolution of the fundamental plane has long been used to study changes in the stellar populations of galaxies (e.g. van Dokkum & Franx, 1996; van Dokkum et al., 2001; Gebhardt et al., 2003; van de Wel et al., 2006). As previously noted, Bernardi et al. (2003c) found evolution in the fundamental plane which is consistent with the passive aging of the stellar population, $\Delta\mu_e \approx -2z$, but without any noticeable difference in the slope. There is a general consensus in the literature regarding the nature of the evolution of the fundamental plane (Gebhardt et al., 2003; Ziegler et al., 2005; Jørgensen et al., 2006). However, as we have shown in the previous section, deviations from the fundamental plane relation are linked to several galaxy properties.

In Fig. 3.18, we plot the model predictions in the Jørgensen et al. (2006) projection of the fundamental plane, $\log r_e = 1.2 \log \sigma + 0.33\mu_e - 9.1$, for galaxies with $M_B - 5 \log h < -19.5$ mag. In this projection, we find no evolution in the slope or offset of the fundamental plane up to $z = 1$, for galaxies with $\log(r_e/h^{-1}\,\mathrm{kpc}) \geq 0.3$, which is contrary to the claims made from observations. As we noted in Fig. 3.11(a), the evolution found in observational studies might be partly due to the correlation between the magnitude and the deviation from fundamental plane, i.e. the act of magnitude-limiting samples might induce the zero-point of the FP to shift. On the other hand, the effective radius of bright early-type galaxies in GALFORM is smaller
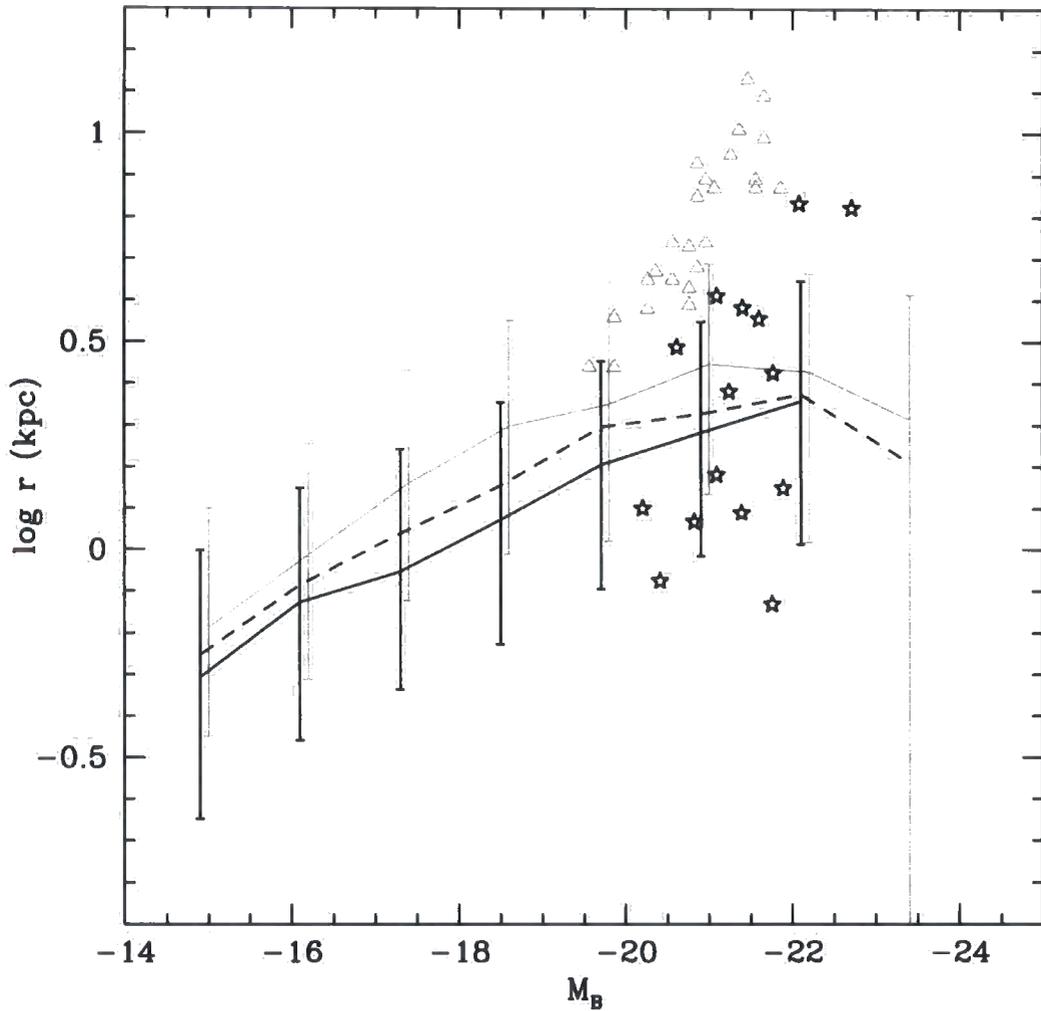
Figure 3.17: The predicted evolution with redshift of the relation between radius and luminosity. The grey solid line shows the relation in the local universe, the black dashed line at $z = 0.5$ and the black solid line at $z = 1$. The errorbars show the 10 to 90 percentile range of the model predictions. The triangles show local data from Reda et al. (2006) and the stars are galaxies in the K20 survey by di Serego Alighieri et al. (2005), with $z \approx 1$.
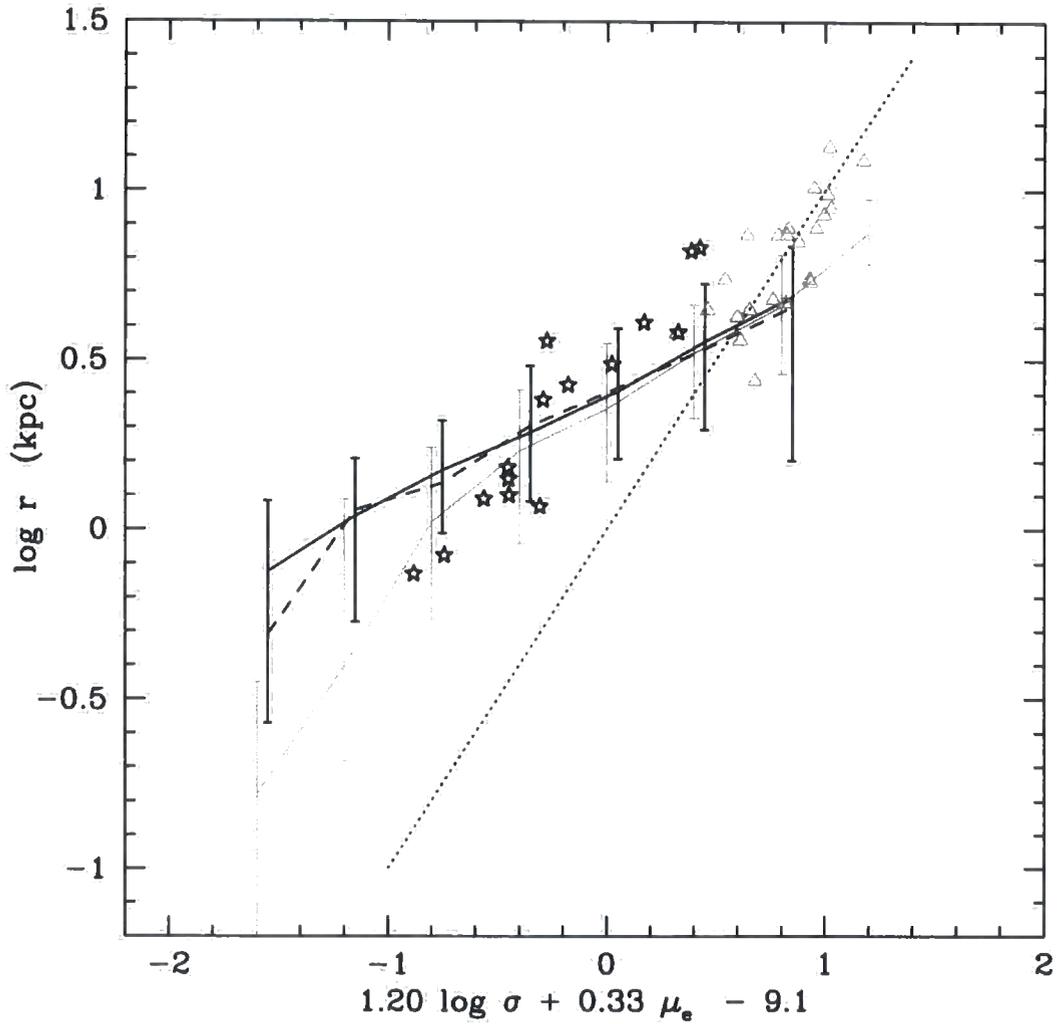
Figure 3.18: The predicted evolution of the fundamental plane with redshift for galaxies with $M_B - 5\log h < -19.5$ mag. The grey solid line shows the relation in the local universe, the dashed black line at $z = 0.5$ and the black solid line at $z = 1$. The errorbars show the 10 to 90 percentile range of the model predictions. The dotted line shows the relation found by Jørgensen et al. (1996). The gray triangles represent $z = 0$ data from Reda et al. (2006) and the stars show data from di Serego Alighieri et al. (2005), at $z \approx 1$.

than observed (see Fig. 3.6), and evolves with redshift, which will complicate any inferences drawn from the evolution of the FP. Interestingly, in the small radius regime, the evolution of our predicted fundamental plane shows an offset similar to that observed.

3.6 Discussion and conclusions

In this chapter, we presented tests of the model proposed by Cole et al. (2000) to calculate the scale sizes of the disc and bulge components of galaxies. This is currently the most sophisticated model in use in semi-analytical codes to compute the radii of galaxies. In brief, the model assumes that galactic discs have an exponential profile and that spheroids follow an $r^{1/4}$ law in projection. The hot gas atmosphere in dark matter haloes is assumed to have the same specific angular momentum as the dark matter. Gas is assumed to retain its angular momentum as it cools to form a galactic disc. The size of a merger remnant is computed by conserving the sum of the binding and orbital energies of the merging galaxies and applying the virial theorem. The self-gravity of the baryons and their impact on the distribution of dark matter in the central parts of the halo are taken into account. Cole et al. demonstrated that this model predicts scale length distributions for galactic discs which are in excellent agreement with observations.

We have carried out the first tests of the model predictions for the structural properties of early-type galaxies and the evolution of these relations with redshift, using the published models of Baugh et al. (2005) and Bower et al. (2006). The Baugh et al. and Bower et al. models differ in a number of ways, as set out in Chapter 2. Two of the main differences are the manner in which the models prevent the overproduction of bright galaxies and in the IMF assumed in starbursts. Bower et al. use AGN heating to switch off the cooling-flow in haloes with a quasistatic hot gas atmosphere, whereas Baugh et al. invoke a superwind which ejects gas that has already cooled. Perhaps controversially, Baugh et al. adopt a flat IMF in starbursts, and a standard solar neighbourhood IMF for quiescent star formation; in Bower et al., a standard IMF is assumed in all modes of star formation. We

emphasize that, for the majority of the results presented, we have not adjusted any of the model parameters in order to improve the predictions for the fundamental plane and its projections. The one exception is where we exploit the modular nature of semi-analytical models to vary or switch off various physical ingredients of the model in order to assess their influence on the model predictions (Section 3.4).

The model enjoys some notable successes. We demonstrated that the model can match the abundance of early-type galaxies in the SDSS sample of Ber05. We also obtain a reasonable match to the Faber-Jackson relation between velocity dispersion and luminosity and its evolution with redshift, albeit with a shallower slope than measured by Ber05. Furthermore, we find a relation between velocity dispersion and age which is in excellent agreement with recent observations. Perhaps most impressively, the fundamental plane predicted by the model is in good agreement with that inferred for SDSS early-types by Bernardi et al. (2003c). The deviation from the FP relation reveals a strong correlation with luminosity, age, colour, stellar mass and metallicity: galaxies that lie above the mean fundamental plane relation are more luminous, younger, bluer, less massive and metal-poor. Furthermore, the feedback processes and clustering, as given by the pseudo-specific angular momentum of the bulge and the halo mass respectively, seem to play a role in defining the plane.

Nevertheless, despite these achievements, there are some model predictions which disagree with the observations. Formally, the slope of the predicted Faber-Jackson relation is at odds with that measured by Ber05, although the overlap between the model galaxies and observations in this projection remains impressive. However, perhaps the most striking discrepancy is the slope of the radius–luminosity relation; the model predicts a significantly flatter radius luminosity relation than is observed. Whereas the model predictions for the effective radii of faint spheroids are in good agreement with the data, the brightest galaxies are up to a factor of three smaller in the model. Our results suggest that, in the model, the brightest spheroids have less specific pseudo-angular momentum (i.e. $j_b = r_b\sigma$, this is a definition of convenience), than is the case for observed galaxies. This could be due to the model underpredicting the galaxy mass for a given luminosity. Somewhat surprisingly, the predicted slope of the radius–luminosity relation is in much better agreement with

the observations if the adiabatic contraction of the halo is switched off (although, in this case, the model galaxies are uniformly too large without adjusting other parameters). The adiabatic contraction of the halo in response to the presence of condensed baryons has been tested against numerical simulations (e.g. Jesseit et al., 2002; Sellwood & McGaugh, 2005; Choi et al., 2006). Our prescription for computing the size of merger remnants could become inaccurate if there is a significant fraction of mass in the form of cold gas.

The other significant discrepancy is the evolution with redshift of the zero-point of the fundamental plane. The model predicts no evolution in the zero-point of the fundamental plane. This is at odds with the evolution inferred observationally, which is consistent with the shift in the mass-to-light ratio expected for a passively evolving stellar population. This discrepancy is intriguing, as the model *does* predict a decline in the mass-to-light ratio of early-types with increasing redshift of the magnitude expected for passive evolution. The lack of evolution in the predicted fundamental plane therefore points to a compensating change in one of the other projections; the effective radii of galaxies also evolve with redshift in the model. This serves as a cautionary note to observational studies which interpret a shift in the fundamental plane in terms of a corresponding change in the typical mass-to-light ratio. The correlation between luminosity and the deviation from the fundamental plane shows that part of the evolution found by observational studies may in fact be due to the construction of magnitude limited samples.

Chapter 4

Luminous Red Galaxies in Hierarchical Cosmologies

4.1 Introduction

Over the past few years, the most rapidly developing aspect of galaxy formation modelling has been the formation of massive galaxies (see Baugh 2006 for a review). As discussed in Chapter 1, on employing the standard White & Frenk (1991) model for the radiative cooling of gas in massive dark matter haloes, hierarchical models have tended to overproduce luminous galaxies. One pragmatic solution to this problem is to simply stop "by hand" the formation of stars from cooling flows in high circular velocity haloes (Kauffmann et al., 1993). A variety of physical mechanisms have been proposed to account for the suppression of the star formation rate in massive haloes, including: (i) the injection of energy into the hot gas halo to reduce its density and hence increase the cooling time (Bower et al., 2001; McCarthy et al., 2007); (ii) the fragmentation of the hot halo in a multi-phase cooling model (Maller & Bullock, 2004); (iii) the complete ejection of gas from the halo in a "super wind" (Benson et al., 2003); (iv) the suppression of the cooling flow due to heating by an AGN (Croton et al., 2006; Bower et al., 2006); (v) thermal conduction of energy within the hot halo (Fabian et al., 2002; Benson et al., 2003). With such a range of possible physical processes to choose from, it is important to develop tests of the models which can distinguish between them. The mechanisms invoked to suppress the formation of bright galaxies could scale in different ways with redshift, leading

to different predictions for the galaxy properties at intermediate and high redshift.

In this chapter we present new tests of the physical processes invoked to suppress the formation of bright galaxies. At the present day, the bright end of the luminosity function is dominated by early-type galaxies with passively evolving stellar populations (e.g. Norberg et al., 2002). Here we focus on a subset of bright galaxies, luminous red galaxies (LRGs) and test the model predictions for the abundance and properties of these red, massive galaxies. LRGs were originally selected from the Sloan Digital Sky Survey (SDSS, York et al., 2000) on the basis of their colours and luminosities (Eisenstein et al., 2001). The red colour selection isolates galaxies with a strong 4000 Å break and a passively evolving stellar population. The galaxies selected tend to be significantly brighter than L_*. The SDSS sample has a median redshift of $z \sim 0.3$. Recently, the construction of LRG samples has been extended to higher redshifts, using SDSS photometry and the 2dF and AAOmega spectrographs ($z \sim 0.5$, Canon et al. 2006; $z \sim 0.7$, Ross et al. 2007b, Wake et al. 2008) (2SLAQ). On matching the colour selection between the SDSS and 2SLAQ surveys, the evolution in the luminosity function of LRGs between $z \sim 0.3$ and $z \sim 0.5$ is consistent with that expected for a passively evolving stellar population (Wake et al., 2006). This has implications for the stellar mass assembly of the LRGs, with the bulk of the stellar mass appearing to have been in place a significant period before the LRGs are observed (Wake et al., 2006; Roseboom et al., 2005; Brown et al., 2007). Due to their strong clustering amplitude and low space density, LRGs are efficient probes of the large-scale structure of the Universe. The clustering of LRGs has been exploited to constrain cosmological parameters (e.g. Eisenstein et al., 2005; Hütsi, 2006; Padmanabhan et al., 2007). The clustering of LRGs on smaller scales has been used to constrain the mass of the dark matter haloes which host these galaxies and to probe their merger history (Zehavi et al., 2005; Masjedi et al., 2006; Ross et al., 2007a).

To date, surprisingly little theoretical work has been carried out to see if LRGs can be accommodated in hierarchical cosmologies, and only very simple models have been used. Granato et al. (2004) considered a model for the formation of spheroids in which the quasar phase of AGN activity suppresses star formation in massive

galaxies, and found reasonable agreement with the observed counts of red galaxies at high redshift. Hopkins et al. (2006) used the observed luminosity function of quasars along with a model for the lifetime of the quasar phase suggested by their numerical simulations to infer the formation history of spheroids, and hence red galaxies. Conroy et al. (2007) use N-body simulations to study the merger histories of the dark matter haloes that they assume host LRGs. Using a simple model to assign galaxies to progenitor haloes, they argue that mergers of LRGs must be very efficient, or that LRGs are tidally disrupted, in order to avoid populating cluster-sized haloes with too many LRGs. Barber et al. (2007) used population synthesis models coupled with assumptions about the star formation histories of LRGs to infer the age and metallicity of their stellar populations.

Here we present the first fully consistent predictions for LRGs from hierarchical galaxy formation models, using two published models, namely Baugh et al. (2005) and Bower et al. (2006). These models, both based on the GALFORM semi-analytical code of Cole et al. (2000), carry out an *ab initio* calculation of the fate of baryons in a cold dark matter universe. The models predict the star formation and merger histories for the whole of the galaxy population, producing broad-band magnitudes in any specified pass band. Hence, LRGs can be selected from the model output using the same colour and luminosity criteria that are applied to the real observational data. The models naturally predict which dark matter haloes host LRGs. As we will see later, a key element in shaping the "halo occupation distribution" of LRGs is the scatter in the merger histories of dark matter haloes, which has been ignored in previous analyses.

This work extends the previous chapter (see also Almeida et al., 2007) in the sense that LRGs are much rarer and significantly brighter than L_* galaxies and even than the early-type population as a whole. Here, we concentrate on massive red galaxies at low and intermediate redshifts where large observational samples exist. We remind the reader that the key features of the two models were presented in Chapter 2. In Section 4.2, we explain the selection of LRGs and show some basic predictions for the abundance and properties of LRGs. In Section 4.3, we present predictions for the clustering of LRGs and in Section 4.4 we examine how the stellar

mass of LRGs is built up in the models. Our conclusions are given in Section 4.5.

4.2   LRG selection and basic properties

In this section, we present the predictions of the Baugh et al. and Bower et al. models for the basic properties of low and intermediate redshift luminous red galaxies (LRGs) and compare these with observational results from the SDSS and 2SLAQ LRG samples. LRGs are a subset of the overall early-type population with extreme luminosities and colours, so it is essential to match their selection criteria as closely as possible in order to make a meaningful test of the model predictions. We begin by reviewing the colour and magnitude selection used in these surveys, before moving on to examine the predictions for the abundance of LRGs.

4.2.1   Sample selection: SDSS and 2SLAQ LRGs

The basic aim of LRG surveys is to select intrinsically bright galaxies which have colours consistent with those expected for a passively evolving stellar population (Eisenstein et al., 2001). The selection criteria used in the SDSS LRG and 2dF-SDSS LRG and QSO (2SLAQ) surveys are targetted at different redshift intervals and pick up very different number densities of objects. Full descriptions of the design of the respective surveys can be found in Eisenstein et al. (2001) and Cannon et al. (2006).

Below, for completeness, we give a summary of the colour and magnitude ranges which define the LRG samples. In the case of the observational samples, Petrosian magnitudes were used for apparent magnitude selection and SDSS model magnitudes were used for colour selection. The SDSS filter system is described in Fukugita et al. (1996). In the case of GALFORM galaxies, we use the total magnitude. We consider two output redshifts in the GALFORM models, chosen to be close to the median redshifts of the observational samples; $z = 0.24$ to compare with SDSS LRGs and $z = 0.50$ to match the 2SLAQ LRGs.

In the case of the SDSS, two combinations of the $g - r$ and $r - i$ colours are

formed:

$$c_\perp = (r - i) - (g - r)/4 - 0.177, \tag{4.1}$$

$$c_\parallel = 0.7(g - r) + 1.2[(r - i) - 0.177]. \tag{4.2}$$

The following conditions are then applied to select LRGs:

$$r_{\mathrm{petro}} < 19.2, \tag{4.3}$$

$$r_{\mathrm{petro}} < 13.116 + c_\parallel/0.3, \tag{4.4}$$

$$|c_\perp| < 0.2. \tag{4.5}$$

In the case of 2SLAQ, somewhat different colour combinations are used:

$$d_\perp = (r - i) - (g - r)/8, \tag{4.6}$$

$$d_\parallel = 0.7(g - r) + 1.2[(r - i) - 0.177]. \tag{4.7}$$

The selection criteria applied in the case of 2SLAQ are:

$$17.5 < i < 19.8, \tag{4.8}$$

$$0.5 < g - r < 3, \tag{4.9}$$

$$r - i < 2, \tag{4.10}$$

$$d_\perp > 0.65, \tag{4.11}$$

$$d_\parallel > 1.6. \tag{4.12}$$

The colour equations and the conditions applied to them are designed to locate galaxies with deep 4000 Å breaks in the $(g - r)$ vs $(r - i)$ plane over the redshift intervals of the two surveys (see Eisenstein et al., 2001; Cannon et al., 2006, for further details).

As we have already commented, these two sets of selection criteria give quite different number densities of LRGs. Here we *do not* attempt to tune the selection to match objects in the 2SLAQ LRG sample with those from the SDSS LRG sample. This was done by Wake et al. (2006), whose motivation was to study the evolution of the LRG luminosity function. Our aim instead is to test the galaxy formation models, so trying to match the selection to pick out similar objects between the two redshifts is not necessary.

4.2.2 Luminosity Function

The luminosity function is the most basic description of any galaxy population and is arguably the key hurdle for a model of galaxy formation to negotiate before considering other predictions. It is important to bear in mind that LRGs represent only a small fraction of the galaxy population as a whole, as can be seen by comparing the integrated space densities quoted in Table 4.1 with the abundance of L_* galaxies, which is around an order of magnitude higher. Reproducing the abundance of such rare galaxies therefore represents a strong challenge for any theoretical model.

In Fig. 4.1, we compare the predictions of the GALFORM models for the luminosity function of LRGs with observational estimates. This determination of the observed 2SLAQ and SDSS luminosity functions is different from that presented in Wake et al. (2006). Here, we have estimated the observed luminosity functions in such a way as to minimize the corrections necessary to compare to the models. We have restricted both samples to tight redshift ranges around the model output redshifts, $0.22 < z < 0.26$ for the case of SDSS and $0.48 < z < 0.54$ for 2SLAQ. We then use simple K+e corrections derived from Bruzual & Charlot (2003) stellar population synthesis models (see Wake et al., 2006) to correct the SDSS LRG magnitudes to $z = 0.24$ and the 2SLAQ LRG magnitudes to $z = 0.5$. Since the redshift ranges considered here are so close to the target redshift, these corrections are very small, < 0.01 mags. The SDSS catalogue is then cut at $i^{0.24} < 17.5$ and the 2SLAQ catalogue is cut at $i^{0.5} < 19.6$. Both of these cuts are brighter than the magnitude limits of each survey and within the limited redshift ranges effectively produce volume-limited samples. The final samples contain 5217 and 2576 LRGs within $8.5 \times 10^7 \mathrm{Mpc}^3$ and $3.9 \times 10^7 \mathrm{Mpc}^3$ for SDSS and 2SLAQ respectively [assuming $h = 0.7 = H_0/(100 \mathrm{kms}^{-1}\mathrm{Mpc}^{-1})$]. Since the samples are approximately volume-limited it is trivial to produce the luminosity function including a correction for incompleteness in each survey (see Wake et al., 2006). The integrated number densities of LRGs in the two surveys are listed in Table 4.1 after applying the respective completeness corrections.

In view of the fact that no model parameters have been adjusted in order to "tune" the predictions to better match observations of LRGs, both models come surprisingly close to matching the number density of LRGs in the SDSS sample, as
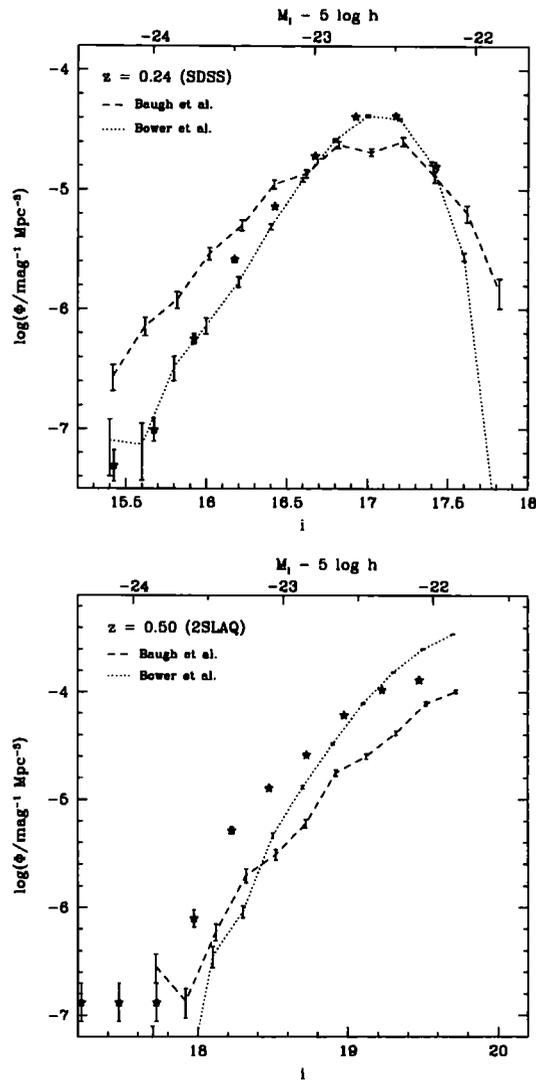
Figure 4.1: The luminosity function of luminous red galaxies plotted as a function of apparent magnitude. The upper panel shows the SDSS LRG luminosity function at $z = 0.24$, the median redshift of this sample, and the lower panel shows the results for the 2SLAQ sample at its median redshift, $z = 0.5$. The upper axis labels show the absolute magnitude for reference (calculated from the apparent magnitude by subtracting the appropriate distance modulus for each panel). In each panel, the dashed line shows the number density of LRGs predicted by the Baugh et al. model and the dotted line shows the prediction of the Bower et al. model. The observational estimates are shown by the stars. The error bars on the model predictions show the Poisson error due to the finite number of galaxies simulated.

Sample	SDSS ($z = 0.24$) (10^{-5} Mpc^{-3})	2SLAQ ($z = 0.50$) (10^{-5} Mpc^{-3})
Observed space density	3.30	8.56
Baugh et al. prediction	3.74	4.75
Bower et al. prediction	2.99	18.11

Table 4.1: The space density of LRGs in the SDSS and 2SLAQ samples estimated at their median redshifts, compared with the predictions of the Baugh et al. and Bower et al. models. The number density in the table are quoted in units of 10^{-5} Mpc^{-3} (the relevant h is used for each model; for the data, $h = 0.7$ is assumed).

Table 4.1 shows. In fact, the Baugh et al. model slightly overpredicts the space density of LRGs at $z = 0.24$ by 13%, whereas the Bower et al. model underpredicts only by 10%. However, Fig. 4.1 shows that the Baugh et al. actually gives a poor match to the shape of the luminosity function, predicting too many bright LRGs. Although the difference looks dramatic on a logarithmic scale, the discrepancy has little impact on the integrated space density.

At the median redshift of the 2SLAQ sample, the comparison with the observational estimate of the luminosity function of LRGs is less impressive. The Baugh et al. model now underpredicts the abundance of LRGs by a factor of 2. Alternatively, the discrepancy is equivalent to a shift of about one magnitude in the i-band. The Bower et al. model fares analogously, predicting around 100% more LRGs than are seen in the 2SLAQ sample. This suggests that neither model is able to accurately follow the evolution of the luminosity function of very red galaxies over such a large lookback time (when the age of the universe is only around 60% of its present day value). The abundance of LRGs is therefore quite sensitive to the way in which feedback processes are implemented in massive haloes.

4.2.3 Properties of LRGs

In this section, we present a range of predictions for the properties of galaxies which satisfy the LRG selection criteria defined in Section 4.2, for both the Baugh et al. and Bower et al. models, comparing with observational results whenever possible. We remind the reader that the models do not reproduce exactly the shape and normalization of the observed luminosity function of LRGs as seen in the previous subsection, but instead bracket the observed abundances. Rather than perturb the selection criteria applied the model galaxies to better match the observed abundances, we have retained the full LRG selection criteria so that the model galaxies have the same colours and magnitudes as observed LRGs.

Stellar mass

The predicted stellar masses of LRGs are plotted in Fig. 4.2. As expected from the high luminosities of LRGs, these galaxies exhibit large stellar masses. At $z = 0.24$ the stellar masses range from $\sim 1 \times 10^{11}$ to 5×10^{11} $h^{-1}M_{\odot}$, with a median of 1.7×10^{11} $h^{-1}M_{\odot}$. At $z = 0.50$, the distribution shifts to lower stellar masses, with a median value of $\sim 1 \times 10^{11}$ $h^{-1}M_{\odot}$. The scatter in the distribution of stellar masses predicted by the Baugh et al. model is somewhat larger than that in the Bower et al. model at $z = 0.24$. The median stellar mass is a property for which the two models agree closely, indicating that stellar mass is a robust prediction which is fairly insensitive to the details of the implementation of the physics of galaxy formation at these redshifts. The difference in the selection criteria applied to the two samples is responsible for picking up objects of quite different stellar masses. We shall see in subsequent comparisons that this basic difference between the LRG samples is responsible for differences in other model predictions.

Morphological mix

The bulge-to-total luminosity ratio, B/T, is often used as an indicator of the morphological type of a galaxy in semi-analytical models (see e.g. Baugh et al., 1996). The B/T ratio is correlated with Hubble T-type, a subjective classification parameter relying upon the identification of features such as spiral arms and galactic bars,
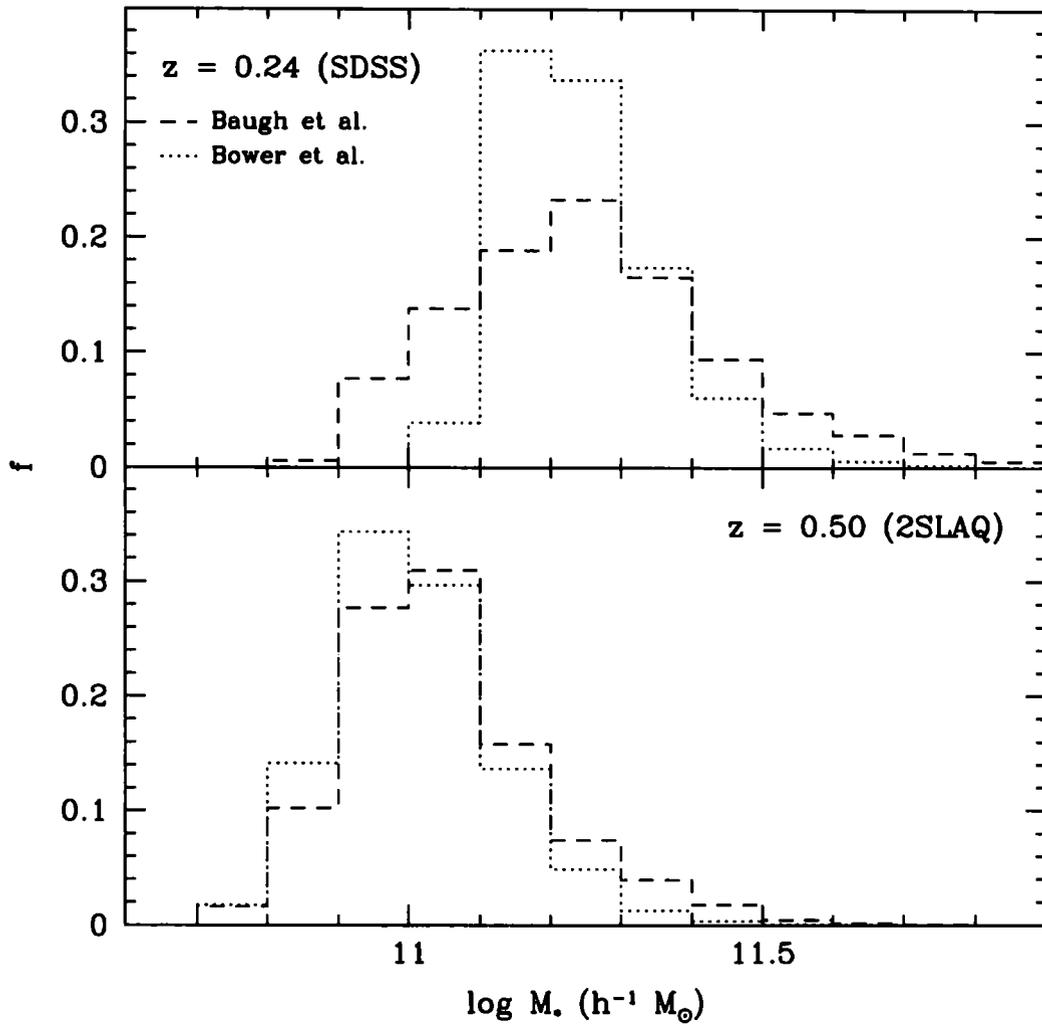
Figure 4.2: The predicted stellar masses of model luminous red galaxies at $z = 0.24$ (upper panel) and $z = 0.50$ (lower panel). The predictions of the Baugh et al. model are shown by the dashed histograms and those of the Bower et al. model by the dotted histograms. The distributions are normalized to give $\sum_i f_i = 1$.
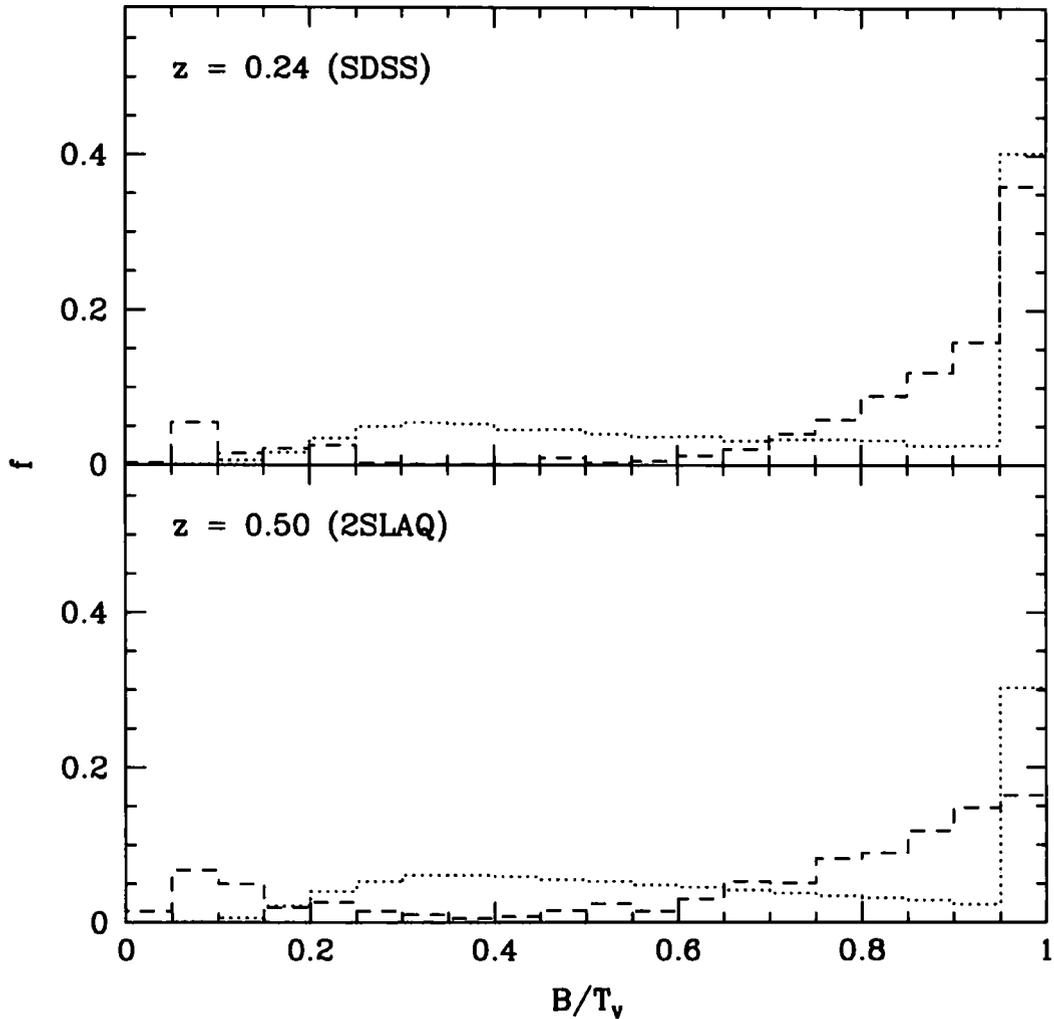
Figure 4.3: The predicted bulge-to-total luminosity ratio in the rest-frame V-band for luminous red galaxies. The upper panel shows the predictions of the Baugh et al. (dashed histogram) and Bower et al. models (dotted histogram) at $z = 0.24$. The lower panel displays the distributions predicted for LRGs at $z = 0.50$. The distributions are normalized to give $\sum_i f_i = 1$.

