November 2015

An empirical approach to understanding of star formation in dark matter halos

Zhankui Lu
University of Massachusetts Amherst

AN EMPIRICAL APPROACH TO UNDERSTANDING OF STAR FORMATION IN DARK MATTER HALOS

A Dissertation Presented

by

ZHANKUI LU

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2015

Astronomy

AN EMPIRICAL APPROACH TO UNDERSTANDING OF STAR FORMATION IN DARK MATTER HALOS

A Dissertation Presented

by

ZHANKUI LU

Approved as to style and content by:

H. J. Mo, Chair

Neal Katz, Member

Martin Weinberg, Member

Andrea Pocar, Member

Grant Wilson, Graduate Program Director
Astronomy

DEDICATION

To my family and friends

ACKNOWLEDGMENTS

Foremost my gratitude will go to my advisor, Prof. Houjun Mo, for his supports over the past five years. Houjun is full of interesting ideas and always stays positive and optimistic, and the inspiring discussions with him helped me overcome many difficulties in my thesis project and encouraged me to strive forward. Houjun also gave me a lot of freedom to pursue different projects on my own, for which I feel very grateful. I would also like to thank Prof. Daniel Wang, my first advisor at UMass, for his guidance in my early graduate career.

I extend my thanks to other members of the thesis committee: Prof. Neal Katz, Prof. Martin Weinberg and Prof. Andrea Pocar. Neal and Martin's enthusiastic discussion greatly broadened my knowledge on my research area and helped me finished my journal papers related to this thesis. I'm also grateful to my collaborators including Dr. Yu Lu and Prof. Frank van den Bosch, who made great contributions to those papers. Yu is also the one who introduced me to my current research field and who lent a lot of help to code up my own computing library. Additionally my thanks goes to Gretchen Zwart, our high perfomance computing administrator. She is always ready to help, even during her vacation times.

Lastly, I wish to express my thanks to my supportive family, including my parents and my sister, as well as many friends made at UMass. In addition to the professional support from the professors and colleagues, the emotional support from my family and friends also helped me to go through many hard times in my graduate career.

ABSTRACT

AN EMPIRICAL APPROACH TO UNDERSTANDING OF STAR FORMATION IN DARK MATTER HALOS

SEPTEMBER 2015

ZHANKUI LU

B.Sc., NANJING UNIVERSITY

M.Sc., UNIVERSITY OF MASSACHUSETTS

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor H. J. Mo

We present a data-driven approach to understand the star formation in dark matter halos over cosmic time. With a simple empirical model and advanced tools for Bayesian inference, we try to constrain how galaxies have assembled their stars across cosmic time using stellar mass functions (SMFs) and the luminosity function of cluster galaxies. The key ingredients of the empirical model include dark halo merger trees and a generic function that links star formation rate (SFR) to the host halos. We find a new characteristic redshift $z_c \approx 2$ above which the SFR in low mass halos $< 10^{11}\,\mathrm{M}_\odot$ must be enhanced relative to that at lower z. This leads to some interesting predictions, for instance, a significant old stellar population in present-day dwarf galaxies with $M_\star \sim 10^8\,\mathrm{M}_\odot$ and steep low-mass end slopes of high redshift SMFs. The constrained empirical model can be combined with other other observational constraints to infer the physics behind the evolution of galaxies. The classical bulge mass could be derived from the major mergers of the host galaxies. Applying the central black hole (BHs) - classical bulge relation, it

predicts all galaxies with $M_\star < 10^{10.5}\,\mathrm{M}_\odot$ host intermediate mass BHs ($M_{\rm BH} < 10^7\,\mathrm{M}_\odot$). Using the gas phase metallicity we study the evolution of gas and metal content of star forming galaxies and the inflow and outflow rates. About 60% of the metals produced have been lost. At low redshift ($z < 1$) the accretion of pristine gas should be lower by a factor of few than expected and the loading factor of gas outflow that is not recycled is of order of unity. The empirical model also serves as basis to study the evolution of satellite galaxies. The progenitors of present-day satellites can be initialized using this empirical model. The physically motivated models of quenching of star formation in satellites, including strangulation, ram pressure stripping of cold gas disks, and tidally triggered starburst, are tested against statistics of the group catalogs.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS . . . . . **v**