Redshift	Baugh et al.	Bower et al.
	[0,0.4]:[0.4,0.6]:[0.6,1.0]	[0,0.4]:[0.4,0.6]:[0.6,1.0]
$z = 0.24$	12:2:86	21:17:62
$z = 0.50$	21:6:73	24:22:54

Table 4.2: The predicted morphological mix of luminous red galaxies, at $z = 0.24$ and $z = 0.50$, for the Baugh et al. and Bower et al. models. The three values quoted for each model show the percentage of galaxies with bulge-to-total luminosity ratios of $B/T < 0.4$, $0.4 \leq B/T \leq 0.6$ and $B/T > 0.6$.

though there is considerable scatter around this relation (Simien & de Vaucouleurs, 1986). In the B-band, galaxies with $B/T < 0.4$ correspond approximately to the T-types of late-type or spiral galaxies, those with $B/T > 0.6$ overlap most with elliptical galaxies and the intermediate values, $0.4 < B/T < 0.6$ correspond to lenticulars. In Fig. 4.3 we plot the predicted distribution of the bulge-to-total ratio in the rest-frame V-band, B/T_V, for the Baugh et al. and Bower et al. models, at $z = 0.24$ (upper panel) and $z = 0.50$ (lower panel). The results are also summarized in Table 4.2, where the fraction of galaxies in intervals of B/T_V ratio are calculated.

In both models, the SDSS and 2SLAQ LRG samples are predicted to be mainly composed of bulge-dominated galaxies, which account for more than $\sim 60\%$ of the LRG population. However, the models suggest that the LRG samples contain an appreciable fraction ($\sim 20\%$) of late-type, disk-dominated systems. These galaxies meet the LRG colour selection criteria primarily because they have old stellar populations. Another prediction is that the fraction of bulge dominated galaxies is higher in the SDSS sample than in the 2SLAQ sample.

The distributions of B/T ratios predicted by the two models show substantial differences, particularly in the intermediate ratio range, which corresponds roughly to S0 types. This difference is not due to any single model ingredient, but is more likey to be the result of the interplay between several phenomena. As outlined in the Chapter 2, both models invoke galaxy mergers as a mechanism for making spheroids, either by the rearrangement of stellar discs or through triggering additional star for-

mation. Bower et al. also consider starbursts resulting from discs being dynamically unstable to bar formation. In the Baugh et al. model, around 30% of the total star formation takes place in merger driven starbursts. This figure is much lower in the Bower et al. case, because galactic discs tend to be gas poor at high redshift in this model. In the Bower et al. model, starbursts resulting from the collapse of unstable discs dominate over bursts driven by galaxy mergers. Baugh et al. allow minor mergers to trigger starbursts, if the primary disc is gas rich. We have tested that removing these starbursts does not have a major impact on the distribution of B/T values.

Stellar populations

The luminosity-weighted age of a stellar population is a measure of the age of the stars in a galaxy. The predicted distributions of the rest-frame V-band luminosity-weighted age of LRGs are plotted in Fig. 4.4. This plot reveals that LRGs have stellar populations with luminosity-weighted ages ranging from 4 to 8 Gyr in the Bower model, and from 2 to 6 Gyr in the Baugh et al. model, at $z = 0.24$ (i.e. when the universe was $\sim 80\%$ of its current age). At $z = 0.50$ (60% of the current age of the universe), galaxies in the Bower et al. model show, again, older stellar ages than those in the Baugh et al. model: the median of the distribution for LRGs in the Bower et al. model is ≈ 4.2 Gyr, whereas for the Baugh et al. model it is 3.2 Gyr. The difference in the age of the universe between these redshifts is around 2 Gyr (the value is slightly different for each model due to the different choice of the values of the cosmological parameters), and therefore accounts for the bulk of the difference in the ages of the SDSS and 2SLAQ samples. The model LRGs are therefore composed predominantly of old stellar populations and resemble those of observed early-type galaxies (e.g. Trager et al., 2000; Gallazzi et al., 2006). Furthermore, our results are in excellent agreement with the analysis by Barber et al. (2007). Barber et al. used stellar population synthesis models to fit the spectra of 4391 LRGs from the SDSS, finding matches for ages in the range from 2 to 10 Gyrs, with a peak in the distribution around 6 Gyrs. The difference in the ages predicted by the Baugh et al. and Bower et al. models has its origins in the different implementations of gas
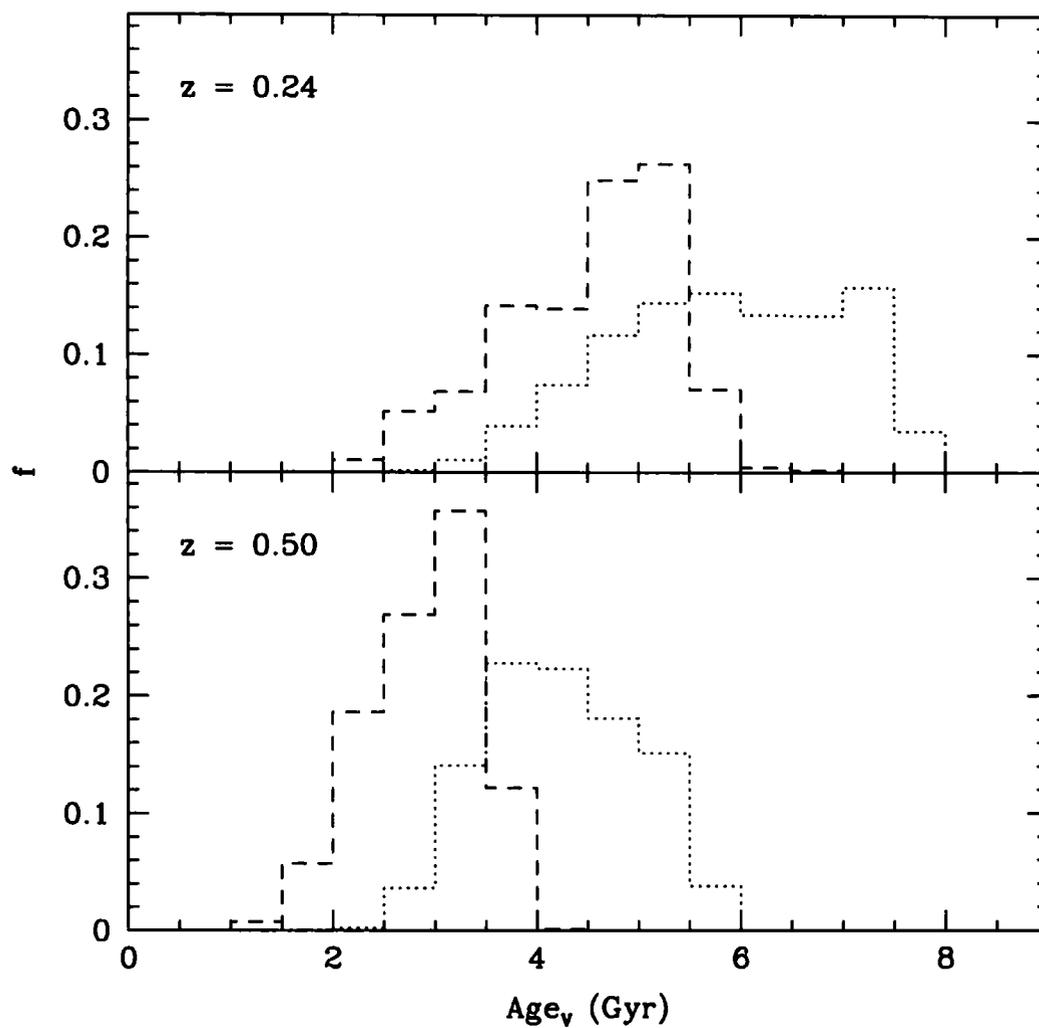
Figure 4.4: The rest-frame V-band luminosity-weighted age of the stellar populations of LRGs at $z = 0.24$ (upper panel) and $z = 0.5$ (lower panel). As before, the predictions of the Baugh et al. model are shown by the dashed histograms and those of the Bower et al. model by the dotted histograms. The histograms are normalized as in Fig. 4.2.

cooling and feedback applied in massive dark matter haloes. De Lucia et al. (2006) show that the suppression of gas cooling due to AGN feedback tends to increase the age of the stellar population in the galaxies hosted by haloes with quasistatic hot gas atmospheres. Note that in hierarchical models, the age of the stars in a galaxy is not the same as the age of the galaxy: the age of the stellar population can greatly exceed the age of the galaxy, with stars forming in the galaxy's progenitors, which are later assembled into the final galaxy through mergers (e.g. Baugh et al., 1996; Kauffmann, 1996; De Lucia & Blaizot, 2007). We revisit this point in Section 4.4.

Fig. 4.5 shows the predicted distribution of the V-band luminosity-weighted metallicity for LRGs. There is little change in the luminosity weighted-metallicity between the two LRG samples in the Baugh et al. model. In the Bower et al. model, there is a modest decrease in metallicity of almost +0.2dex between $z = 0.24$ and $z = 0.5$. To help in the interpretation of these predictions it is instructive to plot the metallicity–stellar mass relation for spheroids at different redshifts (N.B. here, we consider any galaxy with a bulge-to-total stellar mass ratio in excess of 0.6, not just LRGs; we note that the metallicity–stellar mass relation is similar for galaxies with bulge-to-total ratios below 0.6). We recall that the typical stellar mass of LRGs is predicted to change by a factor of two between the $z = 0.5$ and $z = 0.24$ samples, from $\log(M_*/h^{-1}M_\odot) \sim 11$ to $\log(M_*/h^{-1}M_\odot) \sim 11.3$.

The predicted evolution of the stellar mass–metallicity relation for the Baugh et al. and Bower et al. models is shown in Fig. 4.6. There is little evolution in the locus of the metallicity–mass relation. Between $z = 0.5$ and $z = 0.24$, the metallicity–stellar mass relation in the Baugh et al. model flattens at the high mass end. Hence, the change in metallicity expected due to an increase in stellar mass by a factor of 2 using the metallicity–mass relation predicted at $z = 0.5$ is largely cancelled out by the change in slope of the metallicity – mass relation at $z = 0.24$. In the Bower et al. model, the metallicity–mass relation has a kink at $\log(M_*/h^{-1}M_\odot) \sim 11$, and is flat for the most massive galaxies over the whole of the redshift range plotted in Fig. 4.6. The evolution seen in the Bower et al. model is therefore due to the presence at $z = 0.5$ of LRGs with masses $< 10^{11}h^{-1}M_\odot$ which come from the steep part of the metallicity–mass relation; at $z = 0.24$, only LRGs from the flat part of
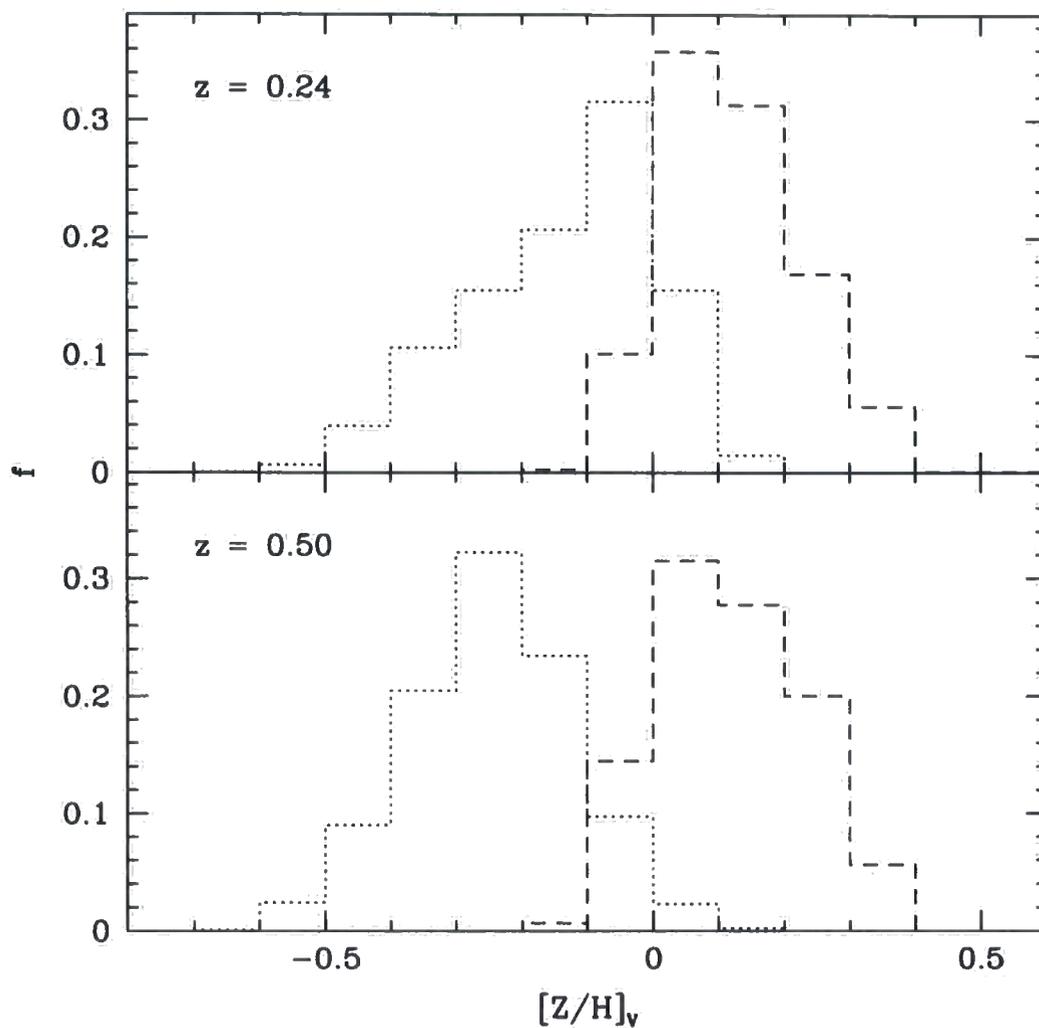
Figure 4.5: The predicted distribution of the rest-frame V-band luminosity-weighted metallicity of LRGs. The upper panel displays the Baugh et al. (dashed histogram) and Bower et al. (dotted histogram) models at $z = 0.24$ and the lower panel shows the predictions at $z = 0.50$. The histograms are normalized as in Fig. 4.2.
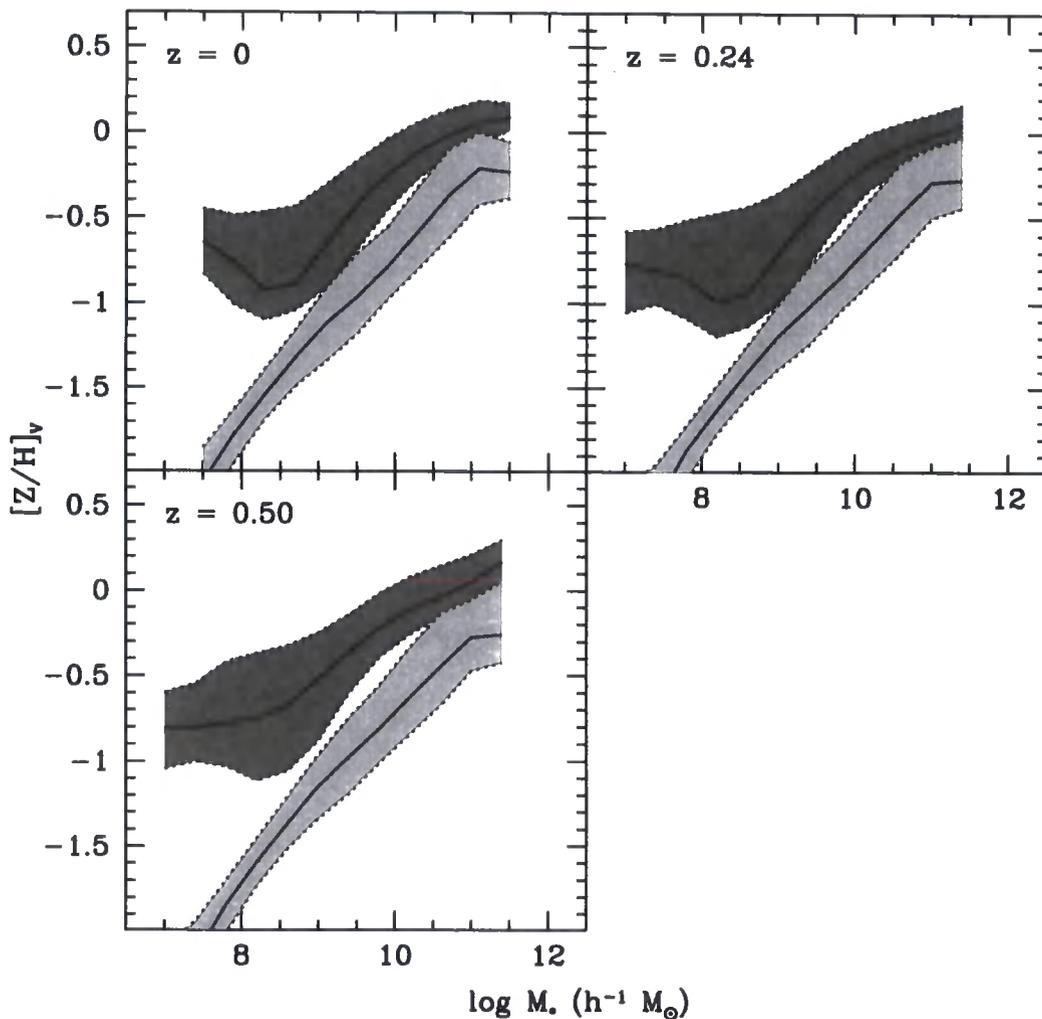
Figure 4.6: The rest-frame V-band luminosity weighted metallicity – stellar mass relation for ellipticals (i.e. galaxies with a bulge-to-total stellar mass ratio greater than 0.6). The light and dark gray shaded regions show the distributions for the Bower et al. and the Baugh et al. models, respectively. Each panel corresponds to a different redshift, as indicated by the key. The solid line shows the median metallicity for stellar mass bins. The shaded regions are enclosed by the 10 and 90 percentiles of the distribution.

Redshift	Baugh et al.	Bower et al.
$z = 0.24$	0.25	0.32
$z = 0.50$	0.21	0.30

Table 4.3: The predicted fraction of satellite galaxies in the luminous red galaxies samples at $z = 0.24$ and $z = 0.50$.

the metallicity–mass relation are sampled due to the increase in stellar mass.

Fig. 4.5 shows that the Bower et al. model displays a different metallicity distribution from the Baugh et al. model, predicting luminosity-weighted metallicities lower by a factor of ~ 2. This difference is entirely due to the choice of IMF used in the models. We remind the reader that, in the Baugh et al. model, stars which form in merger driven bursts are assumed to be produced with a flat IMF, whereas in the Bower et al. model, a Kennicutt (1983) IMF is adopted in all modes of star formation. The yield adopted is consistent with the choice of IMF. For a flat IMF, the yield is over six time larger than the yield expected from a Kennicutt IMF. As noted by Nagashima et al. (2005a), the metal abundances for galaxies in the Baugh et al. model are higher by a factor of 2-3 than is the case for a model using a Kennicutt IMF. Intriguingly, Barber et al. (2007) also favour high metallicities in their simple fits to the spectra of SDSS LRGs, finding best fitting models in the range $-0.6 \leq [Z/H] \leq 0.4$, which they argue is evidence in favour of LRGs forming stars with a top-heavy IMF.

Are LRGs central or satellite galaxies?

In the models, the most massive galaxy within a dark matter halo is referred to as the "central" galaxy and any other galaxies which also reside in the halo are referred to as "satellites" (see Baugh 2006). In the majority of semi-analytical models, this distinction is important because gas which cools from the hot gas halo is directed onto the central galaxy, and satellite galaxies can merge only with the central galaxy. The fraction of luminous red galaxies which are satellite galaxies in the models is given in Table 4.3; more than 25% of LRGs at $z = 0.24$ are satellite galaxies with

a slight decrease in this fraction for the 2SLAQ sample. The fraction of satellite galaxies has important consequences for the small scale clustering of LRGs, as we shall discuss in next Section.

Radii

The distribution of the radii of model LRGs is shown in Fig. 4.7. We plot the half-mass radius of the model galaxies, taking the mass weighted average of the disk and bulge components. The calculation of the linear sizes of galaxies in GALFORM takes into account the conservation of the angular momentum of cooling gas and the conservation of energy of merging galaxies (Cole et al., 2000). This prescription was tested against observations of bulge dominated SDSS galaxies in Chapter 3 (see also Almeida et al., 2007). Overall, we found that the predicted sizes of spheroids in the Baugh et al. model matched the observed sizes reasonably well, except for galaxies much brighter than L_*. These bright galaxies are predicted to be a factor of up to three smaller than observed by Bernardi et al. (2005). In the Bower et al. model, the brightest spheroids are predicted to be even smaller than in the Baugh et al. model. The same trend is seen in the predictions for the sizes of LRGs shown in Fig. 4.7. At $z = 0.24$, the median of the distribution of LRG half-light radii in the Baugh et al. model occurs at $\sim 1.8h^{-1}$kpc. In the Bower et al. model, the median is at a radius that is around a factor of 2 smaller. At $z = 0.5$, the median of the two model distributions differ by a smaller factor, ≈ 1.8, although there is little evolution in the distributions from $z = 0.5$ to $z = 0.24$. The observed radii of SDSS LRGs (as extracted from the online database) are larger than the model predictions, with a median de Vaucouleur's radius of $6.6h^{-1}$kpc. The observational estimate of the LRG radius is obtained by fitting a de Vaucouleur's profile convolved with a seeing disc. In this case the seeing is restricted to be no worse than 1.4 arcseconds, which corresponds to a scale of $3.7h^{-1}$kpc at the median redshift of the SDSS sample. Thus the tail of small scale-length galaxies predicted by the models would not be observable. However, the observed distribution of LRG sizes has few galaxies close to the seeing limit, so this is does not affect the estimation of the median size of SDSS LRGs.
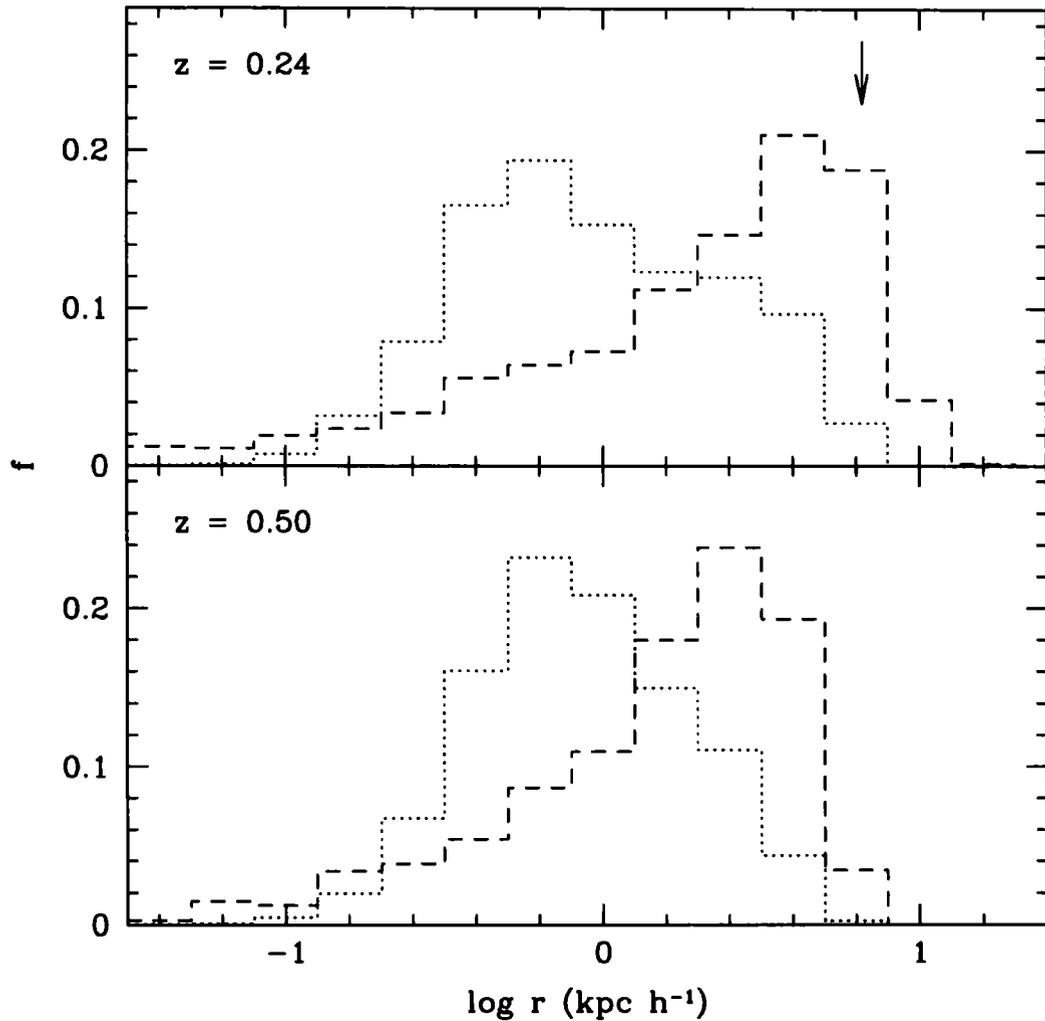
Figure 4.7: The predicted half mass radii of luminous red galaxies. The upper and lower panels display the predictions of the Baugh et al. (dashed histograms) and Bower et al. models (dotted histograms), at $z = 0.24$ and $z = 0.5$ respectively. The histograms are normalized as in Fig. 4.2. The arrow shows the median de Vaucouleur's radius of the observed SDSS LRGs.

Velocity dispersion

Fig. 4.8 shows the distribution of the one-dimensional velocity dispersion of the bulge component of model LRGs, σ_{1D}. This is calculated from the effective circular velocity of the bulge, $V_{c,bulge}$, using $\sigma_{1D} = (1.1/\sqrt{3})V_{c,bulge}$, where σ_{1D} is assumed to be isotropic. The circular velocity at the half mass radius of the bulge is a model output which is computed taking into account the angular momentum and mass of the disk and bulge, and the gravitational contribution of the baryons and dark matter (see Chapter 2 or Cole et al., 2000, for further details). The factor of 1.1 is an empirical adjustment introduced by Cole et al. (1994) which we have retained to facilitate comparison with predictions for the more general population of spheroids presented in the previous chapter. At $z = 0.24$, both models predict velocity dispersions in the range 220-400kms^{-1}, with a median around 320kms^{-1}. Between $z = 0.24$ and $z = 0.5$, the predicted distribution of LRG velocity dispersions shifts to lower values by $\Delta \log \sigma \approx 0.1$. The bulk of this evolution is due to the change in stellar mass between the LRG samples (see Fig. 3.16). The median velocity dispersion for SDSS LRGs is $\sigma = 250$kms^{-1}, which is somewhat smaller than the model predictions.

4.2.4 Why do the two models give different predictions?

In §4.2.3, we demonstrated that the predictions of the Bower et al. model for the luminosity function of LRGs are in better agreement with the observations than those of the Baugh et al. model. In particular, at $z = 0.24$, the Bower et al. model gives a very good match both to the shape of the observed luminosity function and the integrated number density of LRGs, matching the observed density at the 10% level. The Baugh et al. model predicts too many LRGs at this redshift, particularly at the bright end.

We saw in Chapter 2 that there are several areas in which the input physics and parameter choices differ between the two models. Although our aim in this Chapter is to test published models and not to tweak the results to fit the LRG population, it is instructive to vary some of the parameters in the Baugh et al.
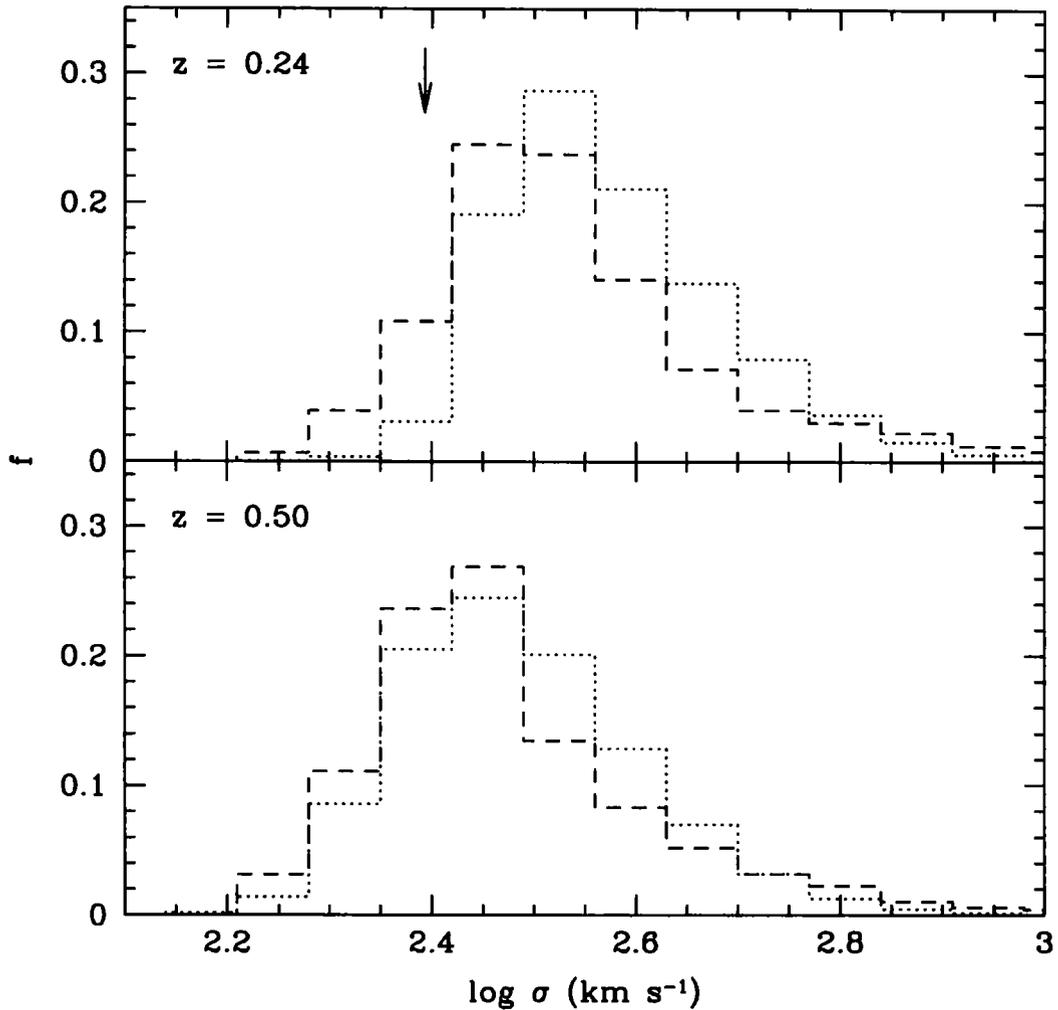
Figure 4.8: The predicted one-dimensional velocity dispersion of luminous red galaxies in the Baugh et al. (dashed lines) and Bower et al. models (dotted lines). The upper panel shows the distributions at $z = 0.24$ and the lower panel at $z = 0.50$. The histograms are normalized as in Fig. 4.2. The median velocity dispersion of the observed SDSS LRGs is shown by the arrow.

model to see if the predictions for the number of LRGs improve. We varied several model ingredients (e.g. strength of superwind feedback, star formation timescale, burst duration, choice of IMF in bursts, criteria for triggering a starburst following a galaxy merger) and found that in all cases, the resulting change in the luminosity function of LRGs was driven by a change in the overall luminosity function, i.e. if the number of bright LRGs increased, the luminosity function of all galaxies was found to brighten by a similar amount. Since a model is only deemed successful if it reproduces as closely as possible the overall galaxy luminosity function, none of these variant models is acceptable without further parameter changes to reconcile the overall luminosity function with observations. Hence, the apparent gain in the abundance of LRGs is likely to be cancelled out by the additional parameter changes which compensate for the brightening of the overall luminosity function. We note that adopting a hot gas density profile with a fixed rather than evolving core radius in the Baugh et al. model does not improve the predictions for the abundance of LRGs. With a fixed core radius, more gas cools in massive haloes than in the evolving core case, which leads to more bright galaxies (see Cole et al., 2000). However, these galaxies are also bluer and so do not match the LRG selection.

The success of the Bower et al. model can be traced to the revised gas cooling prescription adopted in massive haloes. The suppression of the cooling flow in haloes with a quasistatic hot atmosphere and the dependence of this phenomenon on redshift are the key reasons why this model matches the evolution of the LRG luminosity function better than the Baugh et al. model. LRGs in the Bower et al. model are older than their counterparts in the Baugh et al. model, because the supply of cold gas for star formation is removed. In the Baugh et al. model, the superwind feedback acts to effectively suppress star formation in massive haloes, but still allows some star formation to take place. The choice of a top-heavy IMF in starbursts helps the Baugh et al. model, to mask the recent star formation to some extent, by making the LRG stellar population more metal rich and thus redder.

4.3 The clustering of LRGs

The clustering of galaxies is an invaluable constraint on theoretical models of galaxy formation. The form and amplitude of the two-point galaxy correlation function is driven by three main factors, which play different roles on different length scales: the clustering of the underlying dark matter, the distribution or partitioning of galaxies between dark matter haloes (e.g. Benson et al., 2000; Seljak, 2000; Peacock & Smith, 2000) and the distribution of galaxies within haloes. The number of galaxies as a function of halo mass, called the halo occupation distribution (HOD), controls how the correlation function of galaxies is related to that of the underlying matter (see the review by Cooray & Sheth, 2002). On large scales, the correlation functions of the galaxies and matter have similar shapes, but differ in amplitude by the square of the effective bias. On smaller scales, comparable to the radii of the typical haloes which host the galaxies of interest, it is the number of galaxies and its radial distribution within the same dark matter halo which sets the form and amplitude of the correlation function (see Fig. 10 of Benson et al., 2000). The clustering of the galaxies can be different from that of the matter on small scales as well as large. The predictions for the correlation function in redshift-space can also be affected by the peculiar motions of galaxies as we shall see later on in this section.

Semi-analytical models naturally predict which dark matter haloes contain LRGs. Fig. 4.9 shows the range of dark halo masses that host LRGs in the Baugh et al. and Bower et al. models. Far from being restricted to cluster-mass haloes, in the models, LRGs can occur in a wide range of halo masses, including fairly modest haloes comparable in mass to the halo which is thought to host the Milky Way. This plot reveals important differences between the predictions of the two models. At $z = 0.24$, LRGs in the Bower et al. model occupy haloes with masses in the range $1 \times 10^{12} - 1 \times 10^{15} h^{-1} M_\odot$, with a median of $\sim 1 \times 10^{13} h^{-1} M_\odot$. The Baugh et al. model predicts that LRGs are to be found in haloes which are a factor of ~ 3 more massive than in the Bower et al. model, with the median of the distribution occuring at $3 \times 10^{13} h^{-1} M_\odot$. The prediction from the Baugh et al. model is in excellent agreement with the halo mass estimated for SDSS LRGs using weak
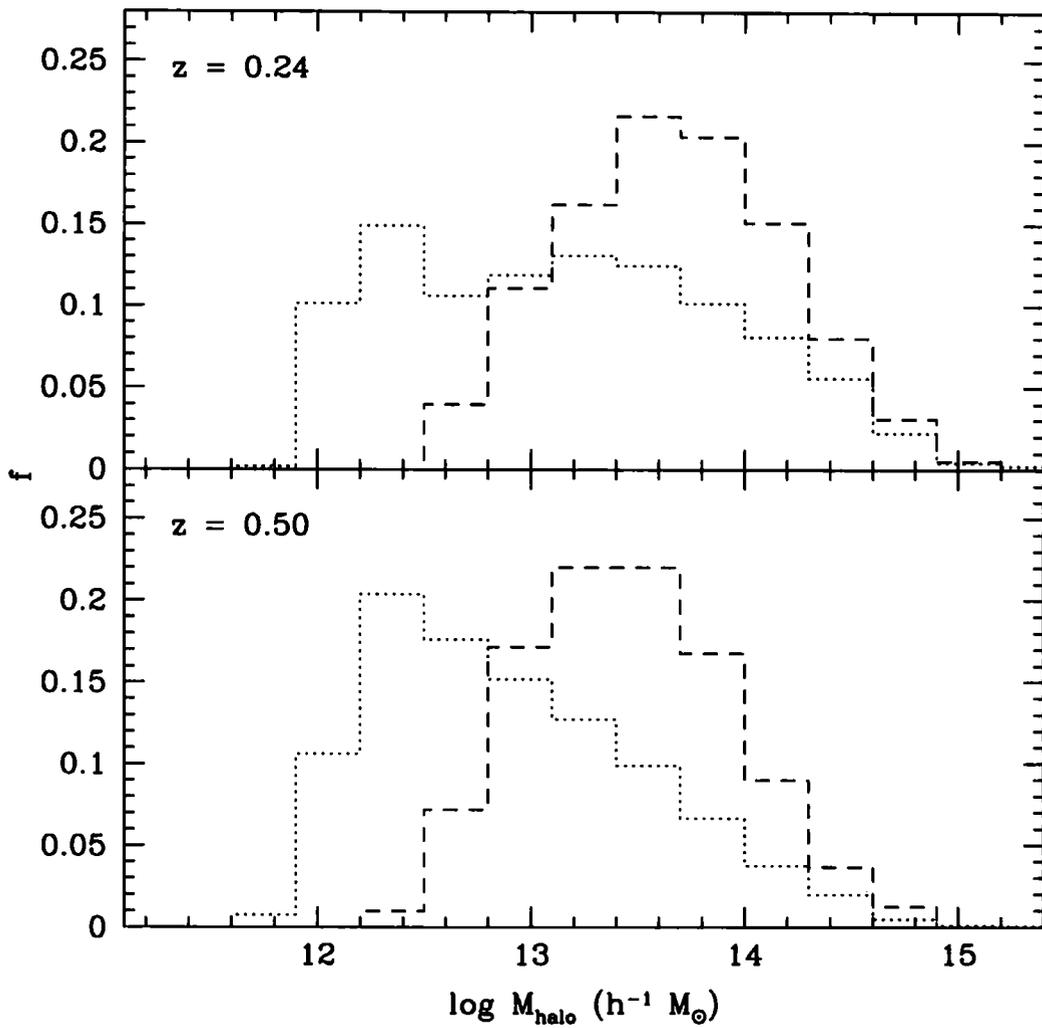
Figure 4.9: The distribution of the masses of haloes which host LRGs at $z = 0.24$ (upper panel) and $z = 0.5$ (lower panel). As before, the dashed histogram shows the predictions of the Baugh et al. model and the dotted line shows the Bower et al. model. The histograms are normalized as in Fig. 4.2.

lensing measurements (Mandelbaum et al., 2006). Despite the differences in the sample selection, there is little evolution with redshift in the distribution of host halo masses from $z = 0.24$ to $z = 0.5$, with a slight shift to lower halo masses seen at the higher redshift. The difference in the range of halo masses predicted to host LRGs will have an impact on the amplitude of the LRG correlation function, and, thus a measurement of the clustering of LRGs can potentially discriminate between the two models.

The distribution of halo masses hosting LRGs plotted in Fig. 4.9 is determined by two factors: the adundance of dark matter haloes, which is a strong function of halo mass for the typical hosts of LRGs, and the number of LRGs which occupy the same dark halo. These factors are separated in Fig. 4.10, in which we compare the overall mass function of dark haloes with the mass function of haloes weighted by the number of LRGs they contain. The solid lines show the mass function of dark matter haloes, which is determined by the values of the cosmological parameters, the form and amplitude of the power spectrum of matter fluctuations and the redshift (e.g. Governato et al., 1999). The dashed lines show the mass function of haloes multiplied by the average number of LRGs as a function of halo mass. At the mass where the solid and dashed lines cross, these haloes host on average one LRG. At lower masses, the mean number of LRGs per halo rapidly falls below one (this is the ratio of the abundance indicated by the dashed line divided by the abundance shown by the solid line). There is a threshold mass which must be reached before there is any possibility of a halo hosting an LRG. At this mass, only a tiny fraction of haloes actually contain an LRG. Nevertheless, these low mass haloes, because they are much more abundant than the more massive haloes, which have a higher mean number of LRGS, make an important contribution to the overall space density of LRGs. Hence it is essential for a model to take into account the scatter in the formation histories of dark matter haloes in order to accurately model the space density and clustering of LRGs. In the high mass tail of the mass function, the amplitude of the dashed curve exceeds that of the solid curve; at these masses, haloes host more than one LRG. Fig. 4.10 shows that, at $z = 0.24$, in the Baugh et al. model there are an average of ~ 3 SDSS LRGs per halo at $M_{halo} \approx 1 \times 10^{15} \, h^{-1} M_{\odot}$. At $z = 0.5$, the
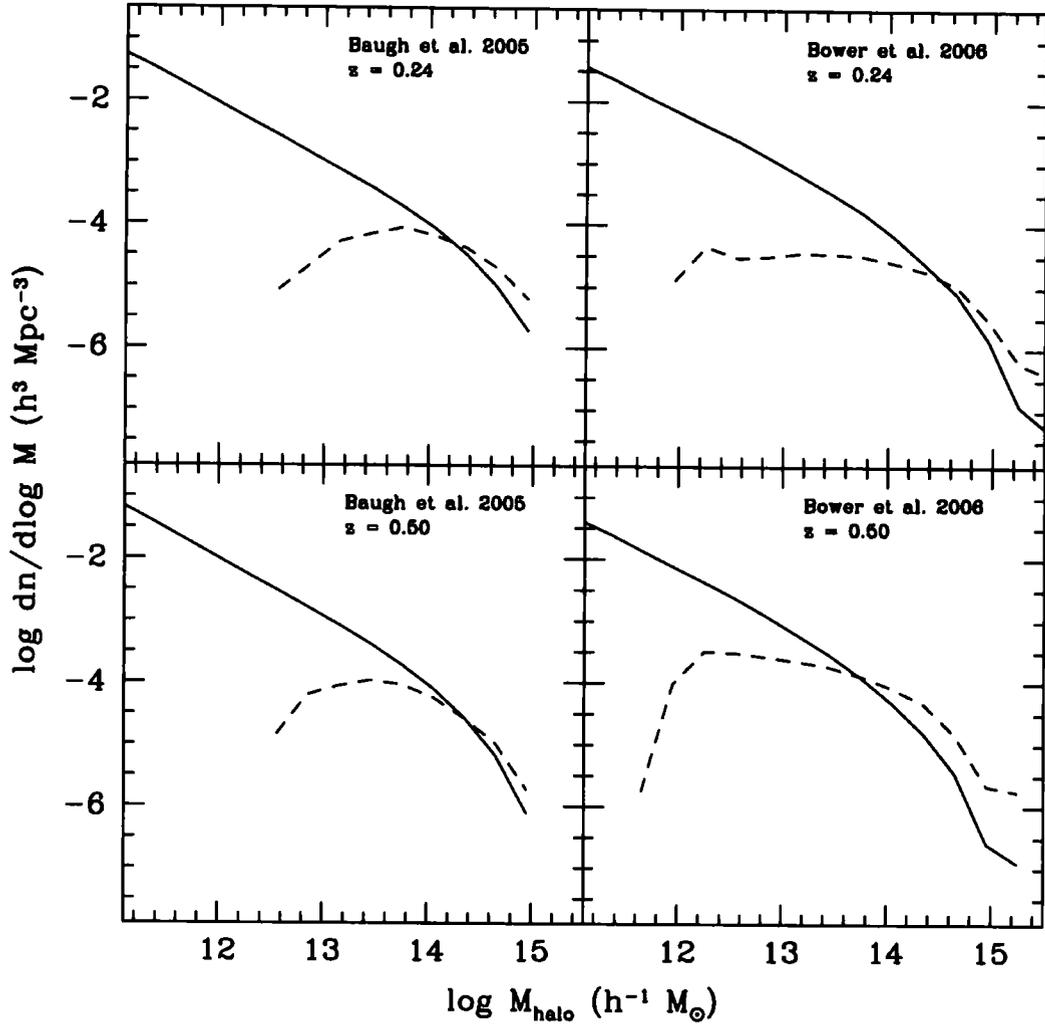
Figure 4.10: The mass function of dark haloes (solid lines) and of haloes which host an LRG (dashed lines). The contribution of each halo to the mass function plotted with a dashed line is directly proportional to the number of LRGs it contains; where the dashed line has a greater amplitude than the solid line, this means that haloes of that mass contain more than one LRG on average. The left-hand panels show the predictions for the Baugh et al. model and the right-hand panels for the Bower et al. model. The top row corresponds to $z = 0.24$ and the bottom row to $z = 0.5$.

mass of haloes which host on average one LRG ($M_{halo} \approx 2.5 \times 10^{14} \, h^{-1} M_{\odot}$) is higher than at $z = 0.24$ and the most massive haloes do not contain as many LRGs as the most massive haloes present at $z = 0.24$. This figure also shows that the Bower et al. model predicts a higher mean number of LRGs per halo than the Baugh et al. model at $z = 0.5$; for haloes of mass $M_{halo} \approx 1 \times 10^{15} \, h^{-1} M_{\odot}$, the Bower et al. model predicts an average of 10 LRGs per halo. The multiple occupancy of LRGs in high mass haloes has important consequences for the form of the predicted correlation function at small pair separations (i.e. $r < 1 h^{-1} \mathrm{Mpc}$).

The mean number of LRGs as a function of halo mass in the halo occupation distribution (HOD), as predicted by GALFORM, is plotted in Fig. 4.11. For comparison, we also plot the function quoted as a description of the HOD for SDSS LRGs by Masjedi et al. (2006), which is reproduced in each panel of Fig. 4.11 to serve as a reference point (dotted line). This measure has been used to quantify statistically the relationship between haloes and galaxies, and basically informs about the probability of finding $< N_{\mathrm{LRG}} >$ LRGs in a halo of mass M_{halo}. Observationally, this occupation distribution can be calculated using data on the clustering of galaxies. The Masjedi et al. HOD was not derived by fitting the model correlation functions to that measured for LRGs. Instead, this is simply the HOD for the brightest luminosity bin of the main SDSS galaxy sample analyzed by Zehavi et al. (2005). Although galaxies in this sample have similar luminosities and colours to LRGs, they have a different redshift distribution and the selection is only crudely matched. Fig. 4.11 shows that whilst this parametric form for the HOD is a reasonable match to that predicted by the models for massive haloes with more than 1 LRG, it is a poor description at lower masses. As we have argued before, even though low mass haloes host a mean number of LRGs below unity, there are more of them than there are high mass haloes, so these objects make a significant contribution to the clustering signal (see a similar discussion in Baugh et al., 1999). Ho et al. (2008) estimated the HOD for a sample of LRGs in clusters and found a result similar to the predictions of the Bower et al. model at $z = 0.5$. A more detailed modelling of the transition from 1 LRG per halo to 0 LRGs per halo is required to describe the model predictions, such as that advocated by Zheng et al. (2005). Wake et al. (2008) have
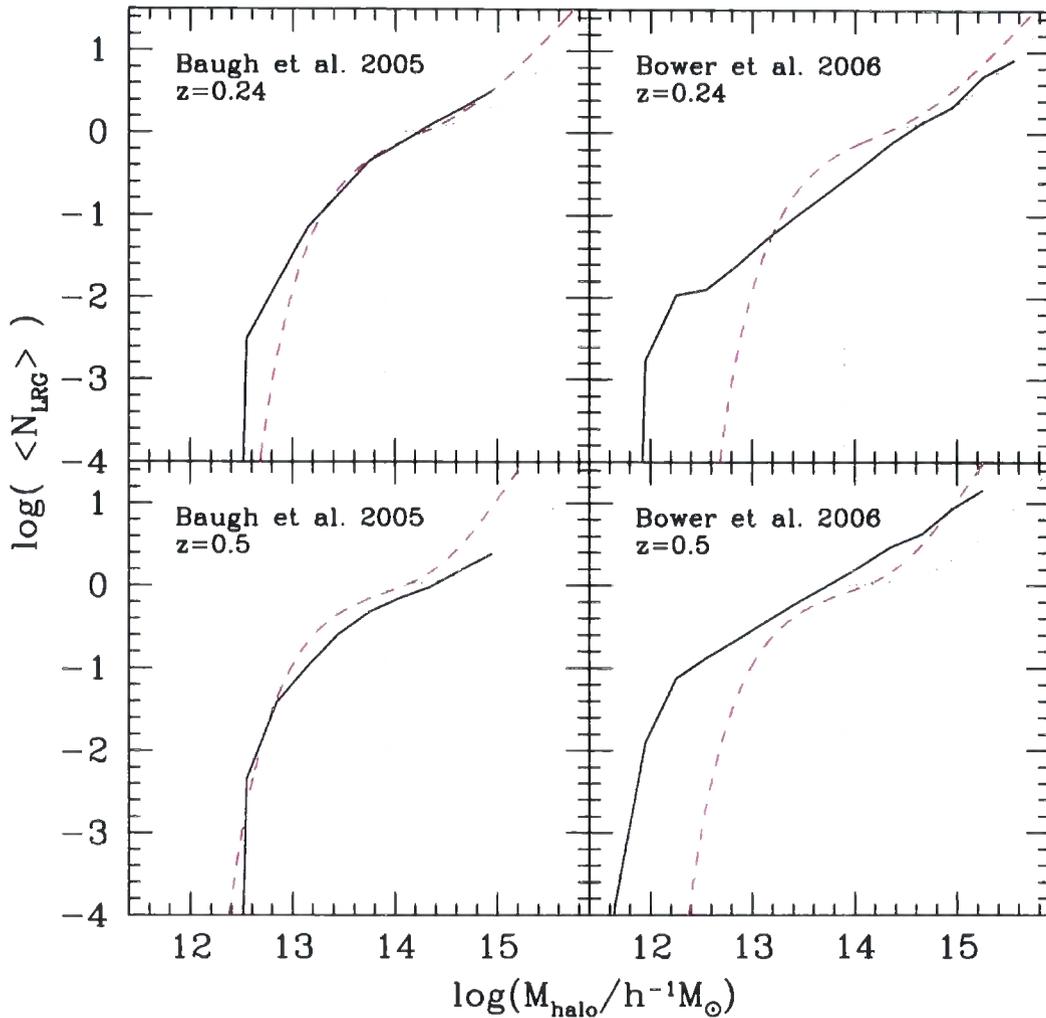
Figure 4.11: The halo occupation distributions (HOD) of LRGs predicted by the models. The left-hand panels show the predictions for the Baugh et al. model and the right-hand panels for the Bower et al. model. The top row corresponds to $z = 0.24$ (SDSS) and the bottom row to $z = 0.5$ (2SLAQ). The dotted line shows the HOD fit used by Masjedi et al. (2006) to describe SDSS LRGs, and is reproduced in each panel. The dashed lines show the HOD fit to the SDSS and 2SLAQ samples obtained by Wake et al. (2008).

carried out such a calculation, fixing the background cosmology to match that used in the two models, and Fig. 4.11 shows that their estimates are in better agreement with the model predictions (Baugh et al. model) than with the HOD advocated by Masjedi et al. (2006) (see also Kulkarni, 2007). Nevertheless, there are still some discrepancies between the fit obtained by Wake et al. and our model predictions. This could be due to the fact that the models do not reproduce exactly the number density of LRGs, whereas Wake et al. include this as a constraint on their HOD parameters.

We plot the frequency of finding a given number of LRGs within a common dark matter halo in Fig. 4.12. Here the number of haloes plotted on the y-axis is the number within the full volume of the Millennium Simulation ($0.125h^{-3}\mathrm{Gpc}^3$), which contain the specified number of LRGs. At $z = 0.5$, nearly ten thousand dark matter haloes in the Millennium contain more than one LRG. Haloes with only one LRG are approximately ten times more common. At the median redshift of the 2SLAQ survey, the tail of haloes with more than one LRG extends to ~ 30, reflecting the higher space density of LRGs in this sample compared with the SDSS LRG sample.

Before presenting explicit predictions for the correlation function of LRGs, it is instructive to compute the asymptotic bias factor, b_{eff}, which quantifies the boost in the clustering of LRGs relative to that of the underlying matter distribution on large scales ($\xi_{\mathrm{LRG}} \approx b_{\mathrm{eff}}^2 \xi_{\mathrm{mass}}$). This will allow us to compare the clustering predictions of the Baugh et al. and Bower et al. models. (We cannot make a direct prediction of the correlation function of galaxies in the Baugh et al. model, because, unlike the Bower et al. model, it is not implanted in an N-body simulation.) The asymptotic bias factor can be calculated analytically using the mass function of haloes which host an LRG, $N(z, M')$, and the bias as a function of halo mass, $b(M', z)$ (e.g. Baugh et al., 1999):

$$b_{\mathrm{eff}}(z) = \frac{\int_M N(z, M')\, b(M', z)\, \mathrm{d}\ln M'}{\int_M N(z, M')\mathrm{d}\ln M'}. \tag{4.13}$$

The integrals are over the range of halo masses which host LRGs. The bias factor, $b(M, z)$, for haloes of mass M, as a function of redshift is computed using the prescription of Sheth, Mo & Tormen (2001). For the Baugh et al. model we calculate an effective bias of $b_{\mathrm{eff}} = 2.45$ at $z = 0.24$, and $b_{\mathrm{eff}} = 2.27$ at $z = 0.50$. In the case
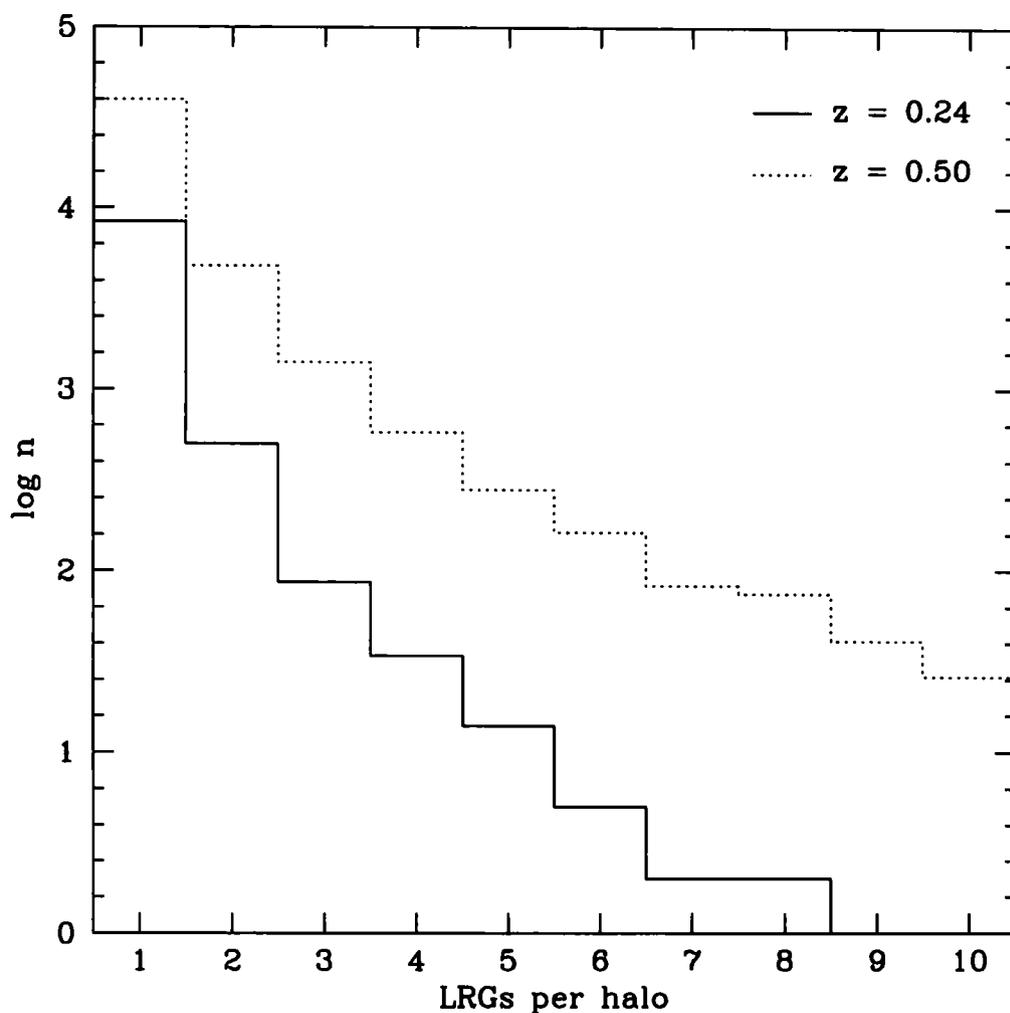
Figure 4.12: T

he number of times a given number of LRGs is found within a common host halo in the Bower et al. model The number of times a given number of LRGs is found within a common host halo in the Bower et al. model, in the whole of the Millennium Simulation volume (note we do not show haloes with zero LRGs). The solid and dotted lines show the distributions at $z = 0.24$ and $z = 0.50$, respectively. Recall that the space density of LRGs selected at $z = 0.5$ is greater than that at $z = 0.24$. Note that the distribution of occupation numbers at $z = 0.50$ extends out to around 30 LRGs per halo.

of the Bower et al. model the values are slightly lower, with $b_{\mathrm{eff}} = 1.82$ at $z = 0.24$ and $b_{\mathrm{eff}} = 1.72$ at $z = 0.50$. Using a sample of 35 000 LRGs from the SDSS, Zehavi et al. (2005) measured a bias of $b = 1.84 \pm 0.11$, which is in agreement with the predictions of the Bower et al. model (see also Kulkarni, 2007; Blake et al., 2007).

For the remainder of this section, we focus on the predictions of the Bower et al. model. As this model is implemented in an N-body simulation, it can be used to produce direct predictions for the spatial distribution of galaxies and hence the two-point correlation function. The published Bower et al. model associates the central (biggest) galaxy in each halo with the largest substructure in the halo. For haloes without resolved substructures, the central galaxy is assigned to the position of the most bound particle. Satellite galaxies are associated with the substructure corresponding to the halo in which they formed or to the most bound particle from the halo in which they formed. Fig. 4.13 shows the predicted correlation function for this model in real-space and in redshift-space. In real-space, the Cartesian coordinates of the LRGs within the simulation box are used to compute pair separations. In the redshift-space, galaxy positions along one of the axes are perturbed by the peculiar velocity of the galaxy, scaled by the appropriate value of the Hubble parameter, H(z). This corresponds to the distant observer approximation, which is reasonable given the median redshifts of the observational samples. Fig. 4.13 shows that the correlation functions predicted by the Bower et al. model agree spectacularly well with the measured correlation functions, both in real and redshift space. The agreement between the model predictions and the observational estimates in real-space is particularly noteworthy. The real-space correlation function of LRGs is very close to a power law over three and a half decades in pair separation, varying in amplitude over this range by nearly eight orders of magnitude. The extension of the power law in the model predictions from $r \sim 1h^{-1}\mathrm{Mpc}$ down to $r \approx 0.01h^{-1}\mathrm{Mpc}$ is a remarkable success of the model. The correlation function on these scales is determined by pairs of LRGs within the same dark matter halo. If the model did not predict that some haloes contain more than one LRG, the correlation function would tend to $\xi \sim -1$ on scales smaller than the typical radius of the haloes hosting LRGs (see Benson et al., 2000, for a discussion of this point). The slope of the
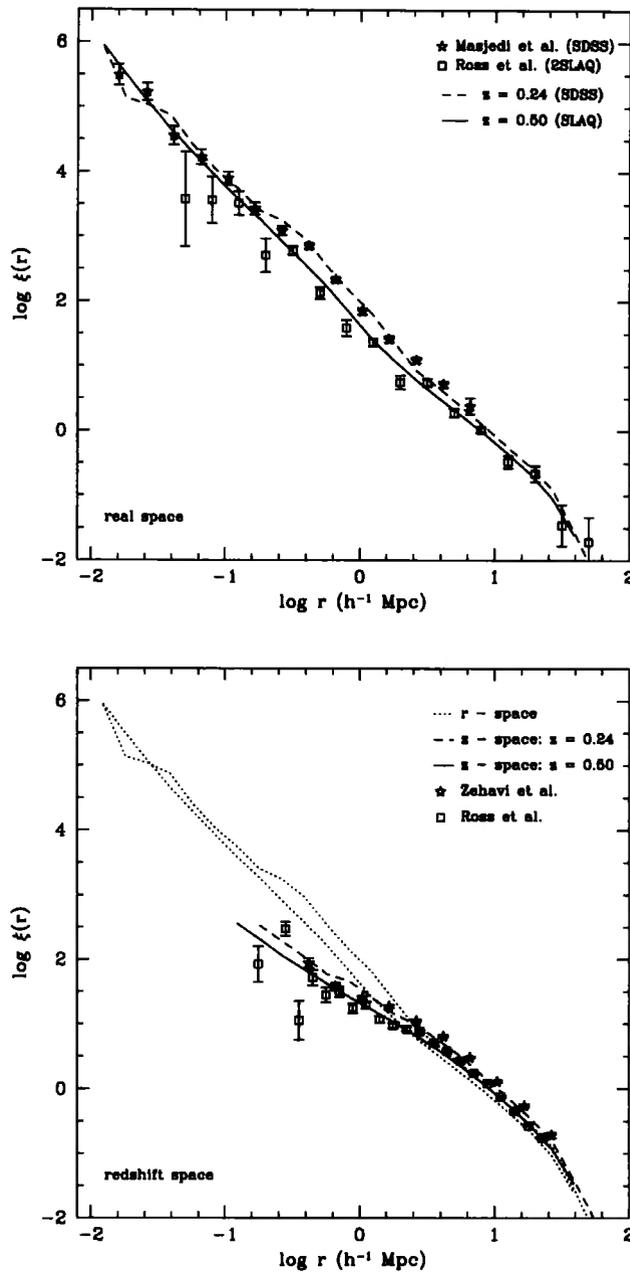
Figure 4.13: The two-point correlation function of LRGs in the Bower et al. model. *Upper panel:* The real-space correlation function at $z = 0.24$ (dashed line) and $z = 0.50$ (solid line). Also shown in this panel are data from Masjedi et al. (2006) (SDSS LRG sample) and from Ross et al. (2007a) (2SLAQ sample). *Lower panel:* The redshift-space correlation function, at $z = 0.24$ (dashed line) and $z = 0.50$ (solid line). For comparison we plot the real-space correlation functions shown in the upper panel using dotted lines. The symbols show the correlation function of LRGs in redshift-space from Zehavi et al. (2005) and Ross et al. (2007a).

correlation function on such small scales is a strong test of the model through the predicted number of LRGs per halo.

In the lower panel of Fig. 4.13, we have retained the same dynamic range on both axes to allow a ready comparison of the clustering signal predicted in redshift-space with that obtained in real-space. To further aid this comparison, we have also reproduced the real-space predictions from the upper panel as dotted lines. The impact of including the contribution of peculiar motions when inferring the distance to galaxies depends on the scale. On intermediate and larger scales($r >$ $3h^{-1}$Mpc), bulk motions of galaxies result in an enhancement in the amplitude of the correlation function measured in redshift-space. This boost is modest because, as we demonstrated above, LRGs are biased tracers of the matter distribution (Kaiser, 1987). On small scales, the clustering signal in redshift-space is significantly lower than in real-space. Again, this feature of the predictions, a damping of the clustering on small scales in redshift-space, is expected if the sample contains haloes which host multiple LRGs; the peculiar motions of the LRGs within the halo cause an apparent stretching of the structure in redshift-space, diluting the number of LRG pairs. The clustering predicted in redshift-space agrees extremely well with the measurements by Zehavi et al. (2005) and Ross et al. (2007a).

Another view of the comparison between the predicted and measured correlation functions is presented in Fig. 4.14, in which we plot the correlation function divided by a reference power law, $\xi(r)_{\rm fit} = (r/r_0)^\gamma$. This way of plotting the results emphasizes any differences between model and data by expanding the useful dynamic range plotted on the y-axis. In both panels, for the reference power law we fit a slope of $\gamma = -2.07$, which agrees with the slope inferred for the real-space correlation function by Masjedi et al. (2006). For the correlation lengths in each panel, we use $r_0 = 8.2\,h^{-1}\,$Mpc at $z = 0.24$ (upper panel) and $r_0 = 7.1\,h^{-1}\,$Mpc at $z = 0.50$ (lower panel). Fig. 4.14 shows clearly the difference between the shape of the correlation function in real-space and redshift space. The level of agreement between the model predictions and the measurements is impressive, particularly in view of the fact that no model parameters were fine-tuned to achieve this match.
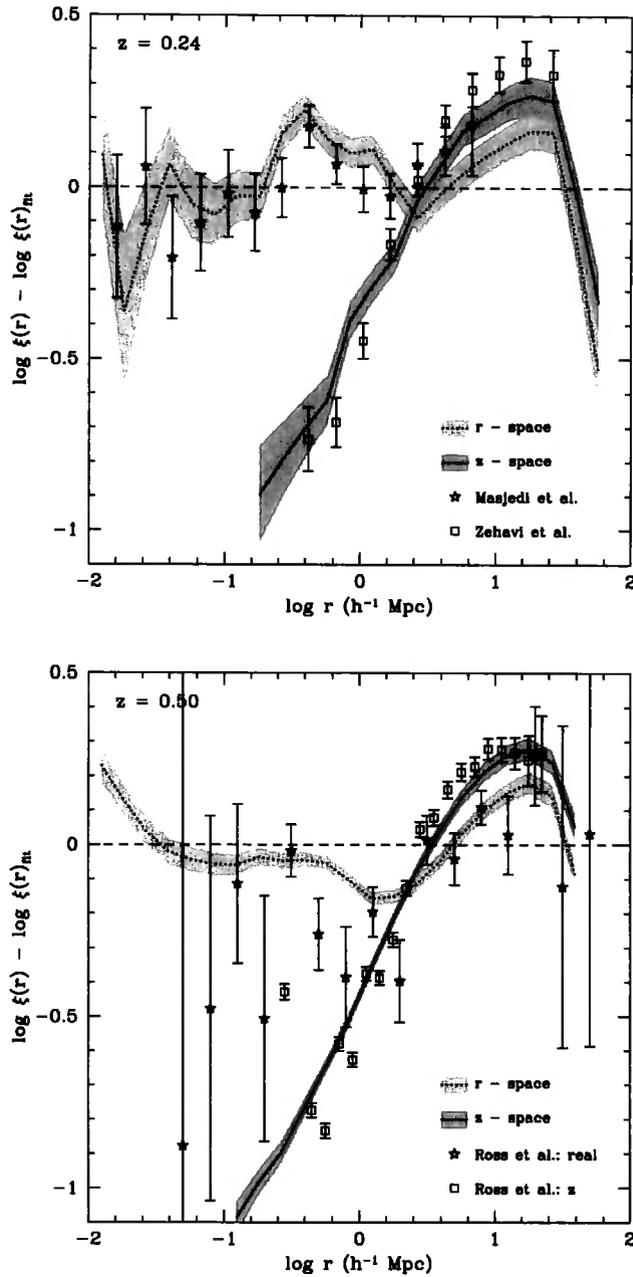
Figure 4.14: The ratio between the two-point correlation function of LRGs and a power law, $\xi(r)_{\text{fit}} = (r/r_0)^{-2.07}$, where $r_0 = 8.2\,h^{-1}\,\text{Mpc}$ at $z = 0.24$ (upper panel) and $r_0 = 7.1\,h^{-1}\,\text{Mpc}$ at $z = 0.50$ (lower panel). The dotted and solid lines show the predictions of the Bower et al. model in real and redshift-space, respectively. The shaded areas show the 1σ Poisson errors derived using the number of pairs expected in the model at a given separation.

4.4 The star formation and merger histories of LRGs

Semi-analytical galaxy formation models trace the full star formation and merger histories of galaxies. This allows us to build up a picture of how the stellar mass of LRGs was assembled and how the LRG population changed between the median redshifts of the 2SLAQ and SDSS surveys. The merging history, in particular, has implications for the clustering expected on small scales, which, as we saw in the previous section, is in excellent agreement with the observational estimates (e.g. Masjedi et al., 2006).

We first consider the star formation and mass assembly histories of LRGs. There are two ways in which a galaxy can acquire stellar mass in hierarchical models: 1) through the formation of new stars and 2) through the accretion of pre-existing stars in galaxy mergers (Baugh et al., 1996; Kauffmann, 1996). A nice discussion of the relative importance of these two processes for brightest cluster galaxies can be found in De Lucia & Blaizot (2007).

In Fig. 4.15 we show some examples of star formation histories for three $z = 0.24$ LRGs in the Bower et al. model. The star formation rate plotted is the sum of the star formation rate over *all* of the progenitor galaxies at each redshift. The star formation history contains contributions from quiescent star formation in galactic discs and from starbursts triggered by galaxy mergers or dynamically unstable discs in the case of the Bower et al. model. In general, the galaxy star formation histories predicted by hierarchical models tend to be more complex than the simple, one parameter, exponentially declining models typically considered in the literature (for some examples of star formation histories generated by semi-analytical models, see Baugh (2006)). Fig. 4.15 confirms that the LRG selection isolates a subset of galaxies in the model with more passive star formation histories, which are closer to exponential models (though these examples still display significant structure in their star formation histories at high redshift). The three examples have similar forms, with a peak at a lookback time of $10 \leq t \leq 12$ Gyrs, followed by a smooth decay. In one of the examples, plotted with the dotted line, there is
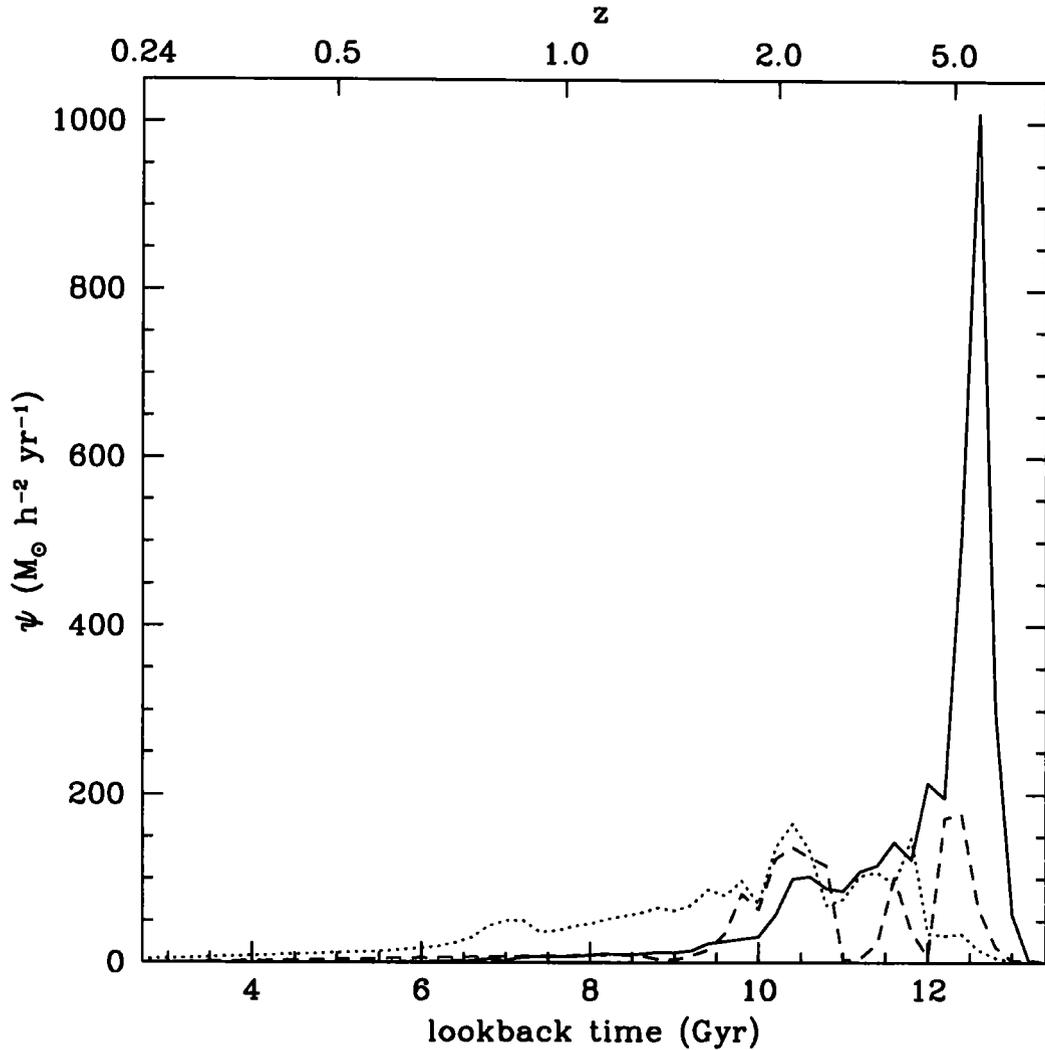
Figure 4.15: Three examples of star formation histories of $z = 0.24$ LRGs extracted from the Bower et al. model, plotted as a function of the lookback time from $z = 0$. The upper axis gives the corresponding redshift. The star formation rate plotted is the sum of the star formation rate in *all* of the progenitor galaxies present at a given redshift.

still appreciable star formation at $z = 0.24$. Around 5% of LRGs in the Bower et al. model display star formation rates at $z = 0.24$ in excess of $0.1\,M_\odot\,h^{-2}\,yr^{-1}$, with the largest being $30\,M_\odot\,h^{-2}\,yr^{-1}$. These low star formation rates indicate that for the bulk of LRGs in the model, ongoing star formation is not an important channel for increasing the stellar mass of LRGs, given the large stellar masses predicted for these galaxies.