ABSTRACT . . . . . **vi**

LIST OF TABLES . . . . . **xii**

LIST OF FIGURES . . . . . **xiii**

CHAPTER

1. INTRODUCTION **1**

1.1 The census of the galaxy population 1
1.2 Theory of galaxy formation and evolution 5

1.2.1 Cosmology 5
1.2.2 Formation of dark matter halos 5
1.2.3 Galaxy formation in dark matter halos 8

1.3 Outstanding problems and challenges 11
1.4 Outline of the dissertation 13

2. THE EMPIRICAL MODEL **16**

2.1 Introduction 16
2.2 The empirical of galaxy evolution 18

2.2.1 Dark halo merger histories 18
2.2.2 Star formation in central galaxies 18
2.2.3 Star formation in satellite galaxies 19
2.2.4 Merging and stripping of satellite galaxies 20
2.2.5 Spectral synthesis and metal enrichment 21

2.3 Observational Constraints and Likelihood Functions 23

2.3.1 The SMFs of field galaxies 23

2.3.2 The composite luminosity function of cluster galaxies 24
2.3.3 The Likelihood Function and Sampling Algorithm 26

2.4 Construct the Parametric Models 27

2.4.1 The Simplest Model Family (Model I) 27
2.4.2 Fixing the Bright-End Problem (Model II) 29
2.4.3 Fixing the faint end problem of the cluster galaxy luminosity function (Model III) 34
2.4.4 Are more general model families necessary? 41

2.5 Summary and Discussion 44

3. MODEL PREDICTIONS 48

3.1 The correlations between galaxies and dark matter halos 48

3.1.1 The stellar mass - halo mass relation of central galaxies 48
3.1.2 SFR and halo mass accretion rate 49
3.1.3 The specific star formation rate 51

3.2 The Stellar Mass Functions 52

3.2.1 The evolution of galaxy SMFs 52
3.2.2 The SMFs of star forming and quenched galaxies 54
3.2.3 The SMF of satellite galaxies 55
3.2.4 The conditional stellar mass functions (CSMFs) 56

3.3 The cosmic star formation history 58

3.3.1 The cosmic star formation rate density 58
3.3.2 The star formation rate function 61

3.4 The projected two point correlation function 64
3.5 Summary 64

4. STAR FORMATION AND STELLAR MASS ASSEMBLY IN DARK MATTER HALOS 66

4.1 General Trends and Characteristic Scales 66
4.2 Star formation and stellar mass assembly 70

4.2.1 Star formation histories 70
4.2.2 Stellar mass assembly histories 72
4.2.3 In-situ star formation versus accretion 73
4.2.4 Downsizing versus upsizing 74

4.3 Merger history and the transformation of galaxies 76
4.4 Stellar populations 82

4.4.1 Stellar ages 82
4.4.2 Halo Stars 84

4.5 Summary and Discussion 88

5. CLASSICAL BULGES, SUPERMASSIVE BLACKHOLES AND AGN FEEDBACK: EXTENSION TO LOW-MASS GALAXIES 92

5.1 Introduction 92
5.2 Implications of galaxy merger histories for supermassive black holes 94
5.3 AGN feedbck from low-mass galaxies 98
5.4 Discussion 101

6. GALAXY ECOSYSTEMS: GAS CONTENTS, INFLOWS AND OUTFLOWS 104

6.1 Introduction 104
6.2 Galaxy Ecology 106

6.2.1 The basic equations 106
6.2.2 The Halo Assembly History 108
6.2.3 The Intrinsic Metal Yield 109

6.3 Observational Constraints 109

6.3.1 Star Formation History 109
6.3.2 Gas Phase Metallicity 111
6.3.3 Gas Fraction in Local Galaxies 111

6.4 Evolution of cold gas content of galaxies 112

6.4.1 The Star Formation Models 112
6.4.2 Disk Size 116
6.4.3 The Cold Gas Contents 117

6.5 Inflow and Outflow 118

6.5.1 Models with strong gas outflow 122
6.5.2 Constraining gas inflow and outflow 125
6.5.3 Metal loss 127

6.6 Conclusions and Implications 128

7. STAR FORMATION AND QUENCHING OF SATELLITE GALAXIES ... 132

7.1 Introduction ... 132
7.2 Progenitors of Satellite Galaxies ... 134
7.3 Evolution of Satellite Galaxies ... 136

7.3.1 Orbits of Satellites ... 136
7.3.2 High-order Substructures ... 138
7.3.3 Star Formation and Outflow ... 138
7.3.4 Ram Pressure Stripping ... 139
7.3.5 Tidally Triggered Starburst ... 141
7.3.6 Environmental Quenching versus Self-Quenching ... 142