The evolution of the stellar mass of LRGs with redshift is shown in Fig. 4.16. In this plot, we take LRGs of similar stellar mass at $z = 0.24$ from the Bower et al model and track the build up of their stellar mass with redshift. The solid line shows how the mean stellar mass of the sample of LRGs builds up over time, expressed as a fraction of the mean mass of the LRGs at $z = 0.24$. Here we sum over all the progenitors of SDSS LRGs. Half of the stellar mass of the $z = 0.24$ LRGs was already in place at $z = 2.2$. The mean change in stellar mass since $z = 0.5$ is around 1%. If, instead, we consider only the most massive progenitor, the figure reveals that half of the mass was already in one object at $z = 1.6$. Since $z = 0.5$, the mean fractional change in stellar mass of the biggest progenitor is just over 0.1; since $z \sim 1$, the average stellar mass has grown by only 25%. This evolution is in agreement with estimates inferred from the observed evolution of the luminosity function for a matched sample of LRGs (i.e. by considering the subset of 2SLAQ LRGs which have similar properties to the SDSS LRGs; Wake et al. (2006)). Fig. 4.16 also shows the contribution to the stellar mass of LRGs from bursts of star formation initiated by galaxy mergers or by the formation of bars in dynamically unstable discs. In total, this channel of star formation is only responsible for around 30% of the stellar mass of SDSS LRGs. Furthermore, the level of contribution from this mode of star formation has changed little since a redshift of $z \sim 1.5$. This reflects the general trend for mergers to become more gas poor (or "dry") with declining redshift in hierarchical models, due to the increasing consumption of gas by quiescent star formation in galactic discs, and the overall decline in the merger rate towards the present day. Given the relatively small star formation rates predicted in model LRGs since $z \sim 1$, the steady increase in the stellar mass of LRGs over the redshift interval $z = 1$ to $z = 0.24$ is driven primarily by gas-poor
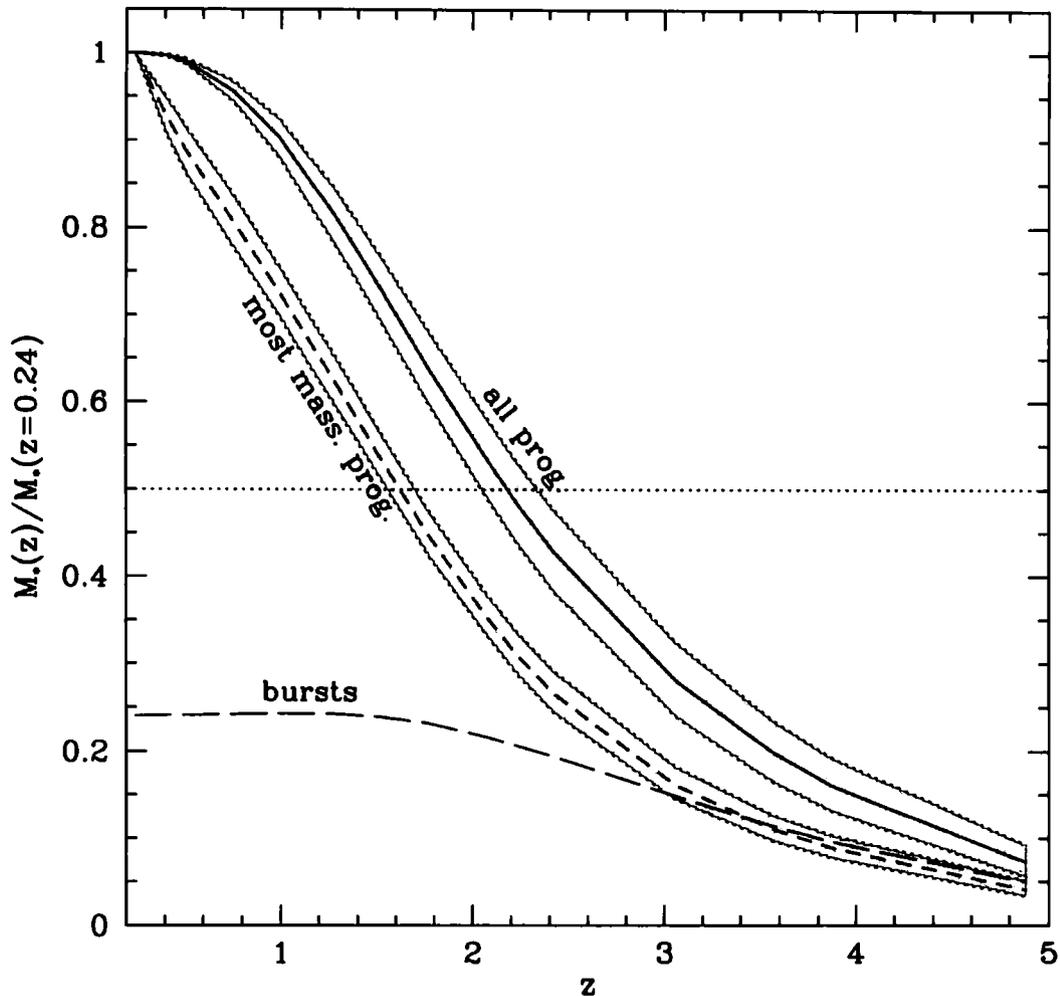
Figure 4.16: The evolution of the total stellar mass of SDSS LRGs, as predicted by the Bower et al. model. The stellar mass in place at a given redshift is expressed as a fraction of the mean stellar mass of SDSS LRGs at $z = 0.24$. The mean fraction of mass in place as a function of redshift, summed over all progenitors and for the most massive progenitor of the SDSS LRGs, is shown by the solid and short-dashed lines respectively. The shaded region indicates the 1σ scatter in this ratio. The long-dashed line shows the fraction of the mean mass accounted for by bursts of star formation, in all progenitors of the $z = 0.24$.

# progenitors > 30% mass	0	1	2	3
(%)	0.3	88.6	11.0	0.1

# progenitors > 30% mass and 2SLAQ LRG				
0	100	27	10	29
1	-	73	55	29
2	-	-	35	29
3	-	-	-	13

Table 4.4: The nature of the progenitors of SDSS LRGs in the Bower et al. (2006) model. The progenitor galaxies are identified at $z = 0.5$ and by definition are required to account for at least 30% of the stellar mass of the LRG at $z = 0.24$, the median redshift of the SDSS sample. The second row gives the percentage of SDSS LRGs with 0, 1, 2 and 3 such progenitors. The rows give the percentage of cases in which a galaxy has the stated number of progenitors which themselves satisfy the LRG selection criteria at $z = 0.5$.

galaxy mergers, which reassemble pre-formed stars. These results differ from that of De Lucia & Blaizot (2007). Using the semi-analytic model described by Croton et al. (2006), the authors found a median formation redshift of the brightest cluster galaxies around $z \sim 5$. As mentioned by Bower et al. (2006), there are several important diferences between these two models, mainly the treatment of feedback from AGN which has a crucial role in shaping the evolution of LRGs.

We now look in more detail at how SDSS LRGs build up their mass in galaxy mergers. The number of progenitor galaxies depends upon whether or not a mass cut is applied before a progenitor is counted. If a mass cut is not used, the number of progenitors obtained is likely to be dominated by low mass galaxies, which only bring in a small fraction of the galaxy's mass. Masjedi et al. (2006) used their measurement of the correlation function on small scales to constrain a simple model for mergers between LRGs, prior to the median redshift of the SDSS sample. Here we do not attempt to consider only those progenitors which are matched to the

SDSS sample selection. Instead, we consider progenitors at $z = 0.5$ which account for 30% of more of the stellar mass of the SDSS LRG at $z = 0.24$. We then further distinguish between progenitors which satisfy the 2SLAQ selection and those that do not.

Table 4.4 shows the number of progenitors of SDSS LRGs predicted by the Bower et al model. The first row gives the percentage of SDSS LRGs which have 0,1,2 or 3 progenitors at $z = 0.5$ which each account for 30% or more of the mass of the $z = 0.24$ LRG. Typically, in the model, SDSS LRGs have one such progenitor at $z = 0.5$. Only 11% of LRGs have more than one progenitor which represents 30% or more of the mass. In the case where a SDSS LRG has only one progenitor with $\geq 30\%$ of the mass, then in 70% of cases this progenitor will be a 2SLAQ LRG. When there are two sizeable progenitors present at $z = 0.5$, then in half of the cases, one galaxy is a 2SLAQ LRG and the other progenitor fails to meet the 2SLAQ LRG definition. In around one third of cases, both progenitors are LRGs.

4.5 Discussion and conclusions

In this chapter we have extended the tests of hierarchical galaxy formation models to include predictions for the properties of a special subset of the galaxy population called luminous red galaxies (LRGs). Given their rarity, bright luminosities and extreme colours, LRGs represent a stern challenge for the models. They are particularly interesting from the point of view of developing the model physics, since the abundance and nature of LRGs probes precisely the regime in which the models are currently most uncertain, the formation of massive galaxies. Historically, hierarchical models have tended to overproduce bright galaxies at the present day (see Baugh, 2006). The phenomena invoked to restrict the growth of large galaxies locally, naturally, have an impact on the form of the bright end of the galaxy luminosity function at intermediate and high redshifts, where LRGs dominate. In addition, LRGs have the red colours expected of a passively evolving stellar population, which restricts the range of possible star formation histories for these galaxies.

It may appear odd to talk about producing model predictions after an observa-

tional dataset has been constructed. The models considered in this Chapter contain parameters whose values were fixed by requiring them to reproduce a subset of the data available for the local galaxy population (for a discussion see Cole et al., 2000). None of the datasets used for this purpose make explicit reference to the redshifts of interest for the LRG surveys discussed here, nor were red galaxies singled out for special attention in the process of setting the model parameters. We do, how-ever, require that our semi-analytical models reproduce as closely as possible the bright end of the present day luminosity function of all galaxies, which does tend to be dominated by red galaxies with passive stellar populations (e.g. Norberg et al., 2002). Therefore, by comparing the models to the observed properties of the LRG population, we are in effect testing the physics which govern the evolution of the bright end of the luminosity function, as traced by objects with the special colours of LRGs.

The two models considered, Baugh et al. (2005) and Bower et al. (2006), enjoy a considerable number of successes and, inevitably, have some shortcomings. It is important to be clear that we have not adjusted or tinkered with *any* of the parameters of the published models in order to improve the comparison of the model output with the observational data. This "warts and all" exercise illustrates the appeal of the semi-analytical approach, in that a given model yields a broad range of outputs which are directly testable against observations. In both models, LRGs are predominantly bulge dominated galaxies (although 20-40% are expected to be spirals with old stellar populations), with velocity dispersions of $\sigma \sim 320 \mathrm{kms}^{-1}$ and stellar masses around $1 - 2 \times 10^{11} h^{-1} M_{\odot}$, which is higher than observed. The models give different predictions for the radii of LRGs, with the Baugh et al model predicting the larger LRGs. Both models fail to produce bright spheroids that are large enough to match the locally observed radius-luminosity relation.

The Baugh et al. and Bower et al. models are two feasible simulations of the galaxy formation process, which differ in several ways, as we saw in Chapter 2. A key difference between the models, in terms of the analysis presented in this chapter, is the form of the physics invoked to quench the formation of massive galaxies. In both models the amount of gas cooling from the hot halo, to provide the raw material

for star formation, is reduced by quite different means. Baugh et al. invoke a wind which expels cold gas from intermediate mass haloes. This gas is assumed to be ejected with such vigour that it does not get recaptured by more massive haloes in the merger hierarchy. Hence, in this model, the more massive haloes contain fewer baryons than expected from the universal baryon fraction, and therefore less gas is available to cool from the hot halo. One controversial aspect of this scheme is the energy source required to drive the wind. Benson et al. (2003) showed that the energy produced by supernova explosions is unlikely to be sufficient to power a wind of the strength required to reproduce the sharpness of the break in the local galaxy luminosity function, and argued that the accretion of gas onto a central supermassive black hole could be the solution. Bower et al. invoked an AGN feedback model in which the luminosity of the AGN heats the hot halo (see also Croton et al., 2006; Granato et al., 2004, for an alternative model). This suppresses the cooling flow in massive haloes which have quasi-static hot gas atmospheres.

The predictions of the Baugh et al. and Bower et al. models bracket the observed luminosity function of LRGs, with the Bower et al. model giving the better overall agreement with the SDSS and 2SLAQ results. The shape and normalization of the $z = 0.24$ LRG luminosity function predicted by the Bower et al. model are in excellent agreement with the observations. This is remarkable when one bears in mind that LRGs are an order of magnitude less common than L_* galaxies. The Baugh et al. model on the other hand, whilst predicting a similar number density of LRGs, gives a poor match to the shape of the observed luminosity function. At $z = 0.5$, the agreement is not so good, with the predictions only coming within a factor of 2 of the observed abundance. This implies that the models may not be tracking the evolution of the bright end of the luminosity function accurately over such a large lookback time (40% of the age of the universe), at least for galaxies matching the 2SLAQ selection. Whilst this discrepancy suggests that there are problems modelling the evolution of the red galaxy luminosity function, it is important to note that the Bower et al. model does give a good match to the inferred evolution of the stellar mass function, to much higher redshifts than that of the 2SLAQ sample. We investigated whether it was possible to tune the predictions of the Baugh et al.

model to better match the LRG luminosity function; this exercise proved to be unsuccessful suggesting that a more substantial revision to the ingredients of this model, involving further suppression of gas cooling in massive haloes, is required.

Semi-analytical models predict the star formation histories of galaxies, based upon the mass of cold gas which accumulates through cooling and galaxy mergers, and a prescription for computing an instantaneous star formation timescale (examples of star formation histories extracted from the models are given in Baugh, 2006). As expected, the stellar populations of model LRGs are old, with luminosity weighted ages in the region of 4-8 Gyr for the SDSS selection, with the Bower et al. model returning the more elderly stars (similar results were reported by De Lucia et al., 2006; Croton et al., 2006, for massive elliptical galaxies). The semi-analytical model can track the build-up of the stellar mass of LRGs, considering all of the progenitor galaxies. There is little recent star formation in any of the progenitor galaxies of SDSS LRGs; averaging over all progenitors, typically 50% of the $z = 0.24$ stellar mass of the LRG has already formed by a redshift of $z \sim 2.2$. However the mass of the main progenitor branch is still growing over this redshift interval. Around half of the mass in the biggest progenitor is put in place since $z \sim 1.5$ through galaxy mergers of ready-made stellar fragments (for a discussion of the difference between the formation time of the stars and the assembly time of the stellar mass, see De Lucia & Blaizot, 2007). On average, only 25% of the stellar mass of the LRG is added after $z \sim 1$, in line with observational estimates of the evolution of the stellar mass function, which indicate that many of the most massive galaxies are already in place by $z \sim 1$ (e.g. Bauer et al., 2005; Bundy, Ellis & Conselice, 2005; Wake et al., 2006).

Perhaps the most spectacularly successful model prediction is for the clustering of LRGs. Masjedi et al. (2006) estimated the two-point correlation function of SDSS LRGs in real space, free from the distortions in the clustering pattern induced by the peculiar motions of galaxies. These authors found that the real-space correlation function of LRGs is a power law over three and a half decades in pair separation, down to scales of $r \approx 0.01h^{-1}\text{Mpc}$. Masjedi et al. argued that current halo occupation distribution models could not reproduce such a steep correlation function on

small scales because these models asume that galaxies trace the density profile of the dark matter halo, which is shallower than the observed correlation function. This line of reasoning is spurious, as HODs can produce realizations of the two-point correlation function with different small-scale slopes for different galaxy samples, even when the different samples trace the dark matter (see, for example, Fig. 22 of Berlind et al., 2003, which compares the correlation functions of old and young galaxies). The small scale slope depends on the interplay between two factors: the number of galaxies within a dark matter halo and the range of halo masses which contain more than one galaxy (e.g. Benson et al., 2000). The Bower et al. model can readily produce predictions of galaxy clustering down to such small scales since it is embedded in the Millennium Simulation (Springel et al., 2005). The correlation function predicted by the Bower et al. model agrees impressively well with the observational estimate by Masjedi et al. The HOD used by Masjedi et al. is actually a poor description of the HOD predicted in the Bower et al. model. Further support for the number of LRGs predicted as a function of halo mass comes from the degree of damping of the correlation function seen on small scales in redshift-space. The virialized motions of LRGs within a common halo gives a contribution to the peculiar velocity of these galaxies, which results in the structure appearing stretched when the distance to the LRG is inferred from its redshift. This damping would not be apparent in the case of a maximum of one LRG per halo.

Overall, the agreement between the model predictions and the observation of LRGs is encouraging, demonstrating the true predictive power of semi-analytical models. The two models we have tested have quite different mechanisms to regulate the formation of massive galaxies, with the Bower et al. model invoking "AGN feedback" and the Baugh et al. model relying on a "superwind"; in the former, the raw material for star formation is prevented from cooling in the first place in massive haloes, whilst in the latter cold gas is expelled from the halo before it can form stars. The Bower et al. model does the best in terms of matching the abundance of LRGs, particularly at $z = 0.24$.

Whilst problems remain in predicting the radii of spheroids and the precise evolution of LRG luminosity function, it is clear that these objects can be accommodated

in hierarchical models.

Chapter 5

Modelling the Dusty Universe I: The Method

5.1 Introduction

A deeper understanding of the processes that determine the formation and evolution of structure in the Universe and increasing computing power have contributed to recent advances in modelling of galaxy formation. The Millennium Simulation is the largest N-body simulation of the hierarchical clustering of dark matter in a ΛCDM universe published to date (Springel et al., 2005). It uses more than 10 billion particles to represent the dark matter, which interact gravitationally in a volume of $0.125\,h^{-3}\text{Gpc}^3$. The potential of the simulation is endless. For example, we are now able to recreate the merger and star formation histories and evolution of millions of simulated galaxies, using semi-analytic models of galaxy formation, along with their spatial distribution. The comparison with observational data is of great potential for improving our understanding of the physical processes involved in the formation and evolution of galaxies.

As introduced in Chapter 2, GALFORM is a semi-analytic model of galaxy formation which has proven to be very successful in reproducing galaxy properties and their evolution. In Granato et al. (2000) the authors presented a new class of semi-analytic models, where they included, for the first time, a detailed treatment of dust emission within a semi-analytic model of galaxy formation (see Chapter 2): a merger between the GALFORM and GRASIL codes. This union extended GALFORM such that

the new model was now able to predict accurately the photometric properties of the galaxies, ranging from the UV to the submillimeter.

In this Chapter we will use this extended version of the GALFORM to predict spectra for all the Millennium Simulation galaxies, to build mock catalogues at wavelengths where the predictions of the standard GALFORM code are inaccurate (e.g. UV or sub-mm), and the GRASIL code needs to be used. This might seem infeasible computationally for a large cosmological volume, taking into account that at each output Millennium snapshot (output redshift), on average more than 20 million galaxies populate the sample, and that GRASIL takes a few minutes to compute the spectrum of a single galaxy, which multiplied by millions of galaxies will easily become thousands of years of computation time – a mission nearly impossible.

In order to achieve our goal we need a new, much faster approach. Our starting point is to note that GALFORM is quick at predicting galaxy properties. The next step is to find a way of relating a galaxy's properties to its spectrum, a correlation between properties and the spectral energy distribution (SED), so that we can avoid the time consuming computation of the SED using GRASIL. To do so, we are going to use *Artificial Neural Networks*, data analysis tools designed to behave like the biological brain, which are excellent at finding hidden patterns in data.

In Section 5.2, we summarize the theory behind the artificial neural networks. Several aspects related to the performance of neural networks are discussed in Section 5.3. Here, we also present the first results: predictions of luminosities in the B, 24 μm and 850μm bands. In Section 5.4, we extend the tests of this new technique, by studying the predicted luminosity functions of Lyman-break ($z = 3$), IR ($z = 0.4$) and submillimeter ($z = 2$) galaxies. Our conclusions are given in Section 5.5.

5.2   Artificial Neural Networks

Artificial neural networks (ANNs) are a powerful data analysis tool. They are mathematical/computational instruments intended to replicate the behaviour of the human brain, i.e. to learn from observations and data. The origins of this technique date back to McCulloch & Pitts (1943), who developed a simple network using *ar-*

tificial neurons to perform logical operations. However, the full concept of learning was only introduced a few years afterwards by Hebb (1949), and implemented by Rosenblatt (1958, 1962). Nowadays, ANNs are widely used in computer science, finance, physics, mathematics, astronomy, and many other areas. Typically, their applications are pattern recognition, prediction and forecasting, function approximation, and categorization. Even though neural networks have traditionally been seen as black boxes, where the user has little information about the internal parameters/coefficients of the network, they offer several advantages over other data mining and analysis tools: the ability to learn, generalization, adaptability and computational parallelism.

5.2.1 Basic Concepts

Since ANNs are inspired by the biological brain, the obvious question to ask is "How does the brain work?" Despite the fact that we do not know the answer in detail, we do know that the brain basically consists of billions of special cells, *neurons*, that process information and are interconnected (through synapses) in a complex net. The important part behind the learning process lies in the way the neurons work: they receive electrochemical signals from other neurons, some of which will excite the cell while others will inhibit it. The neuron adds up all these contributions and if the sum is bigger than a certain threshold, it will transmit the same kind of signal to other neurons (i.e. the neuron is activated).

Artificial neural networks are not much different from their biological counterparts: they consist of simple computational units (also called neurons or nodes), that are connected and organized as a network. In Fig. 5.1 we show a schematic representation of an artificial neuron. For every neuron, $(i_1, i_2, \cdots)$, we need to specify *input* connections and their associated *weight*, w_j. Once these have been defined, the neuron multiplies the input by its weight and adds all the contributions from the interconnected units. The sum is then mapped by the *activation function*, f, to the *output* value, which in turn will become an input to its adjacent neurons.

If we define i_k as the input coming from the neuron k, and w_{jk} as the weight between the input k and neuron j, then, according to the definition given in the
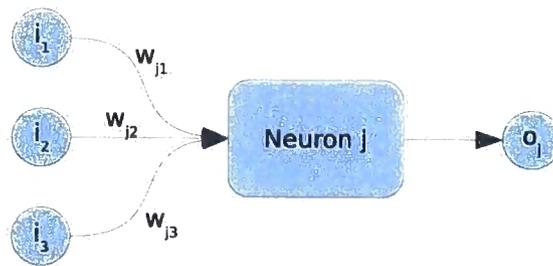
Figure 5.1: Schematic representation of an artificial neuron. i_1, i_2, i_3 represent the input neurons and w_{j1}, w_{j2}, w_{j3} the weights associated with the connections between the input neurons and the neuron j in the next layer. The output neuron is shown by o_j.

paragraph above, we calculate the output o_j from the neuron by:

$$o_j = f\left(\sum_{k=1}^{n} w_{jk}\, i_k\right). \tag{5.1}$$

There are numerous types of ANNs, which differ by the way the neurons are organized and exchange information. In terms of organization, it is common (although not necessary) to group the neurons into *layers*; this help us to better understand how the ANN works. Generally, there is an *input* layer, an *output* layer and *hidden* layers. The input layer is responsible for handling the input data. It is clear that there is no activation function associated with this layer, because the output values of their neurons are simply set to be equal to their input values. Evidently, the output of the network is recovered from the output layer. Using only one input layer and one output layer, it is possible to construct a very simple network, usually called a *perceptron*, which can recognize simple patterns in data (linearly separable data). For more difficult tasks we need the help of hidden layers between the input and output layers. The term "hidden" is used due to the fact that the user does not have direct access to the information dealt by them; this is the reason why neural networks are seen as black boxes.

Networks with more layers than just an input and output layers are called multilayer networks, or multilayer perceptrons (MLPs). They are the most widely used due to their ability to learn nonlinear functions.

In Fig. 5.2 we show a schematic view of three popular network configurations:

the perceptron (no hidden layers), the feed-forward and recurrent network (the two latter cases both use hidden layers). The feed-forward nets are the most widely
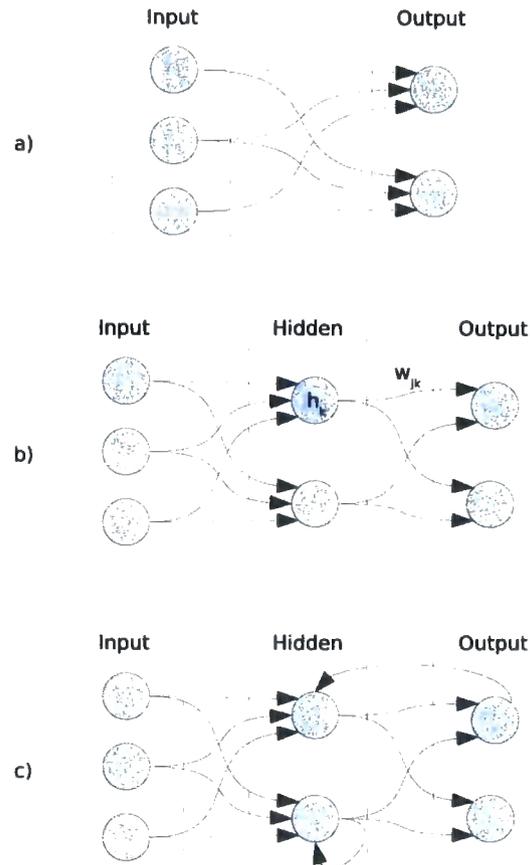
Figure 5.2: Different artificial neural networks configurations: (a) Simple perceptron; (b) Feed-forward network; (c) recurrent network.

used mainly due to their simplicity: they pass the information from the input layer, through the hidden layers, to the output neurons. In recurrent networks, on the other hand, the output from its neurons can be fed backwards, through feedback connections, and act as input; this behaviour is similar to that found in the biological brain. Even though recurrent networks can perform better than the feed-forward nets, they suffer from a major drawback: the training is more difficult due to their oscillatory/chaotic behaviour, resulting in longer computing times.

5.2.2 Learning

Training an artificial neural network is a crucial step: without training there is no learning. The core question of ANN theory is: "How does the neural network learn?" Basically, there are two processes: *supervised learning* and *unsupervised learning*. In supervised learning, the network is presented with a *target*, i.e., we teach the ANN that for a given input we expect a given answer. The net will have to adapt its weights in order to produce the desired output. In unsupervised learning, as the name suggests, the network does not have a target output. It is used to find patterns and group the data, and is generally applied to clustering analysis.

We will focus mainly on supervised learning, whilst bearing in mind that most of the procedures that we discuss here are also applicable to unsupervised nets.

There are several different methods within supervised learning. Most of them share a common base: they learn by comparing the predicted output from the ANN to the target output. The algorithm of the whole process is simple: (i) start with an untrained net; (ii) determine the output from a given input; (iii) compare the output to the target value and compute an error; (iv) adjust the weights in order to reduce the error.

The most widely used learning algorithm is the *backpropagation algorithm*, which was introduced by Rumelhart et al. (1986), and is based on the gradient descent method to find the local minimum of the error function. The concept is simple: we use the *gradient descent* method, or more correctly, a generalization of the *delta rule*, to find the best way of updating the weights in order to minimize the error function, which is given by:

$$\varepsilon = \frac{1}{2} \sum_k (t_k - o_k)^2, \tag{5.2}$$

where t_k represents the desired or target values, and o_k the predicted output from the neuron.

It can be shown that the update to the weights from the hidden layer to the output layer is given by:

$$\Delta w_{jk} = -\eta \left(\frac{\partial \varepsilon}{\partial w_{jk}} \right) = \eta \, \delta_j h_k, \tag{5.3}$$

with η being the *learning rate*, $\delta_j = (t_j - o_j) f'(i_j)$ (with the activation function, f differentiable) and h_k is the output from the preceding hidden neuron (see Fig. 5.2 (b)). A similar expression can be found for the variation of the weights between the input and hidden layers.

It is clear that if the error surface has multiple local minima then this method, as it was originally defined, will guarantee convergence towards one of these minima, but not necessarily to the global minimum. However, this is not an insurmountable problem because, during the first steps of the gradient descent, the weights will gradually move towards the global minimum and, in the worse case scenario, converge to a local minimum in its vicinity. There are several methods to avoid this behaviour: we can add an extra factor to Eq. 5.3, $\beta \Delta w_{jk}^0$, called *momentum*, which has the same direction as the previous step change, Δw_{jk}^0, and is controlled by the coefficient β; or we can train the network various times using the same training sample but with different initial random weights. Another option is the RPROP (resilient backpropagation) algorithm (Riedmiller & Braun, 1993), in which instead of adopting the full Eq. 5.3, we only use the sign of the derivative, multiplied by a constant; this is the method that we will use in our network. RPROP has the advantage of being one of the fastest weight update learning algorithms.

Of course, in an ideal situation the ANN will find the optimal set of weights such that the error function is minimal. However, there is an important aspect that we need to keep in mind: one of the reasons why we use artificial neural networks is that we need generalization; i.e. it is more important to find the network that best fits the testing or validation set, than finding the minimum of the error function for the training set. In fact, if the net is overtrained it might fit the noise associated with the data, instead of the underlying signal, inducing overfitting, and consequently may underperform on the validation sample. One of the procedures to avoid overfitting is to do an early stop, i.e. stop the training process, either when the error function reaches a pre-defined error threshold or by setting a maximum number of iterations (epochs).

Finally, just a brief word about the activation functions. As we saw, the activation function has an important role in the neural network: it allow us to activate

or deactivate neurons and add the nonlinearity needed to solve some problems (by adding a nonlinear activation function); evidently, activation functions only make sense for the hidden and output layers. There are many activation functions, in fact, any nonlinear function would fit the bill. The most common are: the Sigmoid function, $f(x) = \frac{1}{1+e^{-\alpha x}}$, where α is the *steepness*; Gaussian, $f(x) = e^{-(\alpha x)^2}$; Elliot, $f(x) = \frac{\alpha x}{1+|\alpha x|}$ (see Fig. 5.3).
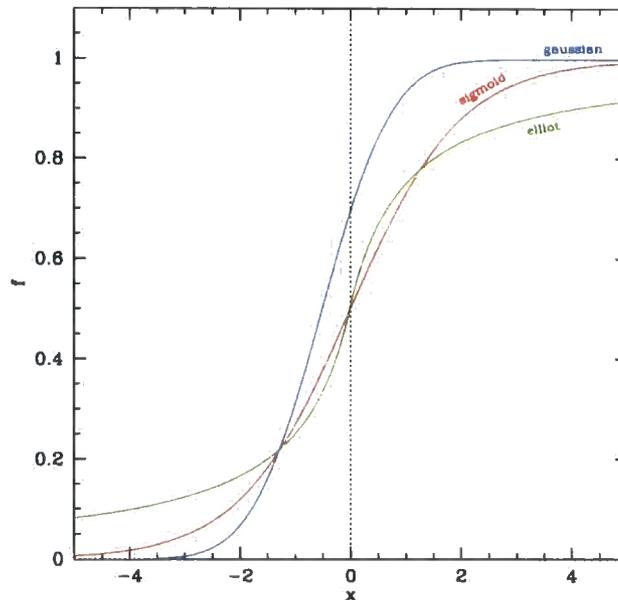
Figure 5.3: Example of the most commonly used activation functions: Sigmoid (red), Elliot (green) and Gaussian (blue).

5.3 First Results and Performance

As mentioned previously, our objective is to predict a galaxy's spectrum given a small set of galaxy properties, using artificial neural networks. The job is far from simple, due to the complexity, dimensionality and variance in the spectra of a population of galaxies. In this section we will explain how we implemented the ANN in order to predict spectra (or luminosities), showing the first results and analysing some performance issues. Throughout this Chapter we will use the variant of the GALFORM model described by Baugh et al. (2005).

5.3.1 Training and Testing Samples

The training process is crucial for artificial neural networks, because the whole process of learning depends on it: the better the network is acquainted with the data, the better it will perform when predicting the output.

The galaxy spectra calculated by GRASIL are far from being a simple function. As noted in Chapter 2, GRASIL has the ability to calculate stellar emission and dust extinction and reemission (both for dense molecular clouds and diffuse cirrus), from given star formation and metal enrichment histories, gas mass and some structural galaxy properties. As a result, GRASIL spectra are quite complex, in terms of shape and normalization, similar to what is observed for real galaxies. In Fig. 5.4 we show some examples of spectra, the very quantity that we are trying to predict. In the top panel, we plot a randomly selected total spectrum (black line), and its underlying components (for this particular galaxy, its interesting to see that the mid-infrared emission is dominated by the PAH molecular bands and the far-infrared by cirrus (diffuse dust) emission); more examples of galaxy spectra are shown in the bottom panel.

Fig. 5.5 shows, in more detail, the complexity of the spectra output by GRASIL. Here we plot the ratio between the standard deviation and the mean of the spectra, for the whole range of wavelengths, for a representative sample of galaxies. The figure reveals that the ultraviolet, infrared (near and mid-infrared) and microwave regions of the spectrum are the main contributors to the total spectral variance; the visible and far-infrared parts of the spectrum do not show much variation in the model spectra.

Each spectrum is composed of 456 flux bins, so as our first approach (Subsection 5.3.2) we will set the number of output neurons in our net to be 456, one for each flux bin. Later on we will try different methods in order to reduce the dimensionality and variance of the output space. We will assign the logarithm of the flux to each of these outputs, in an effort to reduce the dynamical range of the training data.

The selection of the input for the ANN is not so straight forward. The natural choice would be to adopt the same input as used by GRASIL to create the spectra,
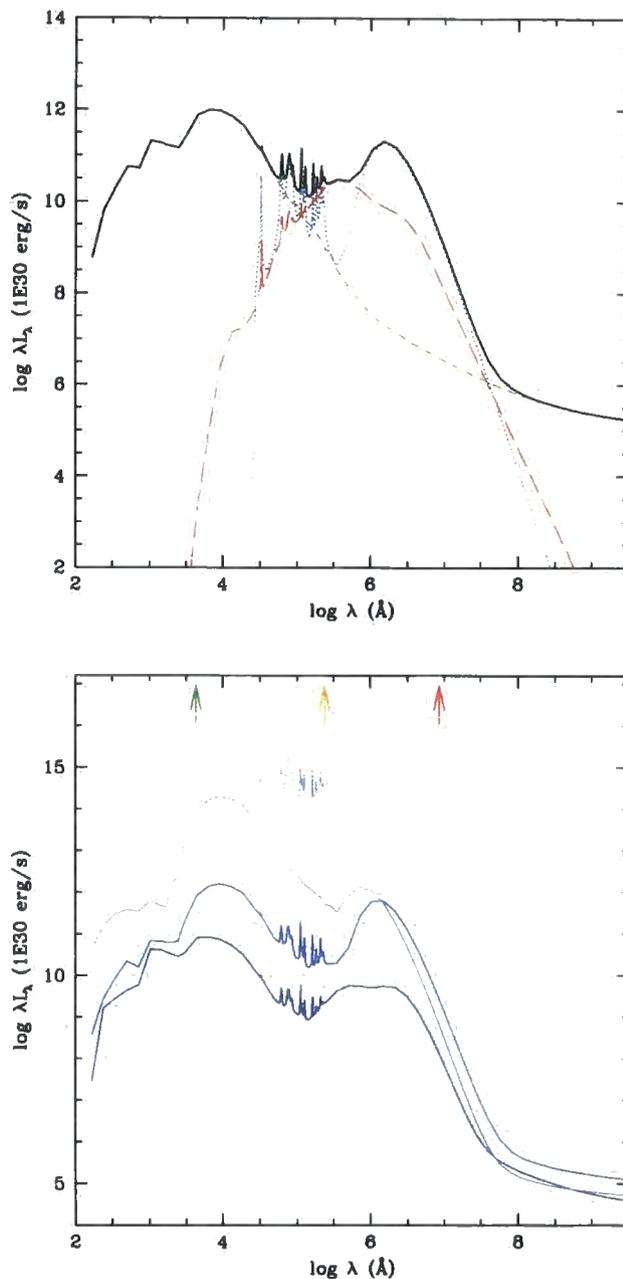
Figure 5.4: Example galaxy SEDs, as given by GRASIL. In the top panel, we show the different components of the spectrum: the black line shows the total SED, which is the result of adding the stellar light extincted (green short-dashed line), cirrus (dotted blue line) and molecular clouds (red long-dashed line) emission. Several examples of spectra are shown in the bottom panel. Here, the green, orange and red arrows show the position of the B, 24 μm and 850 μm filters, respectively.
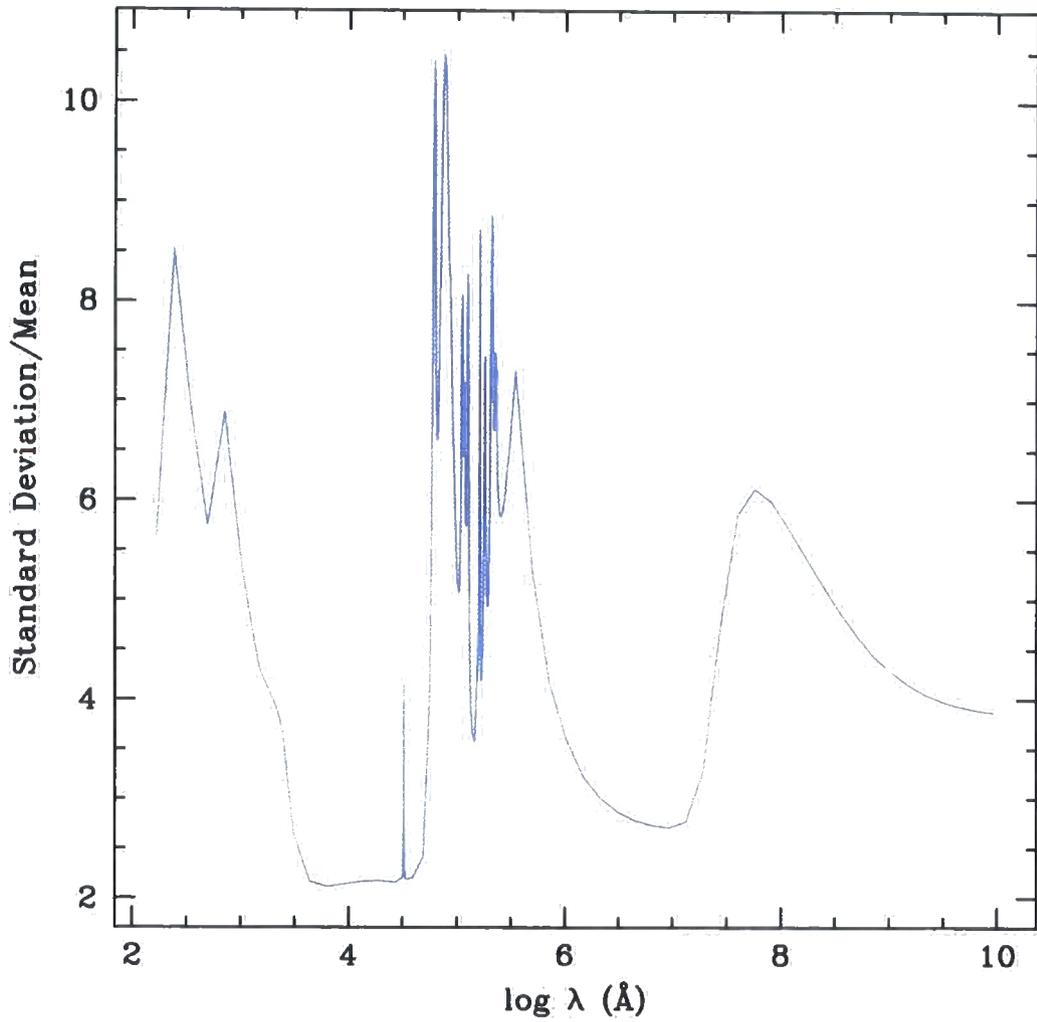
Figure 5.5: The ratio between the standard deviation and mean of model GRASIL spectra, for all the available wavelength range, for a representative sample of galaxies.

i.e. gas mass and the star formation and metal enrichment histories. However, this is hard to implement due to the enormous number of input variables implied (more than 3000), which, in turn, would represent a substantial amount of computing time and trouble for the learning process. To keep things simple, we decided to use a small set of galaxy properties, defined at the galaxy's output redshift. The properties used are: total stellar mass, metallicity of combined disc+bulge stars, bolometric luminosity, circular velocity of the disc and bulge, disc and bulge half-mass radius, V-band luminosity weighted age, V-band extinction optical depth, metallicity of the cold gas, luminosity of dust clouds, disc star formation rate, mass of stars formed in the last burst and time since last major star formation burst (N.B. bursts are triggered by galaxy mergers; see Chapter 2). In total, we construct an input layer with 14 galaxy properties.

Our training and testing samples were extracted from a large catalogue of GAL-FORM model galaxies at $z = 0$, following a similar procedure to Granato et al. (2000). We sample the GALFORM catalogue such that we have equal numbers of galaxies in logarithmic bins of total stellar mass. In a later subsection, we will use also burst galaxies, where the subsampling was done using the logarithmic bins of burst mass instead. This strategy leads to 1945 galaxy spectra for the training set, each of these composed of 456 flux bins, i.e. a 1945×456 data array. 1898 galaxies were used as a testing sample.

5.3.2 Predicting Spectra

To predict luminosities, we will use a supervised feed-forward neural network composed of: 14 neurons in the input layer (which correspond to the 14 selected galaxy properties), 50 neurons in one hidden layer and 456 neurons in the output layer (which are set to be equal to the flux in each of the spectrum bins); the net architecture is therefore 14:50:456.

Unless otherwise specified, the following procedures and parameters were chosen: (i) in order to deal with the different ranges of the input and output parameters, we subtract the mean from each input and output and divide by the standard deviation; (ii) we select a Sigmoid activation function and set the number maximum of training

epochs to 5000 (i.e. the criteria used for stopping the training process); (iii) also, in order to guarantee convergence towards the global minimum of the error function (see previous section), we train the ANN ten times using different initial random weights, and select the one that gives the smallest relative root mean square error (see definition below) for the test sample. Later on in this section, we will analyse how the results change on modifying the ANN parameters.

In Fig. 5.6 we plot four randomly selected examples of predicted spectra. The figure shows that, even without further optimization, the spectra predicted with the ANN agree reasonably well with the original values, particularly in the visible and near-infrared bands. In a small number of cases, however, some galaxies exhibit predicted luminosities which can disagree by a factor bigger than 100 from their true values (see for example the far-infrared region for the galaxy shown in the bottom-right panel).

Fig. 5.7 shows the percentage error of the predicted spectra (as given by 1-predicted luminosity/original luminosity) against the original luminosity, for the B, MIPS 24 μm and SCUBA 850 μm bands; the statistics of this distribution are also summarized in Tab. 5.1. Here, the relative root mean squared error is given by $rrmse = \sqrt{(\sum(L_{\text{predicted}} - L_{\text{original}})^2/L^2_{\text{original}})/n}$ and $P_{|e|<10\%}$ represents the percentage of galaxies with predicted luminosities within 10% of the true values. The figure shows that there is reasonable agreement between the predicted and original luminosities for the analysed bands: most of the galaxies have errors within 20%. Furthermore, we find that there is an increase in scatter for those wavelengths where the contribution of the dust component is dominant, with a $rrmse$ going from 0.12 in the B-band to 1.61 in the SCUBA-850 μm band. Presumably the relative root mean squared error will also have a high value in the ultraviolet bands, as suggested by Figs. 5.6 and 5.5. Note, however, that $rrmse$ exaggerates the cases in which the prediction was significantly greater than the true value. Also, the error has a slight tendency to increase with luminosity. This reveals that the ANN has more difficulty in dealing with bright galaxies, which might be due to the complexity of the star formation histories of these galaxies, which in turn will induce an increase in the amount of information/spectral features modelled by GRASIL.
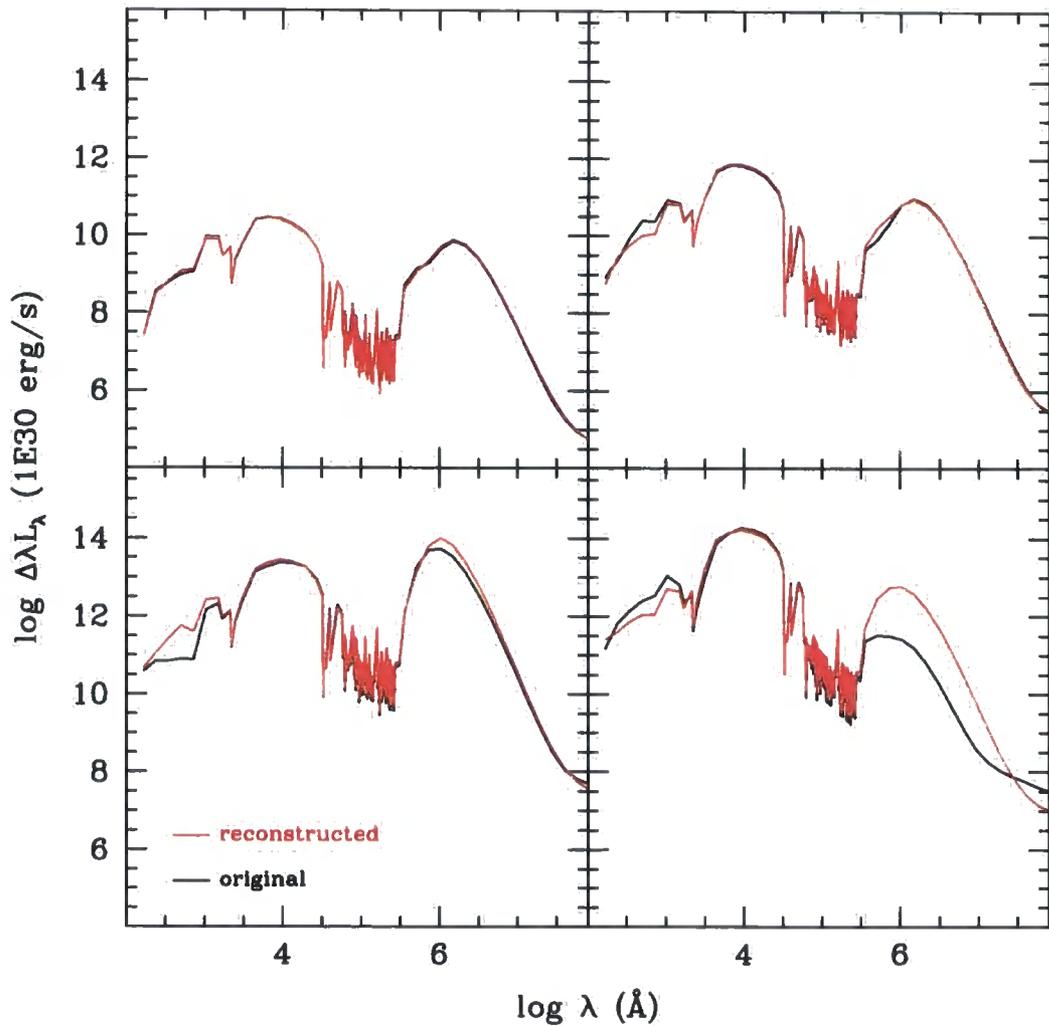
Figure 5.6: Four randomly selected galaxy SEDs predicted by the ANN method. The red and black curves represent the predicted and original spectra respectively.
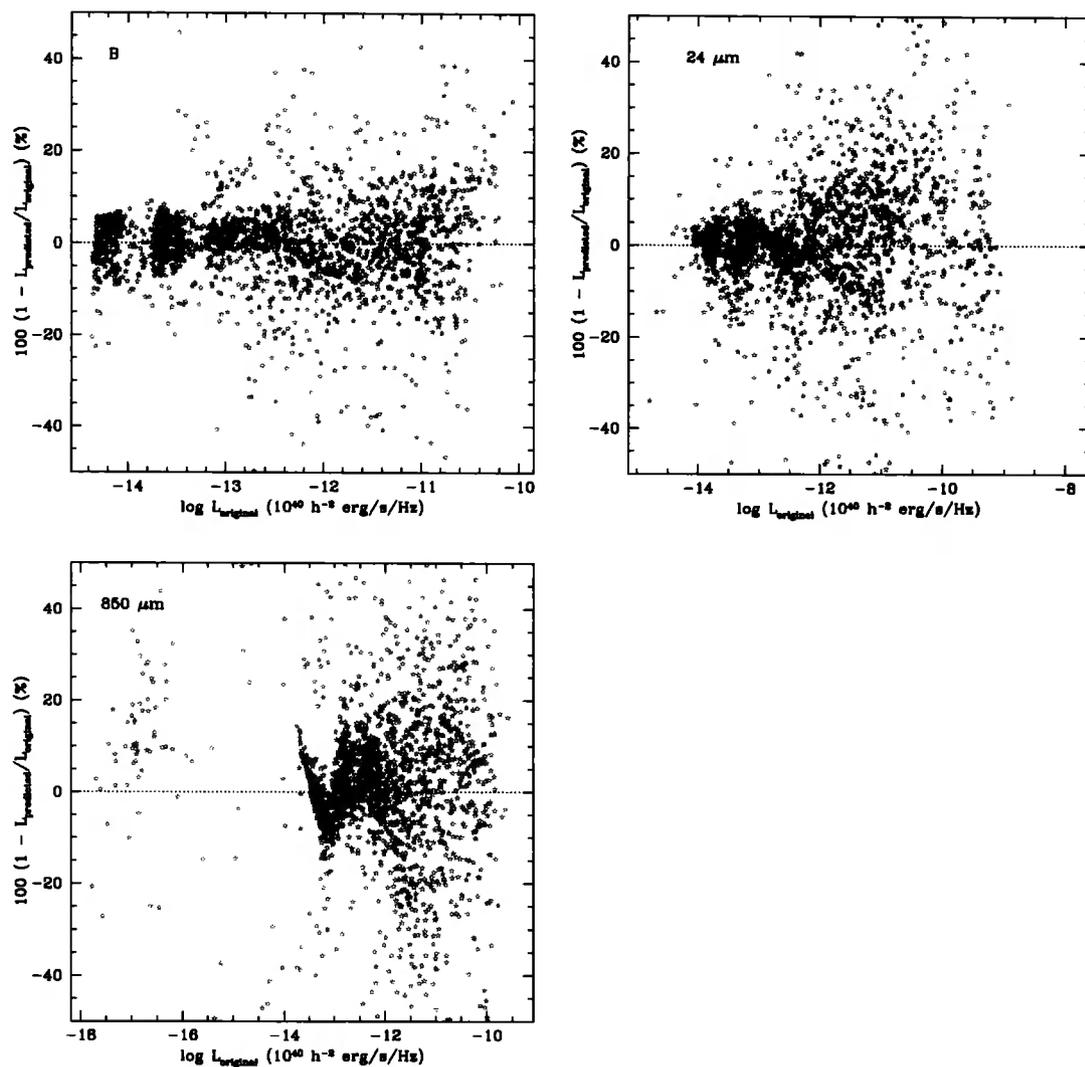
Figure 5.7: The percentage error (as given by 100(1-predicted luminosity/original luminosity)) of the spectra predicted by the ANN, for different bands, when using the entire spectrum as the output layer. From left to right and top to bottom we have: B, MIPS 24 μm and SCUBA 850 μm bands.

Band	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
Bolometric	3.72	81.4	-35.65	-3.10	0.01	4.89	23.77
B	0.12	79.1	-45.35	-4.89	0.10	3.78	23.63
MIPS 24 μm	0.23	76.0	-70.54	-4.53	-0.48	5.31	38.73
SCUBA 850 μm	1.61	60.5	-106.05	-8.79	0.82	7.36	45.65

Table 5.1: Summary statistics for the distribution of error of the spectra predicted by the ANN, when using the entire spectrum as the ANN output layer, for the bolometric luminosity and three different bands: B, MIPS 24 μm and SCUBA 850 μm. $rrmse$ represents the relative root mean square error, as calculated by $rrmse = \sqrt{(\sum(L_p - L_o)^2/L_o^2)/n}$. $P_{|e|<10\%}$ shows the percentage of galaxies with predicted luminosities within 10% of the true values. p_1, Q_1, Q_2, Q_3, p_{99}, represent the 1st percentile, 1st quartile, median, 3rd quartile and 99th percentile of the error distribution, respectively.

In summary, this first attempt to predict spectra (luminosities) from a given set of galaxy properties, using artificial neural networks, has proven to work reasonably well. We find that around 81% of the predicted spectra show no more than 10% of error, and 98% of the population has errors smaller than 35%. However, there is a considerable error associated with this method, particularly for wavelengths where the molecular clouds and cirrus components are dominant over the stellar light.

5.3.3 Using Principal Components Analysis

A simple way of helping the ANN in the prediction of spectra is to reduce the dimensionality of the output, i.e. the spectrum. The effect is clear: the ANN should converge more rapidly to a trained network because there will be fewer weights to be adjusted. The principal components analysis (PCA) is a statistical technique that is designed to do so. In essence, PCA is a linear orthogonal transformation that it is commonly used to identify patterns in data, and express them using a new set of basis-vectors (the principal components), which span in directions of maximum variance.

We now summarize the main steps of this technique, as applied to galaxy spectra. Let us suppose that we have a dataset of n galaxy spectra and each of them can be described by a m-dimensional vector, $\vec{x}$, containing the galaxy's spectrum (for our galaxies, we have $m = 456$). Our data sample will therefore consist of $n \times m$ data points. The first step of the PCA is to subtract the mean from all data dimensions, i.e. the mean of all the m data vectors will be zero:

$$\xi_{ij} = x_{ij} - \bar{x}_j,\tag{5.4}$$

where $i = 1, n$ and $j = 1, m$, and the mean is

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}.\tag{5.5}$$

We then compute the covariance matrix

$$C_{jk} = \frac{1}{n} \sum_{i=1}^{n} \frac{\xi_{ij}\, \xi_{ik}}{s_j\, s_k},\tag{5.6}$$

where the variances of each variable are given by:

$$s_j^2 = \frac{1}{n} \sum_{i=1}^{n} \xi_{ij}^2.\tag{5.7}$$

To find the axes of maximum variance we find the eigenvectors and eigenvalues of the covariance matrix:

$$\vec{C}\, \vec{e}_j = \lambda\, \vec{e}_j.\tag{5.8}$$

The assumption of linearity is present here: we assume that the observed data is a linear combination of the new eigenvectors basis.

The next step is to sort the eigenvectors in decreasing order of their eigenvalues, which translates into decreasing order of variance. It is now, when the reduction of dimensionality comes into place: we can decide how many eigenvectors we want to retain (i.e. how many eigenvectors we think are sufficient to describe the data reasonably). We select the largest eigenvalues, which correspond to the axis where the variance is higher. Once we have selected p eigenvectors, we are ready to project our original data onto this new basis and retrieve the principal components (PCs) of the spectra:

$$\vec{A} = \vec{\xi}\, \vec{E}_p.\tag{5.9}$$

Number of Principal Components	$rrmse$
1	15.64
3	0.28
5	0.12
10	0.04
20	0.01
50	3.3×10^{-4}
100	9.3×10^{-5}
456	0

Table 5.2: The relative root mean square error, $rrmse$, of reconstructed spectra using principal components.

To go back to the original basis, from $p \leq m$ eigenvectors, and reconstruct the original data (at least partially), we use:

$$\vec{x}_{\text{rec}} = \vec{A}\,\vec{E}_p^{-1} + \bar{x}. \tag{5.10}$$

Tab. 5.2 shows how well we can reconstruct the original spectra using its principal components. Evidently, when using the entire set of principal components, we do not lose any information, and hence we are able to recover the entire spectrum; 20 principal components are enough to extract 99% of the spectral information.

The implementation of this method, in the ANN framework, is straightforward. We will still use the same input layer as in Subsection 5.3.2, i.e. 14 neurons corresponding to the selected galaxy properties, and 50 neurons will constitute the hidden layer. The number of output neurons is defined by the number of principal components that we want to use. Once we establish this number, and having calculated the principal components and eigenvectors of the training sample, we use the ANN to predict the principal components instead of the full spectrum, from the given galaxy properties. Therefore, the network architecture is 14:50:n PCs. The final step is to reconstruct the entire spectrum from the predicted components, using Eq. 5.10, where $\vec{E}_p$ and $\bar{x}$ are the eivenvectors and mean of the training sample, and

Band	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
Bolometric	4.49	81.8	-36.35	-3.54	0.01	5.24	26.97
B	0.14	75.4	-49.13	-5.43	0.83	5.29	23.16
MIPS 24 μm	0.27	76.1	-86.33	-4.21	0.24	4.58	43.64
SCUBA 850 μm	2.41	66.1	-105.94	-5.50	0.82	6.33	45.69

Table 5.3: Summary statistics for the distribution of error of the predicted spectra by the ANN, when using 20 principal components as the output layer, for the bolometric luminosity and three different bands: B, MIPS 24 μm and SCUBA 850 μm. See description of quantities in Tab. 5.1.

$\bar{A}$ the predicted PCs. We use the same procedures and ANN parameters as the ones defined in the previous subsection.

In Fig. 5.8 we plot the error of predicted spectra using 20 principal components (see also Tab. 5.3). Using the principal components instead of the entire spectrum does not seem to help the accuracy of the training process much. Quantitatively, the results achieved with this technique are similar to those in the previous subsection: 81% of the predicted spectra show bolometric luminosities that lie within 10% of the original value, and a $rrmse = 4.5$. For the different bands, the figure and table show an analogous distribution. Interestingly, we notice that both methods show similar difficulties when predicting the spectra, i.e. if a particular spectrum was not forecast accurately by the method using the entire spectrum as the ANN output layer, it is likely that it will differ substantially from its original value. This was expected because, as seen in Tab. 5.2, with 20 PCs we only lose 1% of the spectral information.

The main advantage that this process offers is the reduced computing time compared with the method described in the previous subsection: the amount of time is reduced by a factor of ~ 15.

Tab. 5.4 shows the errors on the bolometric luminosity of predicted spectra, using artificial neural networks with different number of principal components. The ANN predictions using PCs do not improve when using more than 10 values, as indicated by the percentage of spectra within 10% of error, approximately a constant value of
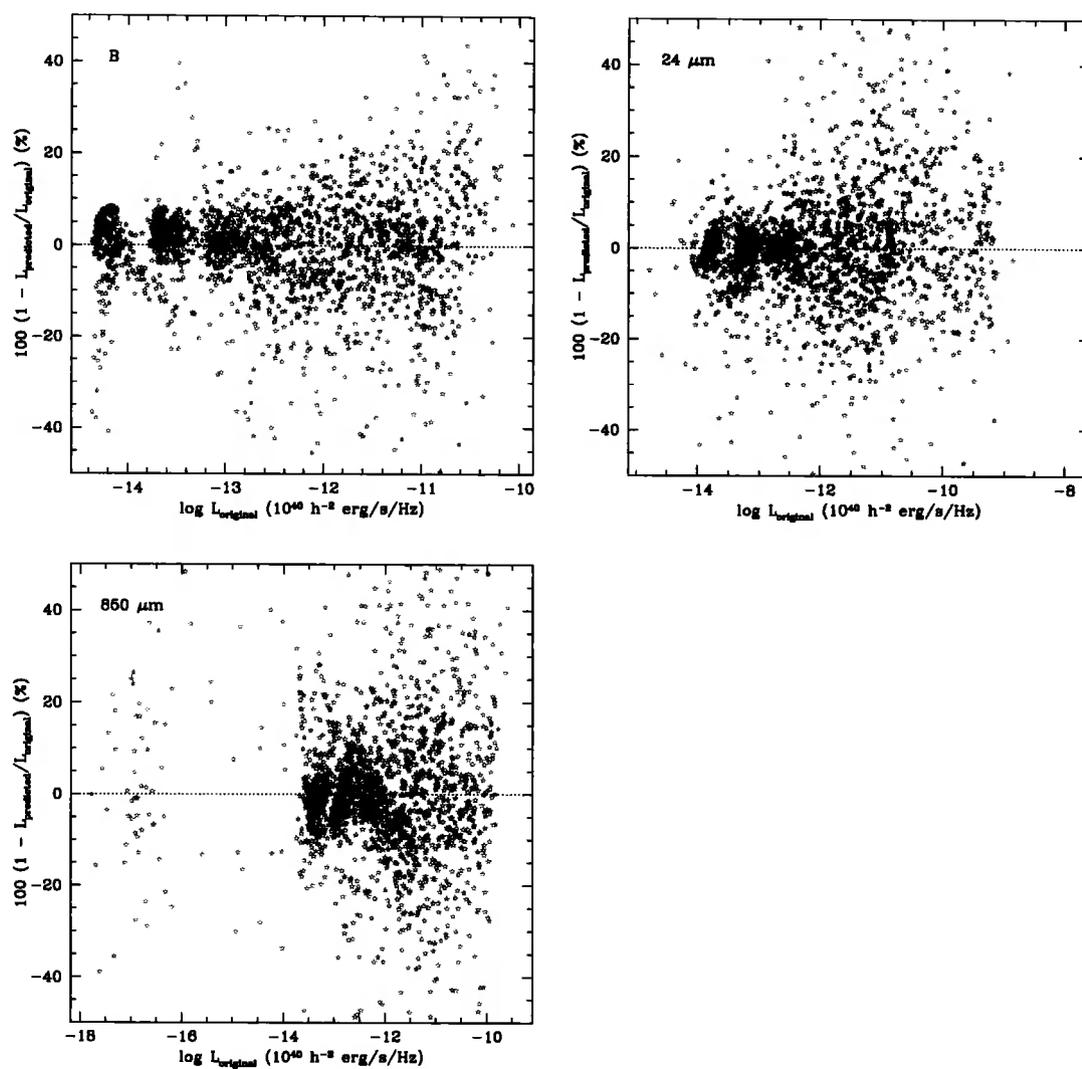
Figure 5.8: The error of the predicted spectra by ANN using the 20 principal components, for different bands. From left to right and top to bottom we have: B, MIPS 24 μm and SCUBA 850 μm bands.

Number of Principal Components	Bol. $rrmse$	$P_{|e|<10\%}$
1	62367.21	24.5
3	262.61	34.2
5	39.07	54.3
10	4.98	81.3
20	4.49	81.8
50	5.42	80.7
100	14.67	81.2
456	57.35	80.0

Table 5.4: Summary statistics of the bolometric luminosity error distribution of predicted spectra, for neural nets using different number of principal components (see Tab. 5.1 for a detailed description of the quantities).

$\approx 80\%$. Once again, we note that Tab. 5.2 showed that most of the information in the spectrum can be reconstructed with 10 or more principal components, hence we expect the ANN to behave in a similar way as when using the entire spectrum for the prediction.

5.3.4 Predicting Luminosities

As we saw in the previous subsection, using just a few principal components of the spectrum instead of the whole 456 flux bins facilitates the training process due to the reduced number of ANN internal weights which need to be adjusted. In this section we explore another possibility: the output of the luminosity in a single given band instead of trying to predict the full spectrum. The neural net becomes simpler in the sense that we only need to predict one variable –the luminosity (later, we also explore the use of more than one luminosity at a time). Evidently, we expect this method to perform better than the previous ones, however it has a drawback: we can only predict the luminosity in a particular band (or bands), not the entire spectrum. This means that, if, eventually, we need to calculate the luminosity of a

Band	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
B	0.06	95.7	-14.50	-1.67	-0.05	1.51	14.34
MIPS 24 μm	0.28	80.9	-55.76	-2.55	0.17	2.76	38.78
SCUBA 850 μm	0.83	88.7	-74.58	-2.40	0.08	2.59	46.59

Table 5.5: Summary statistics for the distribution of error of the luminosities (B, MIPS 24 μm and SCUBA 850 μm) predicted by the ANN, using one output neuron. See description of quantities in Tab. 5.1.

particular galaxy in a different band than the one we use to train the network, we need to retrain the ANN.

To predict luminosities we need to preprocess the spectra to calculate the luminosities in a predefined set of bands (in this study: B, MIPS 24 μm, and SCUBA 850 μm). Initially we will start with a network configuration of 14:50:1, i.e. 14 neurons in the input layer, 50 neurons in the hidden layer and 1 output neuron corresponding to the desired luminosity. The ANN is trained for each of the selected bands, using procedures and parameters similar those used in Subsection 5.3.2. The results are shown in Fig. 5.9, and the performance statistics are given in Tab. 5.5.

This method performs much better than the previous neural nets: the number of spectra (luminosities) with errors smaller than 10% increases substantially from a three-band average of $\sim$ 72% to 88%; similar successes are achieved in terms of the relative root mean square error. The results are particularly impressive for the B-band, where the $rrmse = 0.06$ and more than 95% of the population have predicted luminosities within 10% of error. Just like for the previous methods, for near and far-infrared/submillimeter bands the ANN could not produce as good results as for the B-band, mainly due to the variance of the spectra in these regimes (see Fig. 5.5). No correlation was found between the errors and the galaxy properties at the output time.

As mentioned earlier, we also have the possibility of using the ANN to predict n luminosities at a time. The advantage is that we only need to train the network once, instead of having to train it n times. Tab. 5.6 shows the statistics of the error
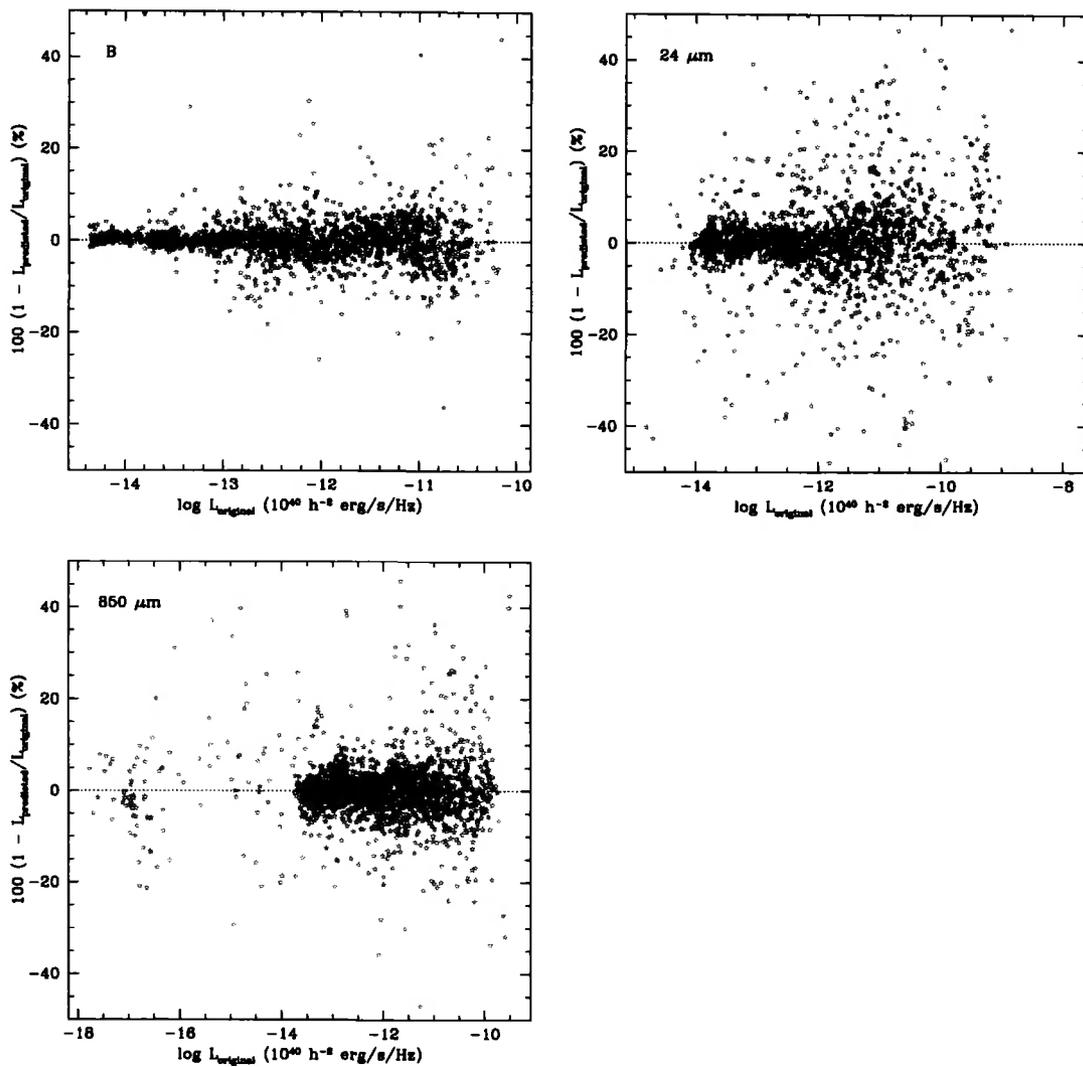
Figure 5.9: The error of the luminosities predicted directly by ANN. From left to right and top to bottom we have: B, MIPS 24 μm and SCUBA 850 μm bands.

Band	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
B	0.09	92.7	-19.49	-2.43	-0.18	1.87	16.07
MIPS 24 μm	0.31	77.4	-62.51	-5.19	-0.19	3.98	33.73
SCUBA 850 μm	1.52	84.3	-59.39	-3.70	-0.18	3.16	39.92

Table 5.6: Summary statistics for the distribution of error of the luminosities predicted by the ANN, using three output neurons: B, MIPS 24 μm and SCUBA 850 μm. See description of quantities in Tab. 5.1.

distribution using this procedure: the ANN is trained using three neurons in the output layer, one for each band (B, MIPS 24 μm and SCUBA 850 μm). As expected, this variant method does not perform as well as when training the ANN three times, one for each of the luminosity bands, with a percentage of 92%, 77% and 84% of galaxies within 10% of luminosity error, in the B, MIPS 24 μm and SCUBA 850 μm bands respectively. Nevertheless, it still outperforms the first two methods explored, where we predict the entire spectrum.

For the remainder of this Chapter, we focus on the predictions of the method described at the beginning of this subsection: we use only one output neuron, which represents the luminosity in a pre-selected band.

5.3.5 Performance

In this subsection we will explore the architecture and some of the parameter-space of the neural networks, in order to justify our default set-up, and also the influence of sample extraction and redshift on the results.

Architectures

So far, we have used a supervised feed-forward neural network with 14 input neurons, one hidden layer with 50 neurons and an output layer with one or more neurons, depending on the method used. The number of input and output neurons are quantities that are reasonably easy to define. On the other hand, the number of hidden layers and their population of neurons are parameters that are more subjective.

Currently, there is no clear view as to how many hidden units should be used (see Scarselli & Tsoi, 1998). Each problem/application will have its optimal parameters, which can only be found by trial and error. In Fig 5.10, the upper and lower panels show the evolution with the number of neurons of the relative root mean square error and percentage of galaxies with predicted luminosities within 10% of the target, respectively. Here, we represent the three different bands in study (B, MIPS 24 μm and SCUBA 850 μm) by different colours, and note that the training process in all the configurations was stopped at 5000 epochs. We see that there is little variation in the performance of the network for architectures with more than 20-30 neurons in the hidden layer. For two hidden layers, the results are similar to those found for one layer. However, for some cases, e.g. for the SCUBA 850 μm band the usage of a two hidden layer net seems to help in terms of the *rrmse*, which changes by a factor of $\approx$ 2 when two layers are used.

Training Epochs

As explained in Section 5.2, two methods can be used to stop the training process in order to avoid overfitting: a pre-defined error threshold or a maximum number of epochs. For this work, we decided to use a maximum number of epochs, 5000, as the threshold for early stopping. We analyse how the *rrmse* and $P_{|e|<10\%}$ evolve with the number of epochs, when using a network configuration of 14:50:1, in Fig. 5.11. All the three networks show a rapid convergence towards the trained values: after 1000 iterations, the *rrmse* in the B, MIPS 24 μm and SCUBA 850 μm are 0.068, 0.21, 0.99, while the $P_{|e|<10\%}$ are 93.5%, 73.9% and 83.9%, respectively. The plot shows that, at 5000 epochs the networks have already converged to their optimal states, after which there is no noticeable change in the *rrmse* or $P_{|e|<10\%}$. Actually, for the MIPS 24 μm network there is evidence of overfitting, as shown by the increase of the relative root mean square error with the number of epochs beyond 7000 iterations.

The value of 5000 epochs was kept as the early stop parameter of our training.
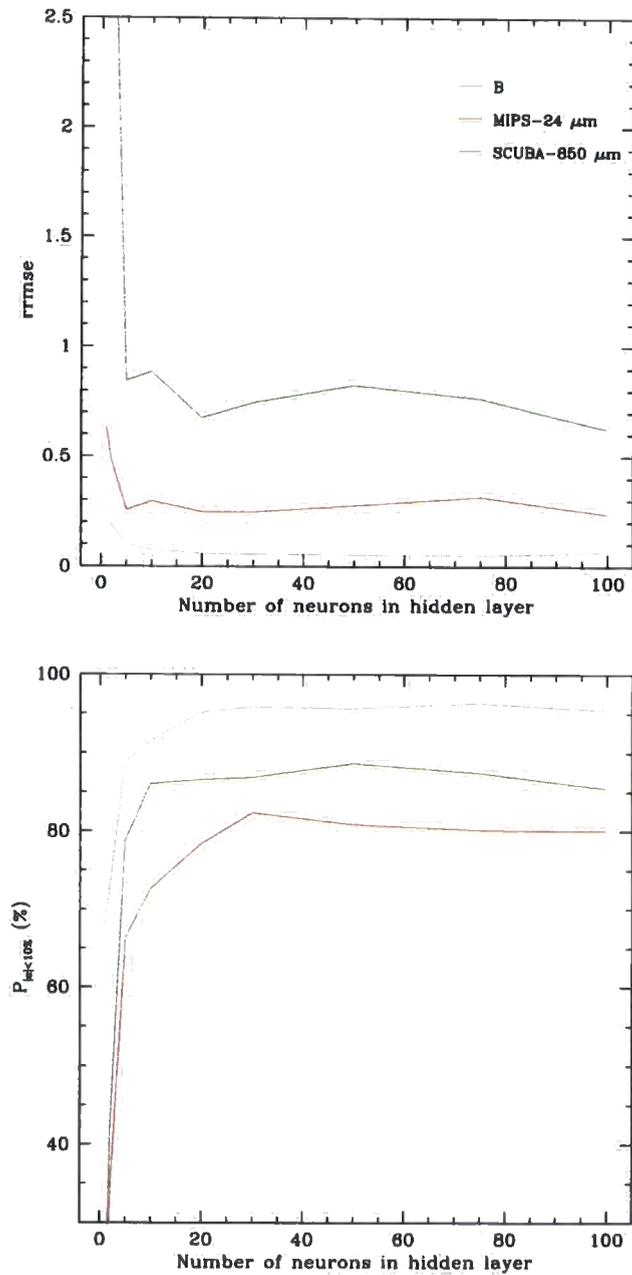
Figure 5.10: The evolution of the *rrmse* (top panel) and percentage of galaxies with luminosities within 10% of the true value (lower panel), as a function of the number of neurons in the hidden layer. For all the tested configurations, we stopped the training process after 5000 epochs. The blue, red and green lines show the results for the B, MIPS 24 μm and SCUBA 850 μm bands, respectively.

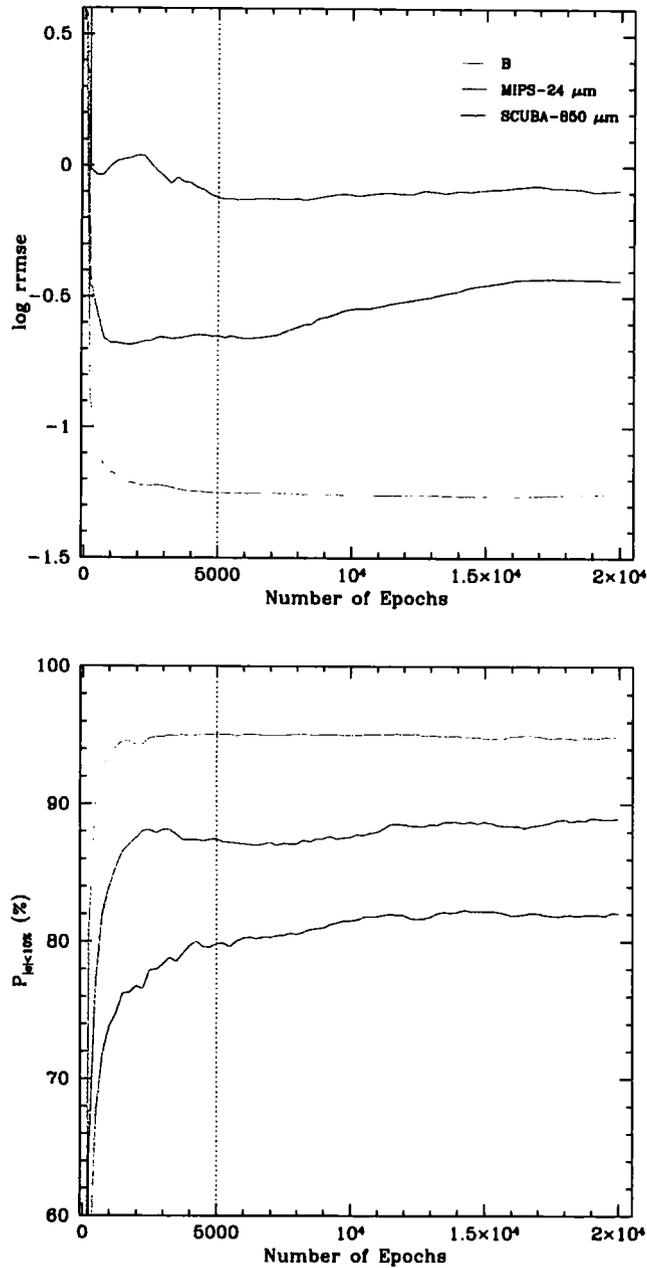Figure 5.11: The evolution of the *rrmse* (top panel) and the percentage of galaxies with predicted luminosities within 10% of the true value (lower panel), $P_{|e|<10\%}$, as a function of the number of training epochs, for a net configuration of 14:50:1. The blue, red and green lines show the results for the B, MIPS 24 μm and SCUBA 850 μm bands, respectively. The dotted line shows 5000 training epochs, for reference.