7.4 Model Predictions ... 143

7.4.1 Quenching by Strangulation ... 143
7.4.2 Quenching by ram pressure stripping ... 149
7.4.3 Quenching by Tidally Triggered Starburst ... 152
7.4.4 Conspiracy between ram pressure stripping and tidally triggered starburst ... 154

7.5 Summary ... 156

8. CONCLUSIONS AND OUTLOOK ... 160

8.1 Conclusions ... 160

8.1.1 The star formation history across cosmic time ... 160
8.1.2 Reverse Engineering ... 161

8.2 Prospective Science Projects ... 163

8.2.1 A Generic Model of Galaxy-Halo Connection ... 163
8.2.2 Modelling the Evolution of Dwarf Galaxies ... 164

8.2.2.1 Star Formation History in Hydrodynamic Simulations ... 165

APPENDIX: MULTINEST SAMPLING OF THE POSTERIOR DISTRIBUTION ... 167

BIBLIOGRAPHY ... 170

LIST OF TABLES

Table **Page**

2.1 Summary of the posterior simulations, including the models, the observational constraints and the marginalised likelihood. 29

2.2 The constrained model parameters of Model II, in terms of the means and the variances. 30

2.3 The same as Table 2.2, but for Model III. 35

6.1 The oxygen yield as a function of initial stellar metallicity. 108

6.2 Parameters in the Kennicutt-Schmidt model, $A_{\rm K}$ and $\Sigma_{\rm c}$, and in the Krumholz model, $\tau_{\rm sf}$ and c, tuned together with the disk-size parameter, $\mathcal{L}$, to match the gas mass to stellar mass ratio of local galaxies (data points in Figure 6.3). 112

LIST OF FIGURES

Figure **Page**

1.1 The evolution of halo mass functions. 7

1.2 Conditional halo mass functions of the halos of mass $10^{13}\,h^{-1}\mathrm{M}_{\odot}$ at present. 7

2.1 The z-band magnitude of a simple stellar population with initial mass $1\,\mathrm{M}_{\odot}$ as a function of age and metallicity. 22

2.2 The posterior predicted SMFs of Model I constrained by the SMF at $z \approx 0$. 28

2.3 The same as Figure 2.2 but for Model I constrained using SMFs at all the redshift bins. 30

2.4 The posterior distribution of the parameters of Model II constrained using the SMFs at different redshift bins. 31

2.5 The same as Figure 2.2 but for Model II constrained using the SMFs at different redshift bins. 32

2.6 The SFR as a function of halo mass (left panel) and redshift (right panel) of Model II, constrained by the SMFs. 33

2.7 Posterior predicted CGLF of Model II. 34

2.8 The posterior distribution of the parameters of Model III constrained by the SMFs and the CGLF. 36

2.9 The same as Figure 2.2 but predicted by the posterior of Model III constrained by both the SMFs and the observed z-band CGLF. 37

2.10 The CGLF predicted by the posterior of Model III constrained by both the SMFs and the observed z-band CGLF. 38

2.11 The SFR as a function of halo mass (left panel) and redshift (right panel) predicted by Model III. 40

2.12 The galaxy SMFs predicted by the posterior of Model IV constrained by both the SMFs and the observed z-band CGLF. 42

2.13 The CGLF predicted by the posterior of Model IV constrained by both the SMFs and the z-band CGLF. 43

2.14 The SFR as a function of halo mass (left panel) and redshift (right panel) predicted by Model IV. 43

3.1 The stellar mass to halo mass ratio as a function of halo mass for Model II (left panels) and Model III (right panels). 49

3.2 The median star formation efficiency as a function halo mass $M_{\rm h}$ for both Model II and Model III. 50

3.3 The specific star formation rate (sSFR) versus stellar mass at different redshifts for Model II (left panel) and Model III (right panel). 51

3.4 The evolution of galaxy SMFs. 53

3.5 The sSFR stellar mass relation predicted by Model III. 54

3.6 The SMF of star forming galaxies (left) and quenched galaxies (right) predicted by Model III. 55

3.7 The SMF of satellite galaxies. 56

3.8 The conditional SMFs for different halo mass bins as indicated. 57

3.9 The cosmic star formation rate density. 58

3.10 The predicted star formation rate density (red and green lines) and the minimal star formation rate density required to keep the Universe ionised (blue lines). 59

3.11 The star formation rate functions at different redshifts predicted by Model II (upper panels) and Model III (lower panels). 62

3.12 Projected 2 point correlation function of the galaxies with mass