Activation Function	B		MIPS 24 μm		SCUBA 850 μm	
	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$
Elliot	0.06	95.7	0.28	80.9	0.83	88.7
Sigmoid	0.07	95.8	0.24	80.5	0.71	89.1
Gaussian	0.09	94.7	0.40	80.2	1.77	83.2
Linear	0.21	70.0	0.73	21.3	5.67	11.2

Table 5.7: The performance of neural nets with different hidden layer activations functions: Elliot, Sigmoid, Gaussian and Linear (see Tab. 5.1 for a detailed description of the quantities).

Activation Functions

Activation functions are needed in order to add nonlinearity to the training process. So far, for the hidden neurons, we have been using the Sigmoid function given by: $f(x) = \frac{1}{1+e^{-\alpha x}}$, where the coefficient α is commonly referred to as the steepness of the activation function; the steepness of choice was $\alpha = 0.01$. For the output neurons, the linear activation function, $f(x) = \alpha x$, was adopted.

In Tab. 5.7 we list the performance of ANNs with different activation functions for the hidden layer in the B, MIPS 24 μm and SCUBA 850 μm bands. The best results are achieved with the Elliot, $f(x) = \frac{\alpha x}{1+|\alpha x|}$, and the Sigmoid activation functions; this last one, seems to slightly outperform the first, particularly in the submillimeter. If a Gaussian function is used instead, $e^{-(\alpha x)^2}$, the $rrmse$ increases by a factor of ~ 2 at 24 μm and 850 μm. We also tried to set a linear activation function for all the neurons in the hidden layer. This should have the effect of removing all the nonlinearity from the neural network training, and consequently the ANN will be no more than a simple perceptron (see Section 5.2). Evidently, the capabilities of neural network are greatly reduced when we use only linear correlations: the errors associated with the predicted luminosities increase, which leads to merely 11% of galaxies with errors smaller than 10%, in the SCUBA 850 μm band; similarly poor results are obtained in the 24 μm band.

In view of these results, we will still use the Sigmoid activation function through-

Network		B		MIPS 24 μm		SCUBA 850 μm	
Train	Test	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$
Normal	Normal	0.07	95.8	0.24	80.5	0.71	89.1
Normal	Burst	0.84	40.6	1.53	15.2	27.71	8.1
Quiescent	Quiescent	0.03	97.9	0.09	87.7	0.04	97.4
Burst	Burst	0.08	89.9	0.21	62.8	0.26	86.1

Table 5.8: The performance of several neural networks, trained with different galaxy samples: "normal", "quiescent" and "bursts". See text for further details, and also Tab. 5.1 for a description of the quantities listed).

out this Chapter.

Normal and Burst Sample

Until now, we have been using a sample extracted from the GALFORM catalogue following a similar procedure as in Granato et al. (2000) for their "normal galaxies" (see definition below). In this subsection, we distinguish between quiescent and burst galaxies (i.e. galaxies that had a burst recently or which are undergoing a burst event) and analyse the performance of the ANN in both cases. Quiescent and bursts galaxies are sampled differently. The reason behind this dichotomy is that typically the GALFORM catalog contains only a few examples of bursts, and so careful processing is needed to produce a statistically significant sample. In the present work, we do not calculate the burst spectrum for a set of times after the start of the burst, as in Granato et al. (2000). Instead, we enlarge the burst sample by increasing the sampling volume.

A given galaxy is considered a burst galaxy if $10\,\tau_e \geq t_{burst}$, t_{burst} is the lookback time to the last burst in a galaxy, and τ_e the burst e-folding time of star formation. In Tab. 5.8 we show the performance of the artificial neural network when applied to the quiescent and burst samples separately. Here, the "normal sample" represents the sample extracted from the GALFORM catalogue using the procedure described previously, sampling the GALFORM catalogue such that we have equal numbers of

galaxies in logarithmic bins of total stellar mass. This selection also picks a small fraction of galaxies that are undergoing a burst event. The "quiescent sample" is a "bursts-clean" version of the "normal sample", i.e. a "normal sample" for which we selected galaxies with $10\,\tau_e < t_{\mathrm{burst}}$. The "burst sample" was extracted differently from the GALFORM catalogue. We selected equal numbers of galaxies in logarithmic bins of burst mass. Then, we filtered the sample, retaining the galaxies that had experienced a recent burst. Tab. 5.8 clearly shows the difficulty experienced by the ANN in predicting the spectra of burst galaxies, mainly due to the large variety in their spectra. When we train the neural network using the "normal sample" and then apply to the "burst sample", the performance of the ANN is greatly reduced. The three-band $P_{|e|<10\%}$ average in this case drops from 88.5% to 21.3%, mainly due to the infrared (IR) bands. In order to improve these results, we constructed a quiescent and a burst sample, as described above. Training and testing the network independently for these two samples produces better results. For quiescent galaxies, the $rrmse < 0.09$ for the three bands studied, with equally impressive results for the distribution of errors ($\sim 90\%$ of the galaxies show predicted luminosities that lie within 10%). The results for burst galaxies are not so remarkable, particularly for the MIPS 24 μm band, where $P_{|e|<10\%} = 62.8\%$. However, they clearly outperform the "normal sample" trained ANN, with a three-band $P_{|e|<10\%}$ average of 79.6%. Therefore, this procedure will be adopted from next Section onwards.

Other Redshifts

It is necessary to analyse how the trained network performs at other redhifts. This is important in the sense that, if we find that the ANN training process is independent of redshift, then there will be no need to retrain the network for new redshifts, saving computing time. So far, we have been analysing the ANN predictions at redshift $z = 0$. Tab. 5.9 compares the relative root mean square error and the percentage of galaxies with predicted luminosities (in the rest frame) that lie within 10% of the original values, for "normal samples" (see above) at redshift $z = 0$ and $z = 2$. The network trained at $z = 0$ performs reasonably when used to predict the luminosities at $z = 2$: it is capable of reproducing $\approx 55\%$ of the luminosities within an error of

Network		B		MIPS 24 μm		SCUBA 850 μm	
Train	Test	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$	$rrmse$	$P_{|e|<10\%}$
$z = 0$	$z = 0$	0.07	95.8	0.24	80.5	0.71	89.1
$z = 0$	$z = 2$	0.33	56.7	0.57	55.1	0.34	59.7
$z = 2$	$z = 2$	0.24	81.1	0.31	74.8	0.17	91.9

Table 5.9: The performance of the ANN trained with galaxy samples extracted at different redshifts: $z = 0$ and $z = 2$. See Tab. 5.1 for a description of the quantities. The luminosities were calculated in the rest frame.

10%, and $rrmse$ smaller 0.57 in the B, MIPS 24 μm and SCUBA 850 μm bands. However, it is strongly advisable to train the ANN at the redshift of choice, as indicated by the table. If we train the net at $z = 2$ instead and apply at $z = 2$, the variance of the error is reduced, and the forecasted luminosities are much more accurate (we achieve a three-band $P_{|e|<10\%}$ average of 82.6%).

5.4 Luminosity Functions

We are now in good shape to predict accurate luminosities for the GALFORM catalogue. In the previous section, we showed the first results of this new technique (ANN). We calculated the luminosities for a sample of GALFORM galaxies in the B, MIPS 24 μm and SCUBA 850 μm bands, and found that more of than 80% of the galaxies show ANN predicted luminosities with errors within 10%. This Section expands the previous tests to show the impact of the errors on the predicted luminosity functions (LFs). It is crucial to closely reproduce the original luminosity function, i.e. the function obtained when using the original spectra (luminosities), because this quantity is the most important statistical description of a galaxy population, and so an essential requisite for any model. Here, we present predictions for the Lyman-break ($z = 3$), submillimeter ($z = 2$) and IR galaxies ($z = 0.4$).

Throughout this section we will use a 14:30:30:1 ANN architecture, which corresponds to 14 input galaxy properties as the input layer, 30 neurons in each of the two hidden layers and one output neuron representing the galaxy luminosity in the

Sample	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
Quiescent	0.21	88.0	-64.79	-1.62	-0.19	1.44	38.15
Burst	3.41	49.0	-244.53	-9.81	0.67	10.63	78.19

Table 5.10: Statistics for the error distribution of the predicted rest frame UV luminosities by the ANN, for quiescent and burst galaxies. The listed quantities are described in Tab. 5.1.

predefined band (see previous section for further details). The Sigmoid function was the activation function of choice for the hidden neurons. Driven by the conclusions, of the previous section, we split our sample into two: quiescent and burst galaxies, and trained the network at the selected redshift.

5.4.1 Lyman-break Galaxies

Lyman-break galaxies (LBGs) have been considered a key population in the understanding of the star formation history of the Universe (Steidel et al., 1999). These high-redshift galaxies, $z \gtrsim 2$, have been detected using a technique that involves imaging through several optical bands, in the observer frame, in order to extract information about the rest frame spectral break around $912\,\mathring{A}$ (which is caused by the absorption of almost all Ly-limits photons with $\lambda < 912\,\mathring{A}$, providing that there is sufficient neutral hydrogen).

In this subsection we use the ANN to predict the luminosities of LBGs in the rest frame UV wavelength of 0.17 μm at $z = 3$, which corresponds to the observer frame R band. We compare with the original luminosities extracted from GRASIL spectra and compare the luminosity functions for both cases.

In Fig. 5.12 we plot the error associated with the ANN predicted UV luminosities against the original luminosities (the target values), for quiescent (top panel) and burst (bottom panel) galaxies. The error distribution is plotted in Fig. 5.13, and the statistics of the distribution are listed in Tab. 5.10. In the UV regime, the neural network shows better performance for quiescent galaxies than for burst galaxies. For most of the quiescent galaxies, 88% of the sample, the ANN predicts luminosities
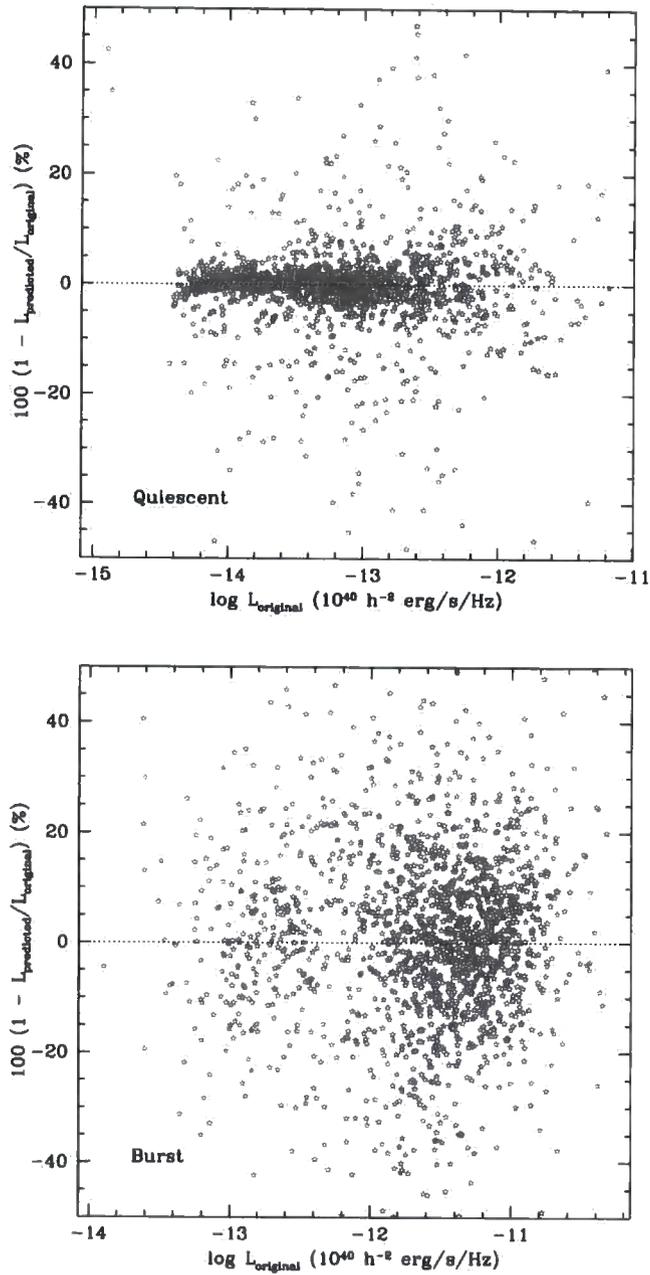
Figure 5.12: Comparison between the ANN predicted rest frame UV ($\lambda = 0.17\,\mu$m) luminosities and the original luminosities, as extracted from GRASIL model spectra. The top and bottom panels show the results for the quiescent and burst galaxies, respectively.
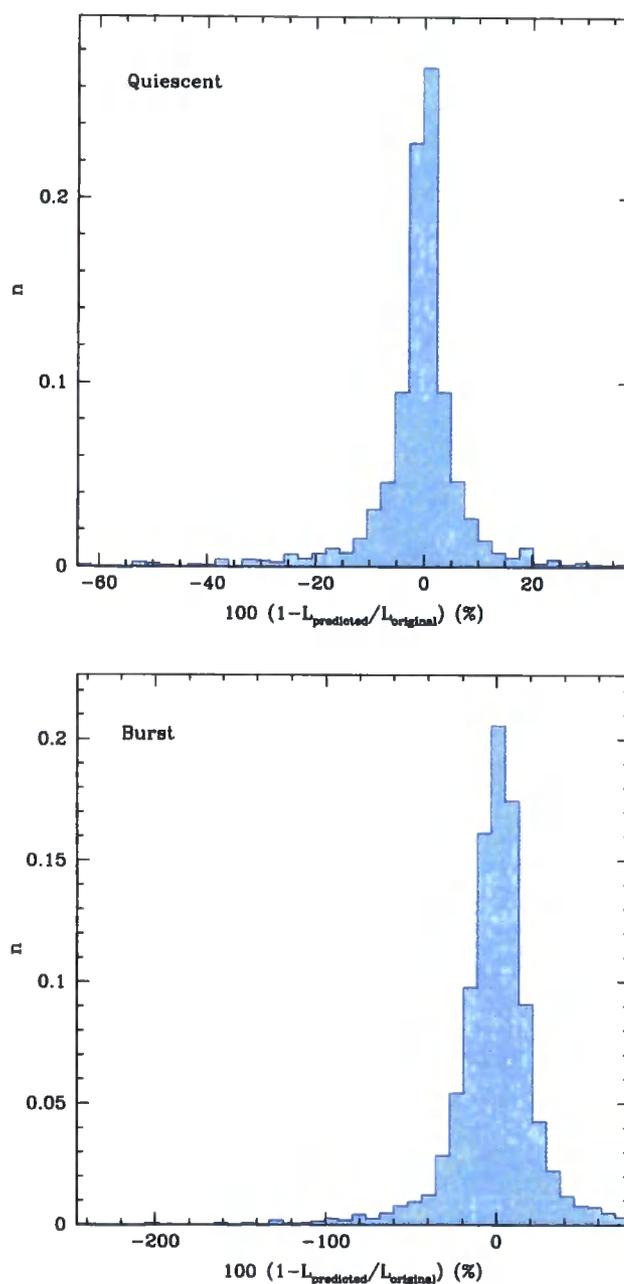
Figure 5.13: Distribution of the error associated with the rest frame UV ($\lambda = 0.17\,\mu$m) luminosities predicted by the ANN. The error is defined by $100(1 - L_{\text{predicted}}/L_{\text{original}})$. The range of the graphs are set to the 1st and 99th percentiles of the distribution, and the distributions are normalized to give $\sum_i n_i = 1$. The top and bottom panels show the results for the quiescent and burst galaxies, respectively.

within 10% of error, and only $\approx$ 4 of galaxies with errors bigger than 20%. Burst galaxies show somewhat a bigger scatter around the expected values, most likely due to the high variance of their spectra at 0.17 μm. The $rrmse$ is rather high, 3.41, while 50% of the burst population have errors below 10%, and 21% of the total lie outside the 20% error threshold. Fig. 5.12 shows another important characteristic of the error distribution: independence of luminosity. This factor will be very important when calculating luminosity-dependent quantities (e.g. luminosity functions) and subsampling using luminosity. In an effort to reduce the error associated with the predicted luminosities, we also investigated the relation between the errors and several galaxy properties for the burst sample. In Fig. 5.14 we plot the error against: bolometric luminosity, central V-band extinction optical depth, disc and bulge size and circular velocities, bulge-to-total mass ratio, disc, bulge and total stellar mass, stellar metallicity, mass of stars formed in the last burst and mass of host halo. Here we only plot the burst galaxies. The figure reveals no clear correlation between the error associated with the predicted 0.17 μm luminosities and the galaxy properties, which shows that any sample drawn from this new method (ANN), using galaxy properties, should be ANN error-independent. Different network configurations were tried in order to improve the performance, but without any success.

The rest frame 0.17 μm luminosity functions at $z = 3$ are presented in Fig. 5.15. The solid lines represent the luminosity functions calculated from the original luminosities, as extracted from GRASIL model spectra, while the dashed lines show the luminosity functions measured from the predicted ANN UV luminosities. The colour scheme differentiates between all (black), quiescent (green) and burst (red) galaxies. Observational data from Steidel et al. (1999) is shown by the purple points. The figure reveals an excellent agreement between the luminosity function calculated from the original 0.17 μm luminosities and the ANN predicted values, for both quiescent and burst galaxies. This result is impressive for burst galaxies, taken into account that a significative fraction of the sample has error bigger than 20%. However, we tested the impact of errors of this magnitude in the LF by randomly perturbing the true luminosities by the distribution of errors expected for the burst galaxies in the UV band. In summary, if the LF is close to a power law we would need much
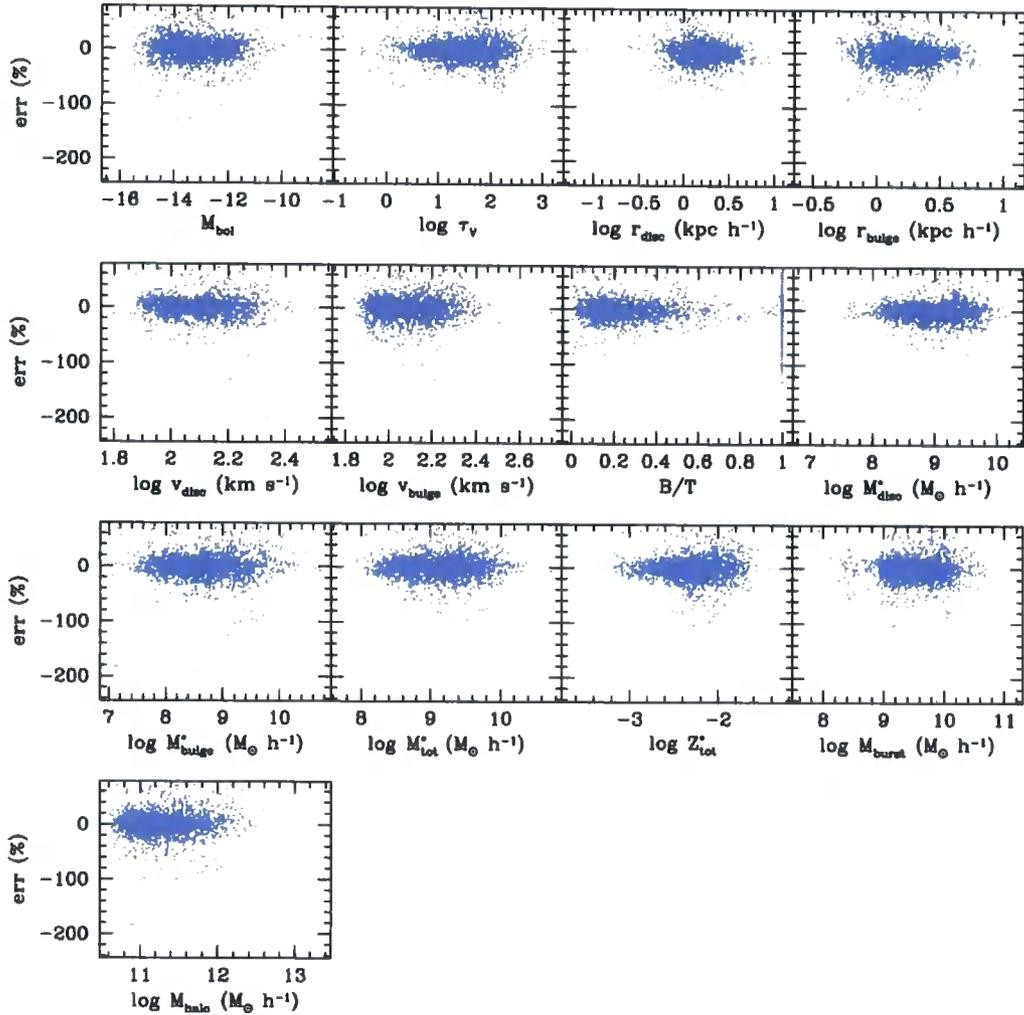
Figure 5.14: The relation between the error associated with the rest frame UV ($\lambda = 0.17\,\mu$m) luminosities predicted by the ANN and various galaxy properties, for the burst sample. From left to right and top to bottom we have: bolometric magnitude, central V-band extinction optical depth, disc and bulge size, disc and bulge circular velocity, bulge-to-total mass ratio, disc, bulge and total stellar mass, stellar metallicity, mass of stars formed in the last burst and mass of host halo. The error is defined by $100(1 - L_{\text{predicted}}/L_{\text{original}})$. The range of the graphs are set to the 1st and 99th percentiles of the error distribution.
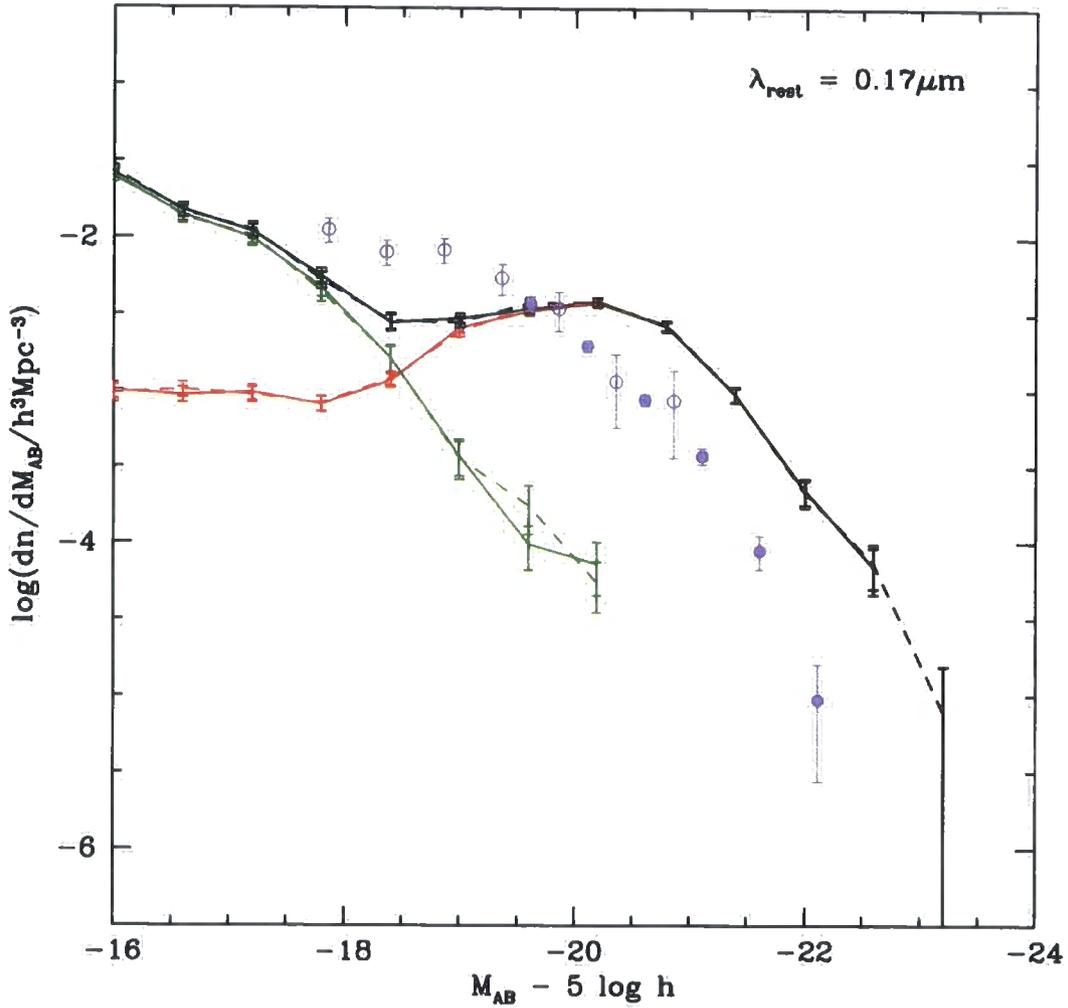
Figure 5.15: Comparison between the rest frame 0.17 μm luminosity function at $z = 3$, calculated from the original luminosities (solid lines), as extracted from the GRASIL spectra, and the luminosity function constructed from the ANN predicted luminosities (dashed lines). The black lines show the total luminosity function, whose components are quiescent galaxies (green lines) and burst galaxies (red lines). The error bars in the model data indicate the Poisson uncertainties due to the number of galaxies simulated. The purple points represent the observational data from Steidel et al. (1999): closed circles from ground-based data and open circles from the *Hubble Space Telescope*.

larger errors than we have to see a discrepancy in the luminosity function. Also, the statistical independence between the error and the UV luminosity helps to keep the luminosity function's shape, while increasing its scatter (see Fig. 5.12).

The model-calculated luminosity functions miss the observed values at both the faint and bright end: at the faint end of the LF, where the function is dominated by quiescent galaxies, the model underpredicts the observed number density of galaxies per luminosity. At the bright end, the LF suffers a boost due to the contribution of burst galaxies, which dominate for $M_{AB} < -18.5$ mag, overpredicting the observed quantity.

In Baugh et al. (2005), the authors analysed the predictions of the GALFORM model for Lyman-break galaxies. Baugh et al. (2005) found that their variant of the model was able to reproduce the UV luminosity functions reasonably well, with a slight underprediction at the faint end (factor $\sim 2-3$). Even though we are using this variant of the GALFORM model, there is an important difference from the 2005 paper: our model is now set in the context of a N-body simulation, with the halo merger trees drawn from the Millennium Simulation (see Chapter 2), which is set in a new cosmology. This increases the number of quiescent galaxies by a factor of ≈ 4, and the burst galaxies by a factor of ≈ 15 at the bright end; at the faint end ($M_{AB} > -20$ mag), the model does not predict enough burst galaxies. Further investigation will be presented elsewhere.

5.4.2 Infrared Galaxies

The advent of new instrumentation able to detect radiation in the infrared, back in the 1980s, revealed a new population of galaxies, whose total luminosity was mostly in the infrared wavelengths. This infrared emission is a product of the heating of the dust (and consequently, emission of radiation) due to the energy emitted by newly formed stars. Further studies made clear that approximately half of the star formation in the local Universe is optically obscured by dust. Later observations by *COBE* (Puget et al., 1996), found that the energy density of the cosmic far-infrared background is similar to that of the optical and near-infrared. These discoveries exposed the importance of this population in the context of galaxy formation and

Sample	$rrmse$	$P_{|e|<10\%}$	p_1	Q_1	Q_2	Q_3	p_{99}
Quiescent	0.12	88.3	-26.45	-2.47	-0.02	2.28	22.68
Bursts	0.34	55.8	-47.87	-9.35	-0.01	8.44	34.45

Table 5.11: Summary statistics for the distribution of error of the predicted 24 μm luminosities by the ANN, for quiescent and burst galaxies. A description of the quantities is given in Tab. 5.1.

evolution, i.e. we cannot have a clear picture of the process of the star formation in the Universe without understanding the evolution of the infrared galaxy population.

This subsection is dedicated to the mid-IR galaxies. Here we predict luminosity in the MIPS 24 μm band, in order to match the recent observations in the infrared by the *Spitzer* satellite (see Lacey et al., 2008, for further comparisons between the GALFORM model and observational data). We selected a galaxy catalogue at $z = 0.4$ for this comparison. This is the predicted median redshift of a sample with a flux limit of $\sim$ 3 mJy (see Fig. A1 of Lacey et al., 2008).

Fig. 5.16 shows the 24 μm predicted luminosity errors, for quiescent and burst galaxies, in the rest frame; the associated statistics are listed in Tab. 5.11. As found in the previous section, the method performs better for the quiescent galaxies, with a relative root mean square error of 0.12 and $P_{|e|<10\%} = 88\%$, at this wavelength. For burst galaxies, the performance is not as good. The percentage of galaxies with errors within 10% is 55.8%. Several ANN architectures and input galaxy properties were tried without any improvement over these figures. As noted in the previous study, the error distribution is not correlated with luminosity. This suggests that the errors might not change the shape of the luminosity function. Also, we found no correlation between the error and several galaxy properties.

In Fig. 5.17, we compare the rest frame 24 μm luminosity functions at $z = 0.4$, constructed from the predicted ANN luminosities and the original luminosities as given by GRASIL. Here, the meaning of the lines is the same as in Fig. 5.15: the black curve represents the total luminosity function, while the green and red lines show the contributions to this from the quiescent and burst galaxies, respectively. The dashed lines show the LF that results from the luminosities predicted by the ANN.
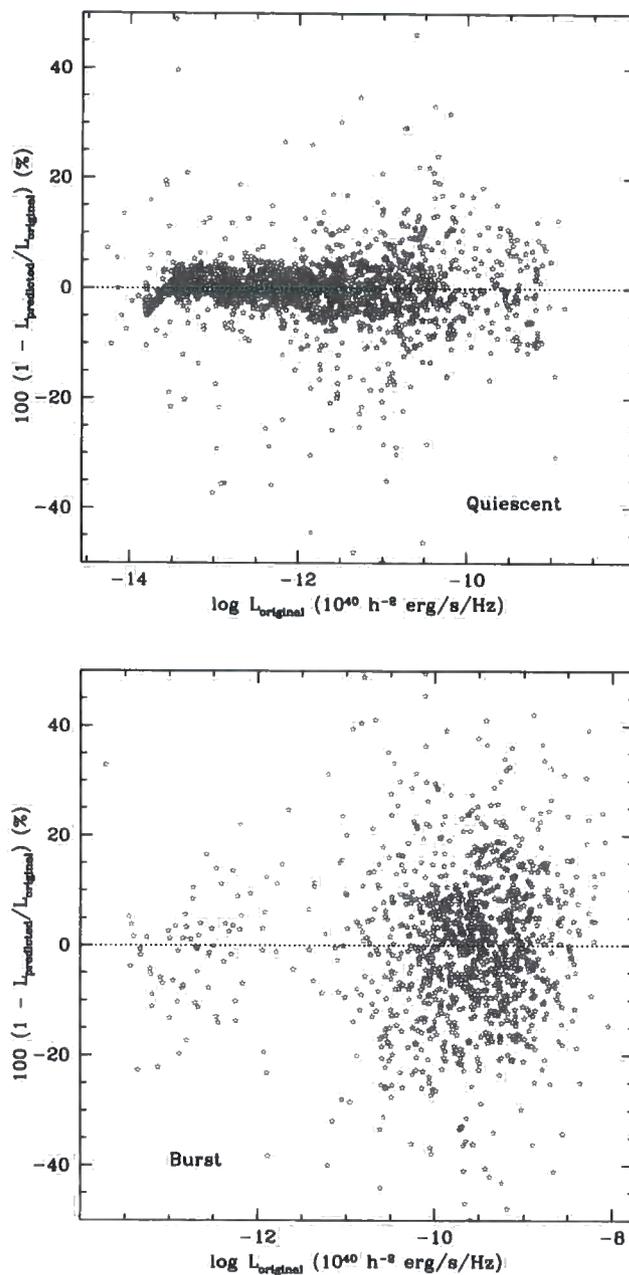
Figure 5.16: The errors associated with predicted rest frame 24 μm luminosities, using ANN, and the their expected value. The results for the quiescent and burst are shown in the top and bottom panel, respectively.
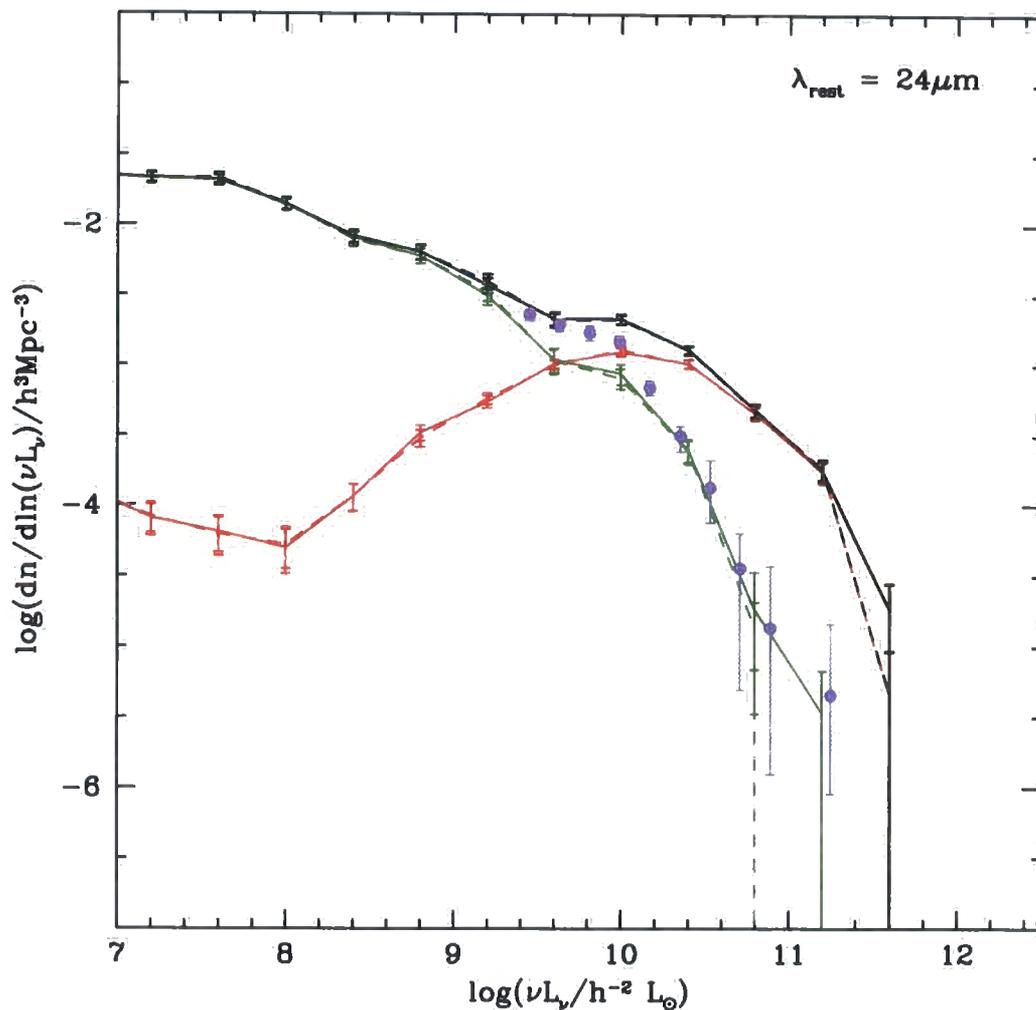
Figure 5.17: Comparison between the 24 μm luminosity function at $z = 0.4$, given by the original luminosities (solid lines), as extracted from the GRASIL spectra, and the luminosity function constructed from the ANN predicted luminosities (dashed lines). The black lines show the total luminosity function, whose components are quiescent galaxies (green lines) and burst galaxies (red lines). The error bars in the model data indicate the Poisson uncertainties due to the number of galaxies simulated. Observational data from Babbedge et al. (2006) is plotted using purple circles.

The purple circles show observational data from Babbedge et al. (2006). Just like in the UV regime, the luminosity functions extracted from the predicted luminosities agree impressively well with the original functions, even for the burst galaxies, where the performance of the ANN was somewhat poorer. As we mentioned earlier, the statistical independence between the errors and the original luminosities, helps to keep the shape of the LF. The largest difference between the predicted and original values comes from the faint end of the luminosity function (which is dominated by the burst galaxies), where there is a factor of $\sim$ 4 discrepancy. However we note that only one galaxy was used to calculate the luminosity function at these fluxes, far from what is desirable for a statistically significant assessment.

As pointed out by Lacey et al. (2008), the model data has some difficulty in reproducing the bright end of the observational 24 μm luminosity function at $z = 0.5$ ($z = 0.4$ in this study). The authors suggested that the discrepancy might be due to uncertainties in the k-corrections applied to the observational data.

5.4.3 Submillimeter Galaxies

Submillimeter galaxies (SMGs) are thought to be dusty starburst galaxies (Smail, 2002). Their emission in the submillimeter region of the spectrum (around 850 μm) is due to the heating of the dust by UV light emitted by young stars (there is also the possibility that part of the submillimeter radiation results from the heating of the dust by an AGN, instead) (Alexander et al., 2003, 2005). These massive galaxies have extreme star formation rates of $\sim 500 - 1000$ $M_{\odot}$ yr^{-1} and are mainly found at redshifts of $z \gtrsim 2$. They provide, therefore, an invaluable opportunity to understand the star formation history of the Universe, when most of the stars were being formed.

Here we predict the rest frame 850 μm luminosities for a sample of galaxies at $z = 2$ using artificial neural networks. We compare our predictions with the correct values extracted from GRASIL spectra, and evaluate the luminosity functions for both the predicted and original luminosities.

A comparison between the 850 μm predicted and original luminosities is presented in Fig. 5.18. More information about the errors is presented in Tab. 5.12.

As shown in the previous section, the ANN predictions for the submillimeter are
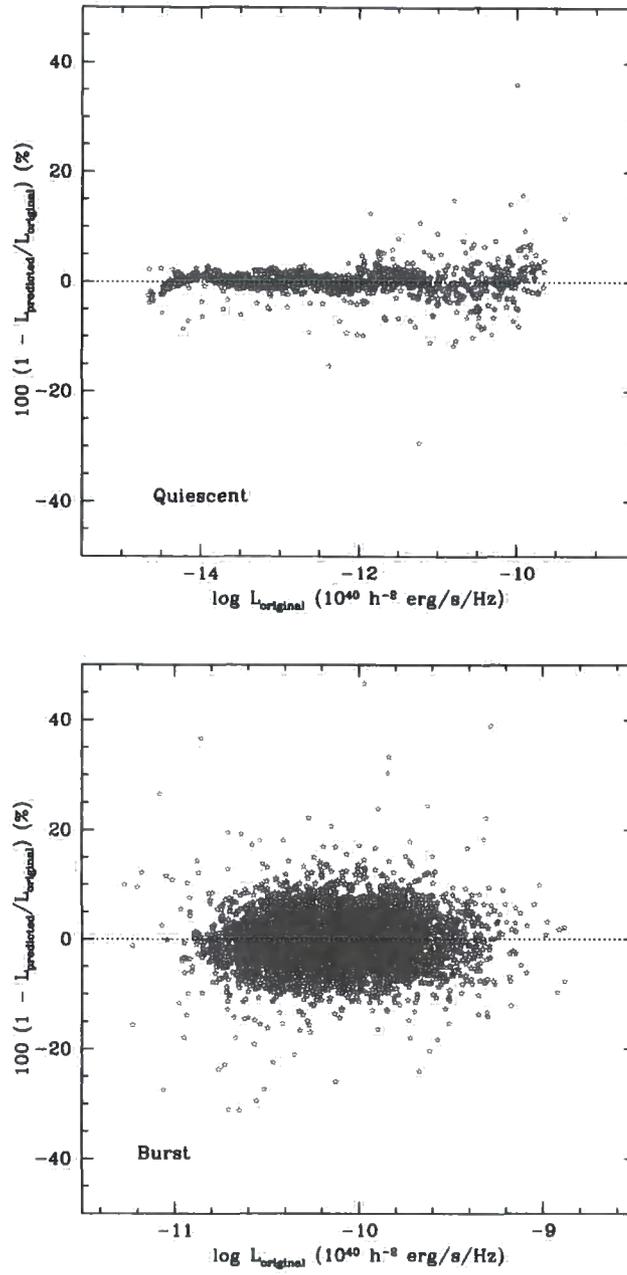
Figure 5.18: The relation between the errors associated with the ANN predicted 850 μm luminosities and the expected values. Quiescent galaxies are shown on the top panel, while the burst sample is plotted in the bottom panel.

Sample	$rrmse$	$P_{|e|<10\%}$	min	Q_1	Q_2	Q_3	max
Quiescent	0.37	97.8	-13.93	-0.95	-0.01	0.93	10.01
Bursts	0.07	94.9	-14.29	-3.39	-0.17	3.10	13.57

Table 5.12: Some statistics from the predicted 850 μm luminosity error distribution, for quiescent and burst galaxies. Tab. 5.1 described the quantities displayed.

remarkable: we are able to reproduce more than 95% luminosities with an accuracy smaller than 10%, either quiescent or burst galaxies. The success of the predictions is also reflected in the relative root mean square error, with 0.37 for quiescent galaxies and 0.07 for the burst sample. No clear correlation was found between the error and galaxy properties.

The luminosity function in the rest frame 850 μm at $z = 2$ is plotted in Fig. 5.19. Similar to what we found for the UV and mid-IR bands, the submillimeter luminosity functions calculated using the predicted luminosities are identical to the functions constructed from the original 850 μm luminosities (extracted from GRASIL spectra), for the whole luminosity range.

5.5 Discussion and Conclusions

In this Chapter we implemented a new method to predict galaxy spectrum from a few galaxy properties, using artificial neural networks. We were motivated by the cost of running GRASIL, in terms of computing time, to generate model spectra for all galaxies in the Millennium Simulation box. The artificial neural networks provide a fast, simple and flexible platform to calculate galaxy spectra. Here, we have carried out the first tests and predictions. The next Chapter will be devoted to applications of the method.

Neural networks show good performance when predicting galaxy spectra from galaxy properties. Using a simple supervised feed-forward net, composed of 14 input galaxy properties, two hidden layers with 30 neurons each, and 456 (the number of spectral flux bins) neurons in the output layer, we achieve an accuracy of 80% of galaxies with predicted bolometric luminosities within 10% of the true values.
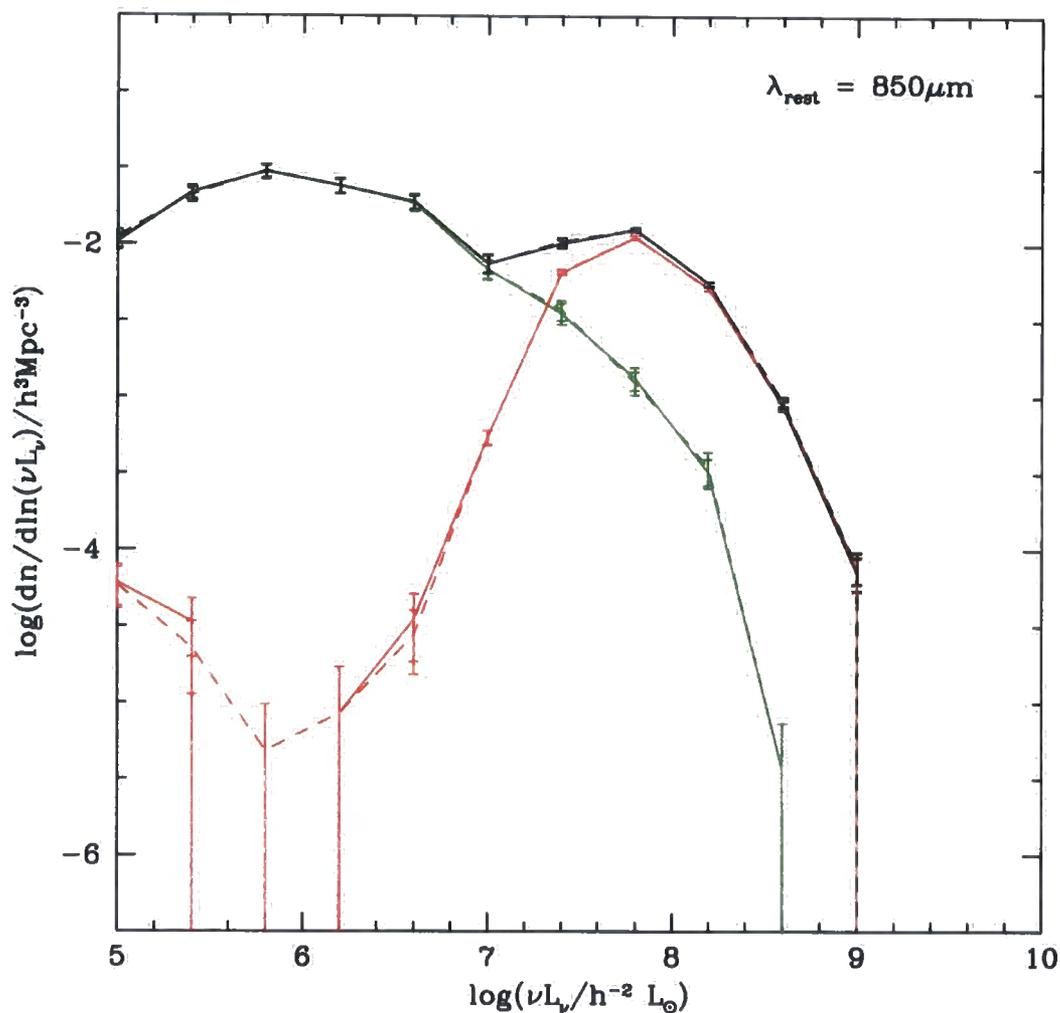
Figure 5.19: The luminosity function at 850 μm. The luminosity functions calculated from the original luminosities are shown by the solid lines, while the ANN predicted values are given by the dashed lines. The colour scheme for the different lines obeys to: the green and red lines show the LF for the quiescent and burst galaxies, respectively, which are the components of the full sample, represented by the black line. The error bars indicate the Poisson uncertainties due to the number of galaxies simulated.

A principal components analysis was used to reduce the dimensionality of spectra. This is done by calculating new basis-vectors, which represent directions of maximum variance; the projection of the spectra on these vectors generate the principal components. We can select as many principal components as we consider necessary (obviously limited by 456) to represent the spectra. Using 20 principal components as the output to be predicted by the ANN, leads to a performance of $\approx 76.4\%$ for $P_{|e|<10\%}$ for bolometric luminosities of galaxies at $z = 0$, but with more than an order of magnitude gain in speed.

The best performance is achieved with only one output neuron, set to the luminosity required. At $z = 0$, in the B, 24 μm and 850 μm we found an accuracy (as given by $P_{|e|<10\%}$) of 95.7%, 80.9% and 88.7%, respectively.

We also compared the luminosity functions constructed from the luminosities predicted by the ANN and the LFs extracted from the true luminosities, as given by GRASIL spectra. We showed that the ANN is capable of reproducing with great precision the luminosities of quiescent Lyman-break galaxies at $z = 3$, mid-IR at $z = 0.4$, and submillimeter galaxies at $z = 2$. For burst galaxies, the ANN reveals difficulties in reproducing the luminosities at the UV and IR bands, with only $\approx$ 50% of predicted luminosities within 10% of error. We proved, however, that this accuracy is sufficient when the sample is statistically significant, and introduces no changes in the shape of the LF. This is possible because there is statistical independence between the error, associated with the predicted luminosities, and the galaxy properties, including the true luminosities.

Chapter 6

Modelling the Dusty Universe II: The Clustering of Galaxies in the Submillimeter

6.1 Introduction

In the previous chapter, we described a new method to rapidly predict large numbers of spectra from a few galaxy properties. The method of choice was artificial neural networks, which we proved to be a good technique to predict galaxy spectra with great accuracy. In the submillimeter region of the spectrum, this new approach performs extremely well, both for quiescent and burst galaxies, predicting accurate luminosities (within 10% error) for more than 95% of galaxies.

This Chapter is devoted to using the model to predict UV–submillimeter colours and the clustering of the submillimeter galaxies, in the Millennium Simulation, at $z = 2$, which is approximately equal to the observed median redshift of this population (Chapman et al., 2000).

Submillimeter galaxies (SMGs) where firstly identified by Smail et al. (1997), using the *SCUBA* instrument. This population of luminous high-redshift galaxies emit most of their light in the submillimeter wavelengths, and are believed to be fundamental for a coherent view of the formation and evolution of galaxies. They have been proposed as the progenitors of massive elliptical galaxies observed in the local Universe today (Blain et al., 2002).

159

The main source of submillimeter radiation from galaxies is thermal emission from dust grains. The heating of the dust is thought to be due to one of two processes: UV light emitted by young massive stars (during episodes of intense star formation, starbursts), or energy released by the accretion of matter on to a central massive black hole. Recently, Alexander et al. (2005) found that 75% of their sample of submillimeter galaxies showed X-ray emission, powered by a central AGN. Nevertheless, these authors concluded that the starbursts are the dominant power responsible for the submillimeter emission in their sample. Determining the connection between the submillimeter galaxies and Lyman-break galaxies (whose luminosity is also dominated by their UV emission) is, therefore, an important goal. Observations of known Lyman-break galaxies at 850 μm have not been very successful (e.g. Chapman et al., 2000), with a very low detection rate (less than 10% of the observed Lyman-break galaxies showed some emission in the submillimeter). In this Chapter, we try to shed some light on this issue by analysing the UV–submillimeter colours and luminosity functions for a sample of galaxies drawn from the Millennium Simulation, i.e. a sample of model galaxies predicted by GALFORM, where the halo merger trees used by this galaxy formation model were extracted from the Millennium Simulation. We should mentioned that, in the current version of the GALFORM–GRASIL code, we do not model the radiation emitted by AGN, hence all the results presented in this Chapter assume that all the submillimeter emission is the result of the reprocessing of the UV radiation (emitted by young massive stars) by dust.

The clustering of galaxies encodes important information about a population. It is driven by three factors: the clustering of the underlying dark matter, the distribution of galaxies between the haloes and the distribution of galaxies within the dark matter haloes. The spatial distribution of submillimeter galaxies is important in the sense that, due to their high redshifts, they can impose key constraints on theoretical models of galaxy formation.

Throughout this Chapter we use a 14:30:30:1 ANN configuration, with 14 galaxy properties in the input layer, two hidden layers with 30 neurons each, and one neuron in the output layer set as the luminosity required (see previous Chapter for further

details). We use the variant of the GALFORM model described in Baugh et al. (2005), set in the context of N-body halo merger trees (see Chapter 2).

In Section 6.2 we present the predictions for the UV–submillimeter colours of the galaxies in the Millennium Simulation, while in the clustering of this population is derived in Section 6.3. Conclusions are presented in Section 6.4.

6.2 UV–Submillimeter Luminosity Functions and Colour

The quest for finding the power source that is driving the submillimeter emission in galaxies has just begun. As mentioned in the previous section, recent findings show that most of the submillimeter emission is due to the heating of dust by UV light emitted by young stars, and only a small percentage is fueled by AGN activity Alexander et al. (2005). In this section, we try to shed some light on this problem by predicting the fraction of SMGs with UV emission and showing the UV–submillimeter colour distribution of the galaxies in the Millennium Simulation, using luminosities predicted by an artificial neural network. We need to bear in mind, however, that the current version of our model does not follow the heating of the dust by an AGN, so all the submillimeter radiation is emitted by the reprocessing of UV light by dust.

Firstly, in Fig. 6.1 we plot the luminosity function in the observer frame 850 μm and in the rest frame UV ($\lambda = 0.2 \mu$m) magnitudes; in fact, this later quantity corresponds to the observer frame R band at $z = 2$, which is the quantity plotted. As expected, the figure shows that bright galaxies are rarer than fainter galaxies: submillimeter galaxies with fluxes densities $S_\nu \geq 1$ mJy are ~ 20 times more abundant than galaxies with $S_\nu \geq 5$ mJy. For these brighter galaxies we calculate a space density of 5.2×10^{-5} Mpc^{-3}, which is in reasonable agreement with the value determined by Blain et al. (2004) of 2.7×10^{-5} Mpc^{-3}.

The bottom panel of Fig. 6.1 reveals that, in our model, a substantial fraction of submillimeter galaxies have detectable rest frame UV luminosities, fainter than $M_{\rm AB} - 5 \log h \gtrsim -22$ mag ($R \gtrsim 23.3$ mag). Chapman et al. (2005), using a sample
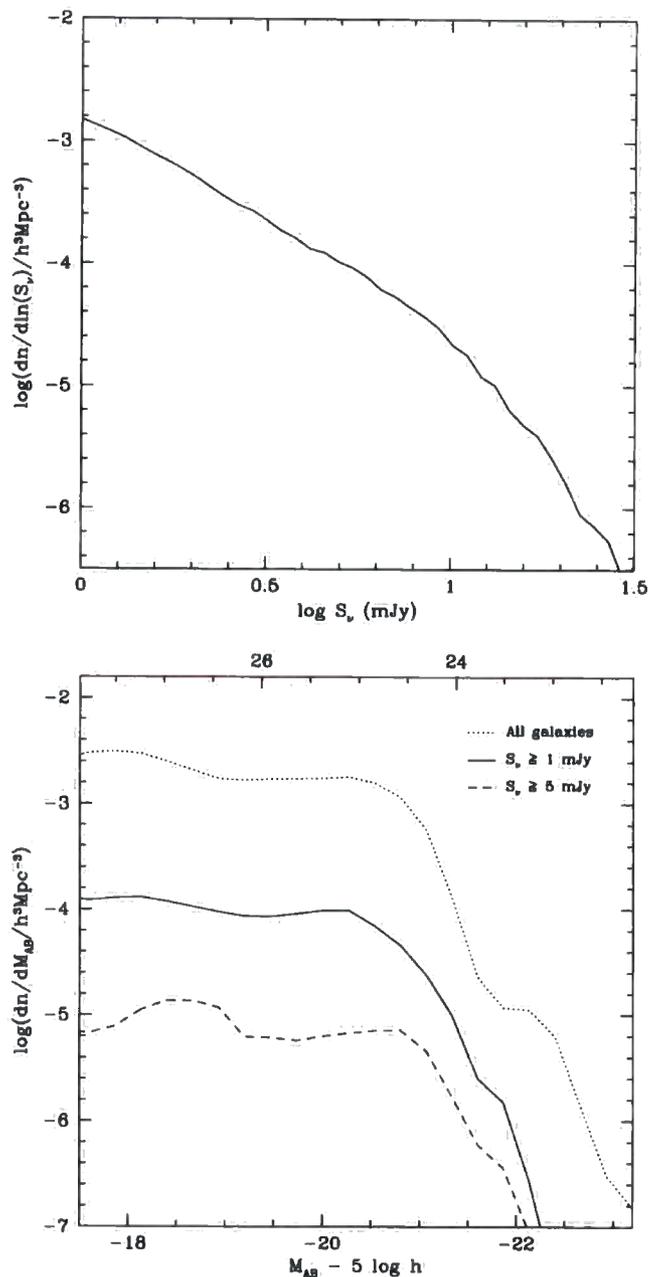
Figure 6.1: The luminosity function in the observer frame 850 μm (upper panel) and R band (bottom panel), which corresponds to the rest frame UV ($\lambda = 0.2\,\mu$m), at $z = 2$. In the lower panel, we distinguish between the total UV luminosity function (dotted line), and the distribution for submillimeter galaxies with $S_\nu \geq 1$ mJy (solid line) and $S_\nu \geq 5$ mJy (dashed line). Here, we show the apparent R band magnitude on the top axis.

Sample	All galaxies	$M_{AB} - 5 \log h \leq -19$ mag
	(10^{-5} Mpc^{-3})	(10^{-5} Mpc^{-3})
Galaxies with $S_{\nu}(850\,\mu m) \geq 1$ mJy	101.1	16.3
Galaxies with $S_{\nu}(850\,\mu m) \geq 5$ mJy	5.2	1.4

Table 6.1: The space density of submillimeter galaxies in the Millennium Simulation box at $z = 2$. We distinguish between galaxies with $S_{\nu}(850\,\mu m) \geq 1$ mJy and brighter galaxies, $S_{\nu}(850\,\mu m) \geq 5$ mJy. In the second column, we further limit our sample be only considering those galaxies with UV magnitudes brighter than -19 mag ($R < 26.3$ mag). The number density in the table are quoted in units of 10^{-5} Mpc^{-3}.

of 73 submillimeter galaxies with a median flux density of 5.7 mJy at a median redshift of 2.2, observed a median magnitude around $M_{AB} - 5 \log h \approx -21$ mag; their values are within our range of predicted magnitudes. This panel also reveals another important characteristic: only $\approx 5\%$ of all the galaxies with $M_{AB} - 5 \log h \leq -19$ mag ($R \leq 26.3$ mag), show submillimeter flux densities brighter than 5 mJy. In other words, we predict that roughly 27% of the submillimeter population with $S_{\nu} \geq 5$ mJy have UV magnitudes $M_{AB} - 5 \log h \leq -19$ mag. This explains the observational difficulty in finding the UV counterparts of submillimeter galaxies (e.g. Chapman et al., 2000). The integrated number densities of submillimeter galaxies is listed in Table 6.1.

We now look in more detail at the performance of the ANN in predicting the UV–submillimeter colours of a sample of galaxies modelled by GALFORM. We define the UV–submillimeter colour as L_{UV}/L_{subm}, where $L_{UV} \equiv \lambda L_{\lambda}$ with the UV luminosity given by the observer frame R band at $z = 2$ (this corresponds to a rest frame wavelength of $\approx 0.2\,\mu m$), and the submillimeter luminosity by the observer frame 850 μm. We calculated the predicted colours training the ANN independently for each band, using UV and submillimeter luminosities. The errors associated with the predicted UV luminosities were presented in Tab. 5.10, while for the observer frame submillimeter band, we find errors similar to those found for the rest frame 850 μm (Tab. 5.12).

Fig. 6.2 shows the performance of the ANN, when combining luminosities to predict colours. Here we compare the predicted and the true (expected) colours for the ANN test sample, at $z = 2$, for quiescent and burst galaxies. The plot shows that, for most of the galaxies, the predicted colours are within 10% of their original values: the first and third quartiles of the error distribution are -5.7% and 5.8%, respectively. Also, we find no correlation between the errors associated with the predicted UV (observer frame R band) and the submillimeter luminosities.

In Fig. 6.3 we plot the distribution of UV–submillimeter colours predicted by the ANN, for the sample extracted from the Millennium Simulation box at $z = 2$. Here, we plot the colour distributions for submillimeter galaxies with flux densities $S_\nu(850\,\mu\text{m}) \geq 1$ and ≥ 5 mJy. We further constrain the plotted data by only selecting galaxies with UV magnitudes brighter than -19 mag ($R < 26.3$ mag). We predict that fainter submillimeter galaxies show a UV–submillimeter colour within the range $10^{-4.5}$–$10^{-2.5}$, with a median colour of $\approx 4 \times 10^{-4}$. SMGs with flux densities $S_\nu(850\,\mu\text{m}) \geq 5$ mJy display redder colours by a factor of ~ 3, with a median of 1.3×10^{-4}, which is in excellent agreement with the values calculated from the median values found by Chapman et al. (2005) for observer frame R band magnitudes and submillimeter luminosities, around 1.4×10^{-4}. This characteristic is a direct consequence of the limit in the UV–magnitude, as seen in the lower panel of the Fig. 6.3 where we plot the predicted colour–magnitude relation for submillimeter galaxies brighter than 1 mJy.

6.3 The Clustering of Submillimeter Galaxies

The characterization of a galaxy population is never complete without knowledge of its clustering. Within the framework of galaxy formation models, comparison with the observed spatial distribution of galaxies can help to constrain model parameters. As mentioned previously in this Chapter, the amplitude of the two-point correlation function is defined by three factors: the spatial distribution of the underlying dark matter, the partitioning of galaxies between the different dark matter haloes, and the distribution of galaxies inside halo. On large scales, the correlation function of
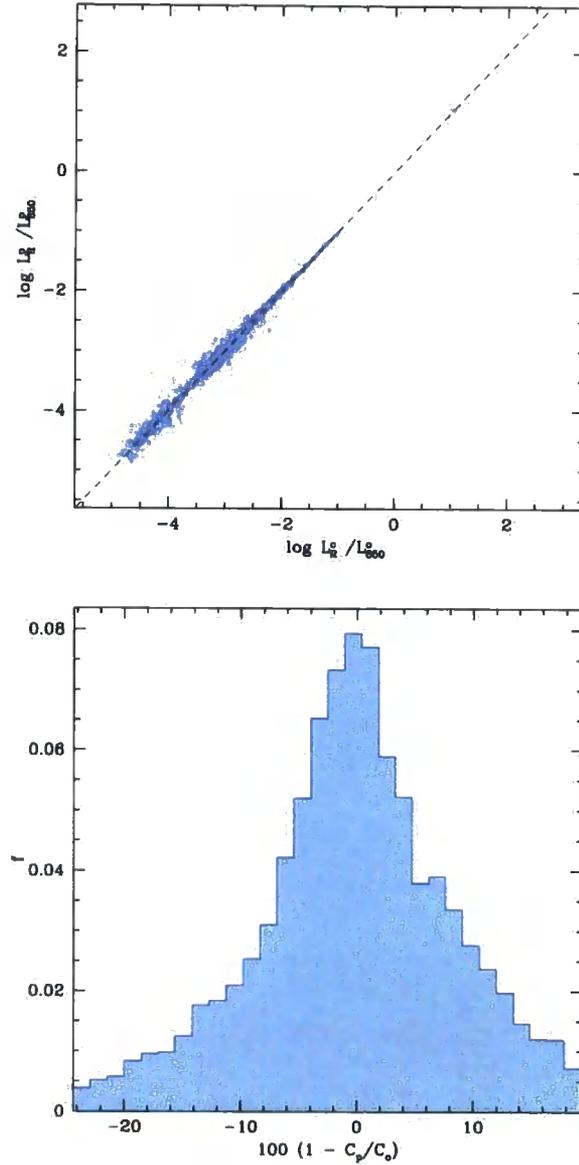
Figure 6.2: The errors associated with the colours, as given by the ratio of luminosities predicted by the ANN. The top panel compares the predicted, $L_{\mathrm{R}}^{\mathrm{p}}/L_{850}^{\mathrm{p}}$, with the true, expected, $L_{\mathrm{R}}^{\mathrm{o}}/L_{850}^{\mathrm{o}}$, colours for a complete (i.e. quiescent+burst galaxies) sample of GALFORM galaxies. As explained in the text, the observer frame R band corresponds to the UV wavelength 0.2 μm, for galaxies at $z = 2$. The error distribution is plotted in the bottom panel. Here, $C_{\mathrm{p}} = L_{\mathrm{R}}^{\mathrm{p}}/L_{850}^{\mathrm{p}}$ and $C_{\mathrm{o}} = L_{\mathrm{R}}^{\mathrm{o}}/L_{850}^{\mathrm{o}}$. We set the range to be equal to the 5th and 95th percentiles of the distribution, and the red dotted lines mark the first and third quartiles. The histogram is normalized such than $\sum_i f_i = 1$.
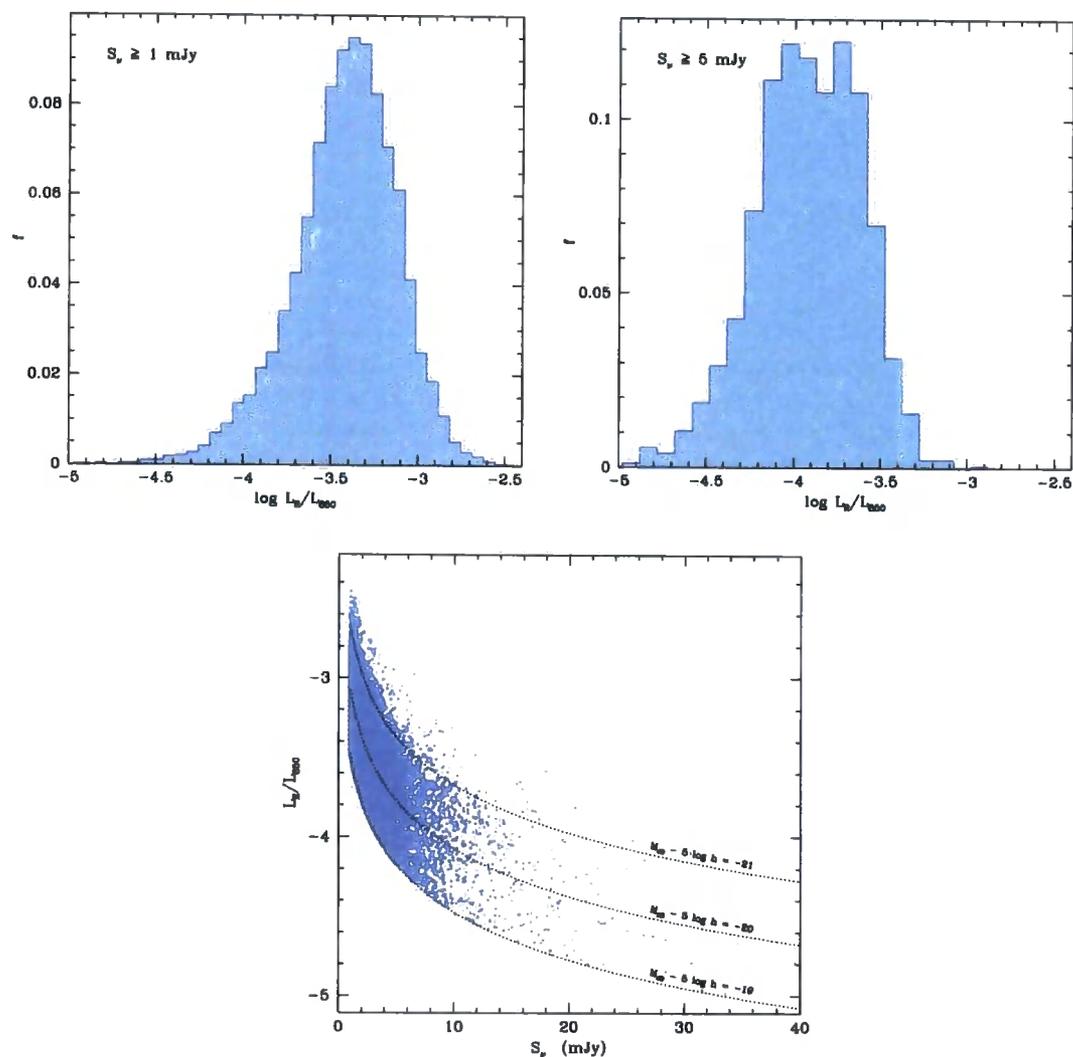
Figure 6.3: The UV–submillimeter colours predicted by the ANN, for a sample of galaxies extracted from the Millennium Simulation at $z = 2$. The top panels show the distribution of galaxies with fluxes densities, $S_\nu(850\,\mu m)$, brighter than 1 and 5 mJy, top-left and top-right respectively. Here, we further constrain the plotted data to have $M_{AB} - 5\log h \leq -19$ mag. The colour–magnitude diagram for the population of galaxies with $S_\nu(850\,\mu m) \geq 1$ mJy is plotted in the lower panel. The dotted lines show curves of constant UV magnitude.

galaxies is, essentially, shaped by the distribution of matter in the Universe. On scales smaller than the typical size of the host dark matter haloes, the function is dominated by the distribution of galaxies within the halo.

Using a sample of 73 observed submillimeter galaxies at $z \simeq 2-3$ with a median 850 μm flux of 5.7 mJy, Blain et al. (2004) determined a correlation length of $\sim$ $6.9 \pm 2.1 \, h^{-1}$ Mpc, which is greater than the separation lengths observed for Lyman-break galaxies and QSO (*quasi-stellar objects*).

In this section, we determine the two-point correlation function for submillimeter galaxies at $z = 2$ in the Millennium Simulation, differentiating between galaxies with submillimeter fluxes $S_\nu \geq 1$ mJy and $S_\nu \geq 5$ mJy.

Firstly, we analyse how the errors associated with the 850 μm luminosities predicted by the neural network affect the determination of the clustering. As noted in Chapter 2, the shortcut calculation gives a reasonable estimate of the 850 μm fluxes and can be used to populate the full Millenium volume quickly. Fig. 6.5 shows a comparison between the 850 μm flux densities computed using the shortcut and the fluxes predicted by the GRASIL code, for a "normal" sample of galaxies extracted at $z = 2$. For most of the galaxies the shortcut calculation is accurate within a factor of ~ 2, with shortcut fluxes being systematically higher than the GRASIL values for brighter submillimeter galaxies. For example, galaxies with $S_\nu \geq 1$ mJy display shortcut flux densities $\approx 43\%$ brighter than the GRASIL code calculations. We plot the correlation function for all the submillimeter galaxies in the Millennium Simulation with fluxes brighter than 1 mJy at $z = 2$. The two-point correlation function derived from the shortcut luminosities is represented in Fig. 6.4 by the black curve. We study the effect of the errors associated with the luminosities predicted by the ANN on the correlation function, by randomly perturbing the shortcut luminosities by the distribution of errors expected for the 850 μm band (blue line in Fig. 5.18). We find no noticeable change in the two-point correlation function when the perturbed luminosities are used instead (the differences between both correlation functions are contained within 3% error), indicating that the clustering is a robust prediction.

In Fig. 6.4 we also show the two-point correlation function calculated from the luminosities predicted by the artificial neural networks (red curve). We find a good
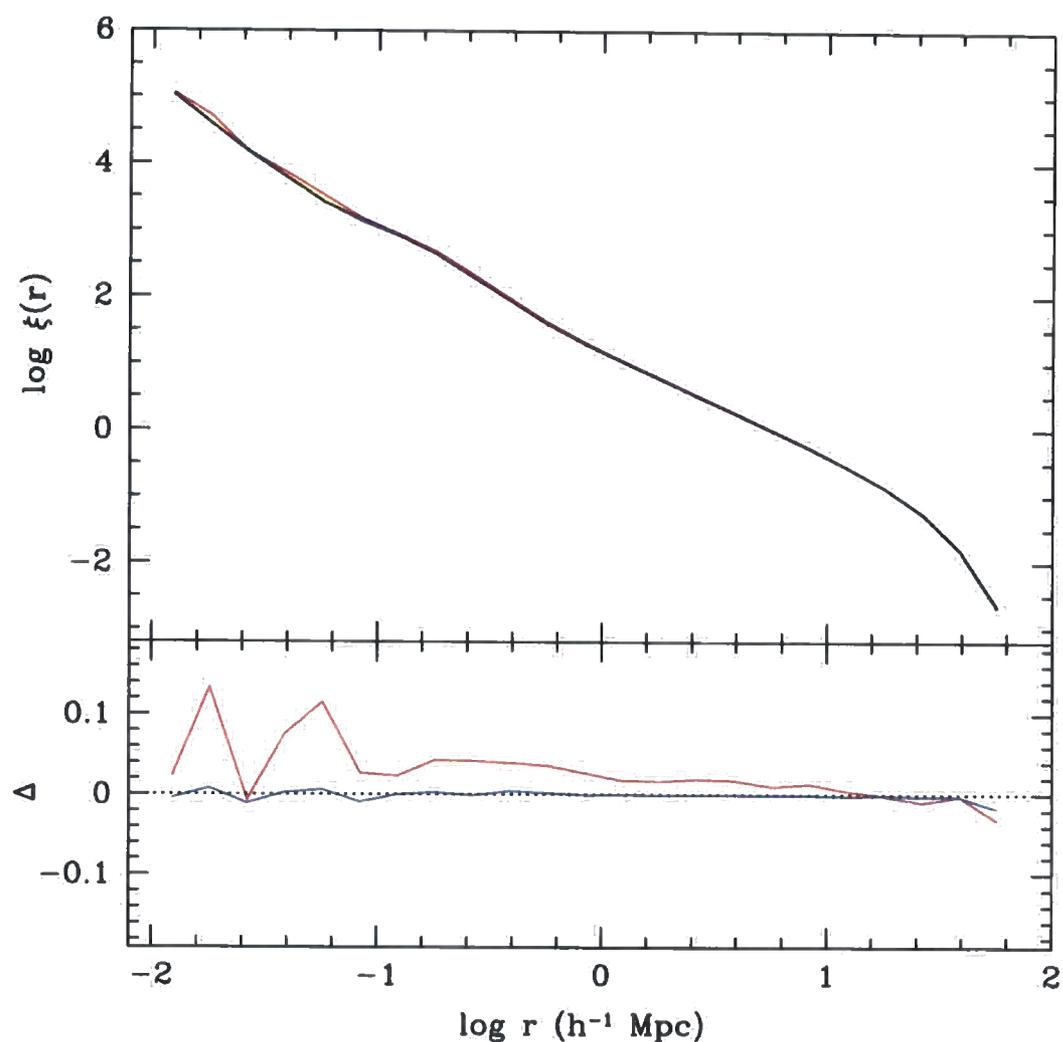
Figure 6.4: The impact of luminosity errors on the real-space two-point correlation function of submillimeter galaxies, with fluxes greater than 1 mJy and at a redshift $z = 2$. The black curve shows the correlation function obtained when using GALFORM shortcut luminosities, while the red line represents the prediction by the neural networks. The blue curve shows the effect of adding a random error to the luminosities predicted by GALFORM, using an error distribution identical to the errors expected in the luminosities predicted by the ANN. In the bottom panel, we plot the ratio between the different curves and the two-point correlation function using the shortcut, $\Delta = \log \xi - \log \xi_{shortcut}$.

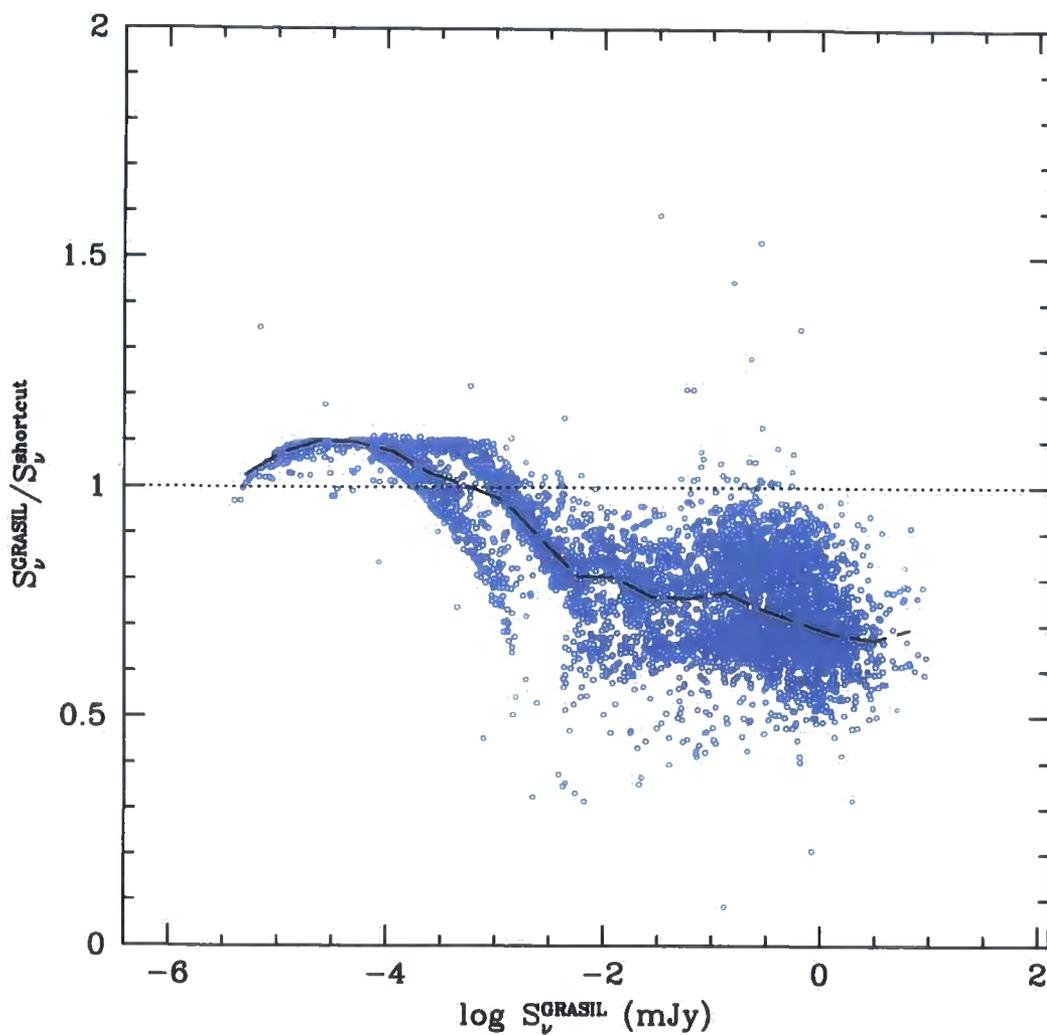Figure 6.5: Comparison between the 850 μm flux densities computed using the shortcut and predicted by the GRASIL code, for a representative "normal" sample at $z = 2$. The dashed line shows the median of the distribution.

agreement between both correlation functions: for most pair separations, the functions do not differ by more than 10%. The only discrepancy is for separations below $100\,h^{-1}$ kpc, where the two-point correlation function given by the ANN-predicted luminosities is ≈ 1.25 times higher than the shortcut calculation. However, this is not a consequence of the errors associated with the new neural networks technique, it is due to the fact that the shortcut luminosities are systematically brighter than GRASIL luminosities for brighter galaxies. Therefore, the clustering of galaxies in the shortcut sample is reduced due to the (erroneous) inclusion of galaxies that GRASIL would predict to be fainter than the flux limit.

Fig. 6.6 shows the predicted two-point correlation function for SMGs in the Millennium Simulation in real and redshift space; here we also differentiate between galaxies with fluxes $S_\nu \geq 1$ mJy and a brighter subsample with $S_\nu \geq 5$ mJy. In real space, the correlation function is calculated using the Cartesian coordinates of the submillimeter galaxies that populate the simulation box. We predict that the correlation function of the SMGs is close to a power law, $\xi(r) = (r/r_0)^\gamma$, over more than three decades in separation. For galaxies with fluxes $S_\nu \geq 5$ mJy, we fit a slope of $\gamma = -1.94 \pm 0.05$ and a correlation length of $r_0 = 6.2 \pm 0.2\,h^{-1}$ Mpc. These values are slightly lower than the values inferred by Blain et al. (2004), but within their errorbars. For fainter galaxies, $S_\nu \geq 1$ mJy, we find an identical slope to that of brighter galaxies, $\gamma = -1.89 \pm 0.04$, and $r_0 = 4.7 \pm 0.2\,h^{-1}$ Mpc, which is 1.3 times smaller than that of galaxies with fluxes $S_\nu \geq 5$ mJy. The increase in the clustering with the luminosity is a consequence of the fact that brighter submillimeter galaxies are found in more massive haloes, which translates into a bigger effective bias and stronger clustering (see e.g. Mo & White, 1996). This is perhaps one of the reasons behind the high value determined observationally, for a flux limited sample, where observed submillimeter galaxies are expected to be affected by cosmic variance. A lower value for the correlation length favors the link between the SMGs and Lyman-break galaxies due to the lower values found for these, around $r_0 = 4\,h^{-1}$ Mpc (Porciani & Giavalisco, 2001), while QSOs are expected to have a correlation function $\sim 5\,h^{-1}$ Mpc (Croom et al., 2002). We remind the reader that, in our model, all the submillimeter radiation is emitted by UV-heated dust, so a similarity
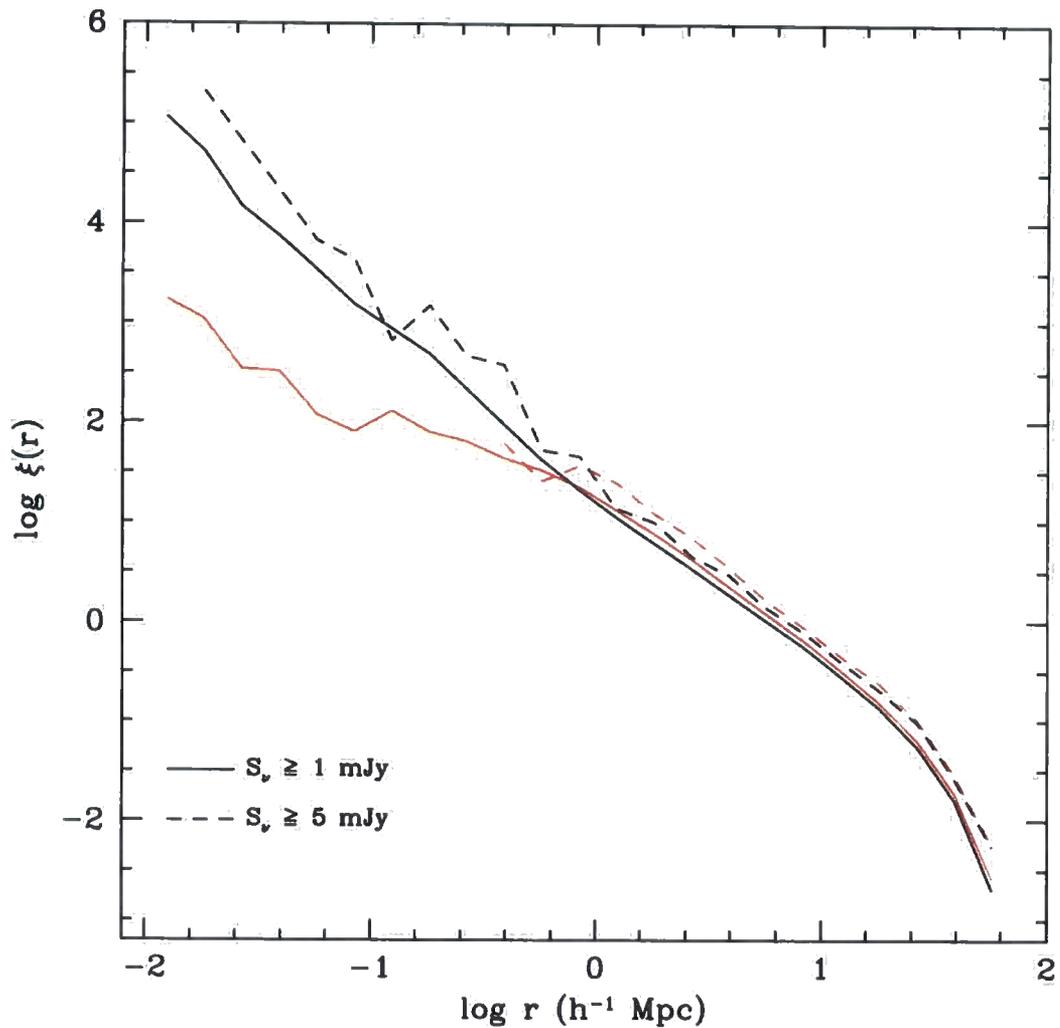
Figure 6.6: The real-space two-point correlation function of submillimeter galaxies. The solid lines shows the function for galaxies with fluxes $S_\nu \geq 1$ mJy, while the dashed lines represent a brighter subsample, $S_\nu \geq 5$ mJy. The real-space calculation is shown by the black lines, while red curves show the function in the redshift-space.

between the correlation function of both populations is expected.

The real-space correlation function on small scales reveals another important characteristic of the population of submillimeter galaxies: our model predicts more than one SMG per halo. If that was not the case, we would expect the correlation function to tend to $\xi \sim -1$ on scales smaller than the typical size of the host haloes (see Benson et al., 2000).

In the redshift space, the redshift positions are used instead of the Cartesian values. These are calculated by perturbing the position of the galaxy along one Cartesian axis using its peculiar velocity, which should give the redshift position of the galaxy as determined by a distant observer. On small scales, the redshift space correlation function falls below the real space function, while at larger scales, the redshift space becomes stronger than the real space. This behaviour is expected: for small pair separations, the peculiar motions of the galaxies within the halo cause an apparent stretching of the structure, which dilutes the number of galaxy pairs; on larger scales, there is an amplification of the correlation function due to the bulk motion of the SMGs (Kaiser, 1987), which for our population is ≈ 1.2.

6.4 Discussion and Conclusions

In this Chapter we use artificial neural networks to predict the submillimeter luminosities for all the galaxies in the Millennium Simulation box at $z = 2$. We used the method described in the previous Chapter, in which we start from a given set of predicted galaxy properties and, with the help of a properly trained ANN, we derive the 850 μm and UV luminosities.

In order to have a clear picture of the whole process of galaxy formation and evolution in the Universe, we need to understand the population of submillimeter galaxies. As we mentioned earlier, submillimeter emission occurs as a result of the heating of the galactic dust. Mainly, there are two processes responsible for the heating of the dust: UV light emitted by young massive stars (during episodes of intense star formation, starbursts), or energy released by the accretion of matter on to a central massive black hole. Nevertheless, recent observations found that

the starbursts are the dominant power responsible for the submillimeter emission (Alexander et al., 2005). We try to shed some light on this problem by predicting the abundance, UV–submillimeter colours and clustering of this population. However, because our model does not include heating of dust by AGNs, the predictions shown in this Chapter concern only the emission of submillimeter radiation by reprocessing of UV light by dust.

We predict that submillimeter galaxies with flux densities $S_\nu(850\,\mu\text{m}) \geq 1$ mJy have a space density of 10.1×10^{-4} Mpc^{-3}, and brighter SMGs $S_\nu(850\,\mu\text{m}) \geq 5$ mJy to be approximately 20 times rarer, which is good agreement with recent observations. These high values of space density reinforce the idea of this population to be one of the most abundant when the Universe was ~ 3.3 Gyr old. We also predict the rest frame UV ($\lambda = 0.2\,\mu\text{m}$) magnitude at $z = 2$, as given by the observer frame R band. We showed that a significant fraction of submillimeter galaxies (≥ 5 mJy) show measurable UV luminosities: approximately 27% of the SMGs have predicted ultraviolet magnitudes brighter than -19 mag ($R < 26.3$ mag). However, this 27% of galaxies represent only 5% of the galaxy population with UV magnitudes $M_{\text{AB}} - 5\log h \leq -19$ mag. Another successful prediction is that of UV–submillimeter colours. We find that most of the submillimeter galaxies have colours, as given by the observer frame ratio L_{R}/L_{850}, around $10^{-5} - 10^{-2.5}$. Galaxies with $S_\nu(850\,\mu\text{m}) \geq 1$ mJy display a median colour of 1.3×10^{-4}, while brighter galaxies ($S_\nu(850\,\mu\text{m}) \geq 5$ mJy) present redder colours by a factor of ~ 3. These values enclose recent observational estimates.

Another important prediction is the clustering of the submillimeter population. We found the correlation function to be a power law down to pair separations of $r \sim 10\,h^{-1}$ kpc. For galaxies brighter than $S_\nu \geq 5$ mJy, we fit a slope of $\gamma = 1.94 \pm 0.05$ and a correlation length of $r_0 = 6.2 \pm 0.2\,h^{-1}$ Mpc, which matches the observations. When fainter galaxies are included in the sample, $S_\nu \geq 1$ mJy, the correlation length decreases by a factor of 1.3, in agreement with the correlation between luminosity, halo mass and clustering amplitude.

In summary, the agreement between the model predictions and the observations favors the connection between submillimeter galaxies and UV emission, where most

of the submillimeter radiation comes from UV light reprocessed by dust. We showed that the correlation function of this population is sensitive to the flux limits of the observational sample, with brighter galaxies exhibiting higher clustering; we hint that this might be the reason behind the high values for the correlation length found observationally.

Chapter 7

Conclusions and Future Work

In this Thesis we have studied early-type galaxies in the framework of hierarchical structure formation. Our approach can be divided into two areas: testing the current galaxy formation model, GALFORM, and making new predictions to help in future observations. We used the published semi-analytical galaxy formation models of Baugh et al. (2005) and Bower et al. (2006). There are a number of important differences between the two galaxy formation models. In summary, to suppress the overproduction of bright galaxies, the Bower et al. model employs AGN heating to stifle gas cooling, whereas the Baugh et al. model invokes a superwind which ejects cooled gas. Also, in the Baugh et al. model, a top-heavy stellar initial mass function is adopted in starbursts.

In the local Universe, we compared the model predictions with observational results derived from the Sloan Digital Sky Survey, in Chapter 3. We tested the prescription outlined by Cole et al. for computing the radii of discs and bulges. We showed that the model enjoys a number of notable successes, such as giving reasonable reproductions of the local Faber-Jackson relation (velocity dispersion–luminosity), the velocity dispersion–age relation, and the Fundamental Plane relating the luminosity, velocity dispersion and effective radius of spheroids. These achievements are all the more remarkable when one bears in mind that none of the parameters were adjusted to refine the model predictions. We studied how the residuals around the fundamental plane relation depend on galaxy properties. We examined, in detail, the physical ingredients of the calculation of galaxy sizes in GALFORM, showing which components have the most influence over our results. We also calculated the evolution of the scaling relations with redshift, finding a rea-

sonable agreement overall. However, in spite of the successes, we found important disagreements between the predictions of the model and observations: the brightest model spheroids have effective radii smaller than observed and the zero-point of the fundamental plane shows little or no evolution with redshift in the model. The solution of these remaining problems will require enhancements of the model to compute galaxy sizes, guided by the results of numerical simulations of the growth of disc galaxies and galaxy mergers (e.g. Okamoto et al., 2005; Robertson et al., 2006).

Luminous red galaxies (LRGs) are much rarer and more massive than L_* galaxies. Coupled with their extreme colours, LRGs therefore provide a demanding testing ground for the physics of massive galaxy formation. In Chapter 4, we presented the first self-consistent predictions for the abundance and properties of LRGs in hierarchical structure formation models. Without adjusting any parameters, the Bower et al. model showed an excellent match to the observed luminosity function of LRGs in the Sloan Digital Sky Survey (with a median redshift of $z = 0.24$) and to their clustering. The Baugh et al. model was less successful in these respects. Both models failed to match the observed abundance of LRGs at $z = 0.5$ to better than a factor of ≈ 2. In the models, LRGs are typically bulge dominated systems with stellar masses of $\approx 2 \times 10^{11} h^{-1} M_{\odot}$ and velocity dispersions of $\sigma \sim 250 \mathrm{kms}^{-1}$. Around half of the stellar mass in the model LRGs has already formed by $z \sim 2.2$ and is assembled into one main progenitor by $z \sim 1.5$; on average, only 25% of the mass of the main progenitor is added after $z \sim 1$. LRGs are predicted to be found in a wide range of halo masses, a conclusion which relies on properly taking into account the scatter in the formation histories of haloes. Remarkably, we found that the correlation function of LRGs is predicted to be a power law down to small pair separations, in excellent agreement with observational estimates. As found for early-type galaxies at $z = 0$, neither the Bower et al. nor the Baugh et al. model were able to reproduce the observed radii of LRGs.

The Millennium Simulation opened a new world of possibilities for hierarchical models of galaxy formation. It is now possible to extract halo merger histories directly from N-body simulations (for the Millennium Simulation) and utilize then in GALFORM. By doing so, we will have additional information about the spatial

distribution of galaxies and hence can connect galaxy properties and evolution with environment. In order to predict the structural and sphectrophotometric properties for the galaxies populating the Millennium Simulation, and make mock catalogues at wavelengths where the basic GALFORM model predictions are less reliable, we need to use the GRASIL code. GRASIL is a state of the art sphectrophotometric model, which includes a comprehensive and self-consistent calculation of the radiative transfer of light through dust. Driven by the large computational time associated with the running of GRASIL code, in Chapter 5 we implement a new method, based on artificial neural networks, to predict a galaxy spectrum given a small set of properties. Overall, the ANN performance is noteworthy: approximately 80% of spectra show predicted bolometric luminosities within 10% from the true values. We proved that predicting luminosities in a given band is more accurate than trying to predict the entire spectral range, with an improvement of $\approx$ 30%: at $z = 0$, we predict 96%, 81% and 89% of galaxies with B, 24 μm and 850 μm luminosities within 10% of the true values. These errors were found not be important when constructing the luminosity function mainly because the errors are not dependent on any galaxy property. In Chapter 6 we apply this new method to predict the abundance, colours and clustering of the population of submillimeter galaxies at $z = 2$. These galaxies are thought to be the progenitors of massive elliptical galaxies observed in the local Universe today. At this redshift, we found submillimeter galaxies with $S_\nu(850\,\mu\text{m}) \geq$ 5 mJy to have space densities around 5.2×10^{-5} Mpc^{-3}, which is in good agreement with observations. Only one third of this population of SMGs show measurable UV emission, i.e. UV magnitudes brighter than -19 mag ($R < 26.3$ mag). Another success of the ANN+GALFORM+GRASIL code is that of UV–submillimeter colours: we predict the colours to be distributed around the median value of 1.3×10^{-4}, which is an excellent match to the observed values. Just like for the luminous red galaxies, we found that the correlation function for the submillimeter galaxies is well described by a power law in the interval $0.01 \lesssim r \lesssim 25\,h^{-1}$ Mpc. For the population of SMGs with flux densities brighter than 5 mJy we determined a correlation length of $r_0 = 6.2\,h^{-1}$ Mpc, which matches the observational estimates. All these successes between the model and observational data hint that most of

the submillimeter radiation in galaxies is due to the reprocessing by dust of UV radiation, produced by young stars.

At the moment, there are two concurrent variants of the GALFORM model, Baugh et al. (2005) and Bower et al. (2006), which have successfully reproduced a wide range of observational data sets; an important future advance would be to merge these two versions of the code. Deeper and wider observations of the Universe are nowadays scratching the early stages of galaxy evolution, the birth of the first stars, creating huge datasets over the whole range of wavelengths. These statistics can, and definitely will, constrain the hierarchical models of galaxy formation and evolution. In this Thesis, we demonstrated that using artificial neural networks associated with the GALFORM+GRASIL code we can now predict with impressive accuracy the luminosities of galaxies in any given band. This created a powerful tool which will enable us to explore and predict the formation and evolution of the Universe, and permits a direct comparison with the forthcoming surveys. For example, an interesting future application will be to create mock catalogues for the *Herschel* satellite to be launched at the end of 2008, which will study the Universe in the far infrared and submillimeter wavelengths (60 - 670 μm).

BIBLIOGRAPHY

Abazajian K., et al., 2005, ApJ, 625, 613

Adelberger K.L., Steidel C.C., Shapley A.E., Pettini M., 2003, ApJ, 584, 45

Alexander D.M., et al., 2003, MNRAS, AJ, 125, 383

Alexander D.M., et al., 2005, ApJ, 632, 736

Almeida C., Baugh C.M., Lacey C.G., 2007, MNRAS, 376, 1711

Almeida C., Baugh C.M., Wake D.A., Lacey C.G., Benson A.J., Bower R.G., Pimbblet K., 2008, MNRAS, 386, 2145

Angulo R.E., Baugh C.M., Lacey C.G., 2008, MNRAS, accepted

Baade W, Gaposchkin C.H.P., 1963, Evolution of Stars and Galaxies, Harvard University Press, Cambridge, MA

Babbedge T.S.R., et al., 2006, MNRAS, 370, 1159

Barber T., Meiksin A., Murphy T., 2007, MNRAS, 377, 787

Barger A.J., Aragón-Salamanca A., Ellis R.S., Couch W.J., Smail I., Sharples R.M., 1996, MNRAS, 279, 1

Barnes J., White S.D.M., 1984, MNRAS, 211, 753

Bauer A.E., Drory N., Hill G.J., Feulner G., 2005, ApJ, 621, L89

Baugh C.M., Cole S., Frenk C.S., 1996, MNRAS, 283, 1361

Baugh C.M., Benson A.J., Cole S., Frenk C.S., Lacey C.G., 1999, MNRAS, 305, L21

179

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

Baugh C.M., 2006, Rep. Prog. Phys., 69, 3101

Bell E.F., et al., 2004, ApJ, 608, 752

Bender R., Burstein D., Faber S.M., 1992, ApJ, 399, 462

Benson A.J., Baugh C.M., Cole S., Frenk C.S., Lacey C.G., 2000, MNRAS, 316, 107

Benson A.J., Lacey C.G., Baugh C.M., Cole S., Frenk C.S., 2002, MNRAS, 333, 156

Benson A.J., Bower R.G., Frenk C.S., Lacey C.G., Baugh C.M., Cole S., 2003, ApJ, 599, 38

Berlind A.A., et al., 2003, ApJ, 593, 1

Bernardi M., Sheth R.K., Annis J., et al., 2003a, AJ, 185, 1817

Bernardi M., Sheth R.K., Annis J., et al., 2003b, AJ, 185, 1849

Bernardi M., Sheth R.K., Annis J., et al., 2003c, AJ, 185, 1866

Bernardi M., Sheth R.K., Annis J., et al., 2003d, AJ, 185, 1882

Bernardi M., Sheth R.K., Nichol R.C., Schneider D.P., Brinkmann J., 2005, AJ, 129, 61

Bertin G., Ciotti L., Del Principe M., 2002, A&A, 386, 149

Bett P., Eke V., Frenk C.S., Jenkins A., Helly J., Navarro J., 2007, MNRAS, 376, 215

Binney J., Tremaine S., 1987, Galactic Dynamics, Princeton Univ Press, Princeton, New Jersey

Blain A.W., Smail I., Ivison R.J., Kneib J.-P., Frayer D.T., 2002, PhR, 369, 111

Blain A.W., Chapman S.C., Smail I., Ivison R., 2004, ApJ, 611, 725

Blake C., Collister A., Lahav O., 2008, MNRAS, in press (arXiv:0704.3377)

Blumenthal G.R., Faber S.M., Flores R., Primack J.R., 1986, ApJ, 301, 27

Bond J.R., Efstathiou G., Silk J., 1980, ApJ, 379, 440

Bond J.R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bournaud F., Jog C. J., Combes F., 2005, A&A, 437, 69

Bouwens R.J., Illingworth G.D., Blakeslee J.P., Broadhurst T.J., Franx M., 2004, ApJL, 611, 1

Bower R.G., 1991, MNRAS, 248, 332

Bower R.G., Benson A.J., Lacey C.G., Baugh C.M., Cole S., Frenk C.S., 2001, MNRAS, 325, 497

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

Boylan-Kolchin M., Ma C., Quataert E., 2006, MNRAS, 369, 1081

Bressan A., Granato G.L., Silva L., 1998, A&A, 332, 135

Bressan A., Silva L., Granato G.L., 2002, A&A, 392, 377

Brown M.J.I., et al., 2007, ApJ, 654, 858

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bundy K., Ellis R.S., Conselice C.J., 2005, ApJ, 625, 621

Cannon R., et al., 2006, MNRAS, 372, 425

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Chandrasekhar S., ApJ, 98, 54

Chapman S.C., et al., 2000, MNRAS, 319, 318

Chapman S.C., Blain A.W., Smail I., 2005, ApJ, 622, 772

Choi J., Lu Y., Mo H.J., Weinberg M.D., 2006, MNRAS, 372, 1869

Ciotti L., Lanzoni B., Renzini A., 1996, MNRAS, 282, 1

Cole S., 1991, ApJ, 367, 45

Cole S., Aragón-Salamanca A., Frenk C.S., Navarro J.F., Zepf S.E., 1994, MNRAS, 271, 781

Cole S., Lacey C.G., 1996, MNRAS, 281, 716

Cole S., Lacey C.G., Baugh C.M., Frenk C.S., 2000, MNRAS, 319, 168

Cole S., Helly J., Frenk C.S., Parkinson H., 2008, MNRAS, 383, 546

Colless et al., 2001, MNRAS, 328, 1039

Connolly A.J., Szalay A.S., 1999, ApJ, 117, 2052

Conroy C., Ho S., White M., 2007, MNRAS, 379, 1491

Cooray A., Sheth R., 2002, PhR, 372, 1

Croom S.M., et al., 2002, MNRAS, 335, 459

Croton D.J., et al., 2006, MNRAS, 365, 11

Davies R.L., Efstathiou G., Fall S.M., Illingworth G., Schechter P.L., 1983, ApJ, 266, 41

De Jong R.S., Lacey C.G., 2000, ApJ, 545, 781

De Lucia G., Springel V., White S.D.M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Dekel A., Cox T.J., 2006, MNRAS, 370, 1445

di Serego Alighieri S., et al., 2005, A&A, 442, 125

Djorgovski S., Davis M., 1987, ApJ, 313, 59

Dressler A., 1980, ApJ, 236, 351

Dressler A., Lynden-Bell D., Burstein D., Davies R.L., Faber S.M., Terlevich R.J., Jackson R., 1987, ApJ, 313, 42

Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069

Eggen O.J., Lynden-Bell D., Sandage A.R., 1962, ApJ, 136, 748

Eisenstein D.J., et al., 2001, AJ, 122, 2267

Eisenstein D.J., et al., 2005, ApJ, 633, 560

Eke V., Cole S., Frenk C.S., 1996, MNRAS, 282, 263

Faber S.M., Jackson R., 1976, ApJ, 204, 668

Faber S.M., Gallagher J.S, 1979, ARA&A, 17, 135

Faber S.M., et al., 2007, ApJ, 665, 265

Fabian A.C., Voigt L.M., Morris R.G., 2002, MNRAS, 335, 71

Ferrara A., Bianchi S., Cimatti A., Giovanardi C., 1999, ApJS, 123, 423

Font et al., 2008, MNRAS, submitted

Forbes D.A., Ponman T.J., 1999, MNRAS, 309, 623

Forbes D.A., Ponman T.J., Brown R.J.N., 1998, ApJ, 508, L43

Frenk C.S., White S.D.M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Frenk C.S., et al., 1999, ApJ, 525, 554

Fukugita M., Ichikawa T., Gunn J.E., Doi M., Shimasaku K., Schneider D.P., 1996, AJ, 111, 1748

Gallazzi A., Charlot S., Brinchmann J., White S.D.M., 2006, MNRAS, 370, 1106

Gamow G., 1956, The physics of the expanding Universe, Vistas in Astronomy, 2, 1726

Gao L., Springel V., White S.D.M., MNRAS, 2005, 363, 66

Gao L., White S.D.M., MNRAS, 2007, 377, 5

Gebhardt K., et al., 2003, ApJ, 597, 239

Governato F., Babul A., Quinn T., Tozzi P., Baugh C.M., Katz N., Lake G., 1999, MNRAS, 307, 949

Graham A. W., Driver S. P., Petrosian V., Conselice C. J., Bershady M. A., Crawford S. M., Goto T., 2005, AJ, 130, 1535

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

Guth A.H., 1981, Physical Review D, 23, 347

Haiman Z., Spaans M., Quataert E., 2000, ApJL, 537, 5

Hatton S., Devriendt J.E.G., Ninin S., Bouchet F.R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75

Hebb D.O., 1949, The Organization of Behavior, John Wiley & Sons

Helly J.C., Cole S., Frenk C.S., Baugh C.M., Benson A.J., Lacey C.G., 2003, MNRAS, 338, 903

Hinshaw G., et al., 2008, ApJSS, submitted

Ho S., Lin Y., Spergel D., Hirata C.M., 2008, MNRAS, submitted (arXiv:0706.0727)

Holden B.P., et al., 2005, ApJ, 626, 809

Hopkins P.F., Hernquist L., Cox T.J., Robertson B., Springel V., 2006, ApJS, 163, 50

Hütsi G., 2006, A&A, 459, 375

Jesseit R., Naab T., Burkert A., 2002, ApJ, 571, 89

Jing Y.P., 2000, ApJ, 535, 30

Jørgensen I., Chiboucas K., Flint K., Bergmann M., Barr J., Davies R., 2006, ApJ, 639, L9

Jørgensen I., Franx M., Kjærgaard P., 1996, MNRAS, 280, 167

Kaiser N., 1987, MNRAS, 227, 1

Kauffmann G., White S.D.M., Guiderdoni B., 1993, MNRAS, 264, 201

Kauffmann G., 1996, MNRAS, 281, 487

Kennicutt R.C., 1983, ApJ, 272, 54

Kobayashi C., 2005, MNRAS, 361, 1216

Kodama T., Arimoto N., Barger A.J., Aragón-Salamanca A., 1998, A&A, 334, 99

Kormendy J., 1977, ApJ, 218, 333

Kulkarni G.V., Nichol R.C., Sheth R.K., Seo H., Eisenstein D.J., Gray A., 2007, MNRAS, 378, 1196

Lacey C.G., Silk J., 1991, ApJ, 381, 14

Lacey C.G., Cole S., 1993, MNRAS, 262, 627

Lacey C.G., Baugh C.M., Frenk C.S., Silva L., Granato G.L., Bressan A., 2008, MNRAS, in press, arXiv:0704.1562

Larson R.B., 1975, MNRAS, 173, 671

Le Delliou M., Lacey C.G., Baugh C.M., Morris S.L., 2006, MNRAS, 365, 712

Le Delliou M., Lacey C.G., Baugh C.M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S.L., 2005, MNRAS, 357, 11

Lemson G., Virgo Consortium, 2006, astro-ph/0608019

Libeskind N.I., Cole S., Frenk C.S., Okamoto T., Jenkins A., 2007, MNRAS, 374, 16

McCarthy P.J., Spinrad H., Djorgovski S., Strauss M.A., van Breugel W., Liebert J., 1987, ApJL, 319, 39

Mandelbaum R., Seljak U., Cool R.J., Blanton M., Hirata C.M., Brinkmann J., 2006, MNRAS, 372, 758

Matković A., Guzmán R., 2005, MNRAS, 362, 289

Malbon R.K., Baugh C.M., Frenk C.S., Lacey C.G., 2007, MNRAS, 382, 1394

Maller A.H., Bullock J.S., 2004, MNRAS, 355, 694

Masjedi M., et al., 2006, ApJ, 644, 54

Mather J.C., et al., 1990, ApJ, 354, 37

McCarthy I.G., et al. 2007, MNRAS, 376, 497

McCulloch W., Pitts W., 1943, Bulletin of Mathematical Biophysics, 5, 115

Mo H.J., Mao S., White S.D.M., 1998, MNRAS, 295, 319

Mo H.J., White S.D.M., 1996, MNRAS, 282, 347

Nagashima M., Lacey C.G., Okamoto T., Baugh C.M., Frenk C.S., Cole S., 2005b, MNRAS, 363, 31

Nagashima M., Lacey C.G., Baugh C.M., Frenk C.S., Cole S., 2005a, MNRAS, 358, 1247

Navarro J., Frenk C., White S., 1997, ApJ, 490, 493

Negroponte J., White S.D.M., 1983, MNRAS, 205, 1009

Nelan J.E., Smith R.S., Hudson M.J., Wegner G.A., Lucey J.R., Moore S.A.W., Quinney S.J., Suntzeff N.B., 2005, ApJ, 632, 137

Norberg P., et al. 2002, MNRAS, 332, 827

Okamoto T., Eke V.R., Frenk C.S., Jenkins A., 2005, MNRAS, 363, 1299

Padmanabhan N., et al., 2007, MNRAS, 378, 852

Pahre M.A., Djorgovski S.G., de Carvalho R.R., 1998, ApJ, 116, 1591

Pahre M.A., Djorgovski S.G., de Carvalho R.R., 1999, ASPC, 163, 283

Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557

Peacock J.A., Smith R.E., 2000, MNRAS, 318, 1144

Peebles P.J.E., 1982, ApJ, 263, 1

Penzias A.A., Wilson R.W., 1965, ApJ, 142, 5258

Pettini M., Shapley A.E., Steidel C.C., et al., 2001, ApJ, 554, 981

Porciani C., Giavalisco M., 2002, ApJ, 565, 24

Press W.H., Schechter P., 1974, ApJ, 187, 425

Puget J.L., Abergel A., Bernard J.P., Boulanger F., Burton W.B., Desert F.X., Hartmann D., 1996, A&A, 308, 5

Reda F.M., Forbes D.A., Hau G.K.T., 2006, MNRAS, 360, 693

Riedmiller M., Braun H., 1993, Proc. of the IEEE Intl. Conf. on Neural Networks, 586

Robertson B., Cox T.J., Hernquist L., Franx M., Hopkins P.F., Martini P., Springel V., 2006, ApJ, 641, 21

Rosenblatt F., 1958, Psychological Review, 65, 386

Rosenblatt F., 1962, Principles of Neurodynamics, Spartan

Roseboom I.G., et al., 2006, MNRAS, 373, 349

Ross N., et al., 2007a, MNRAS, 381, 573

Ross N., Shanks T., Cannon R.D., Wake D.A., Sharp R.G., Croom S.M., Peacock J.A., 2007b, MNRAS submitted, arXiv0704.3739

Rubin V.C., Thonnard N., Ford Jr. W.K., 1980, ApJ, 238, 471

Rumelhart D.E., Hitton G.E., Williams R.J., 1986, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, Vol. 1, 318

Sánchez A.G., Baugh C.M., Percival W.J., Peacock J.A., Padilla N.D. Cole S., Frenk C.S., Norberg P., 2006, MNRAS, 366, 189

Sandage A., Perelmuter J., 1990, ApJ, 361, 1

Sandage A., Visvanathan N., 1978a, ApJ, 223, 707

Sandage A., Visvanathan N., 1978b, ApJ, 225, 742

Scarselli F., Tsoi A.C., 1998, NN, 11, 15

Schade D., Barrientos L.F., López-Cruz O., 1997, ApJ, 477, 17

Seljak U., 2000, MNRAS, 318, 203

Sellwood J.A., McGaugh S.S., 2005, ApJ, 634, 70

Sheth R.K., Mo H.J., Tormen G., 2001, MNRAS, 323, 1

Silva L., Granato G.L., Bressan A., Danese L., 1998, ApJ, 509, 103

Simien F., de Vaucouleurs G., 1986, ApJ, 302, 564

Smail I., Ivison R.J., Blain A.W., 1997, ApJL, 490, 5

Smail I., 2002, Ap&SS, 281, 453

Smoot G.F., et al., 1992, ApJL, 396, 1

Springel V., et al., 2005, Nature, 435, 629

Steidel C.C., Adelberger K.L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1

Sutherland R., Dopita M., 1993, ApJS, 88, 253

Toomre A., Toomre J., 1972, ApJ, 178, 623nn

Tormen G., 1997, MNRAS, 290, 411

Trager S.C., Faber S.M., Worthey G., González J.J., 2000, AJ, 119, 1645

Treu T., Ellis R.S., Liao T.X., van Dokkum P.G., Tozzi P., Coil A., Newman J., Cooper M., Davis M., 2005, ApJ, 633, 174

Trujillo I., Burkert A., Bell E.F., 2004a, ApJL, 600, L39

Trujillo I., et al., 2004b, ApJ, 604, 521

Tully R.B., Fisher J.R., 1977, A&A, 54, 661

van de Ven G., van Dokkum P.G., Franx M., 2003, MNRAS, 344, 924

van de Wel A., Franx M., van Dokkum P.G., Huang J., Rix H.-W., Illingworth G. D., 2006, ApJ, 636, L21

van de Wel A., Franx M., van Dokkum P. G., Rix H.-W., Illingworth G. D., Rosati P., 2005, ApJ, 631, 145

van Dokkum P.G., Franx M., 1996, MNRAS, 281, 985

van Dokkum P.G., Franx M., Kelson D.D., Illingworth G.D., 2001, ApJ, 553, L39

van Dokkum P.G., van der Marel R.P., 2007, ApJ, 655, 30

Wake D.A., et al., 2006, MNRAS, 372, 537

Wake D.A., et al. 2008, MNRAS, submitted, arXiv:0802.4288

Warren M.S., Quinn P.J., Salmon J.K., Zurek W.H., 1992, ApJ, 399, 405

Wechsler R.H., Zentner A.R., Bullock J.S., Kravtsov A.V., Allgood B., 2006, ApJ, 652, 71

White S.D.M., Frenk C.S., 1991, ApJ, 379, 52

White S.D.M., Frenk C.S., Davies M., 1983, ApJ, 274, L1

White S.D.M., Rees M.J., 1978, MNRAS, 183, 341

Wilman R.J., Gerssen J., Bower R.G., Morris S.L., Bacon R., de Zeeuw P.T., Davies R.L., 2005, Nature, 436, 227

Wuyts S., van Dokkum P.G., Kelson D.D., Franx M., Illingworth G.D., 2004, ApJ, 605, 677

York D.G., et al., 2000, AJ, 120, 1579

Zehavi I., et al., 2005, ApJ, 621, 22

Zheng Z., et al., 2005, ApJ, 633, 791

Ziegler B., Thomas D., Böhm A., Bender R., Fritz A., Maraston C., 2005, A&A, 433, 519

Zwicky F., 1933, Helvetica Physica Acta, 6, 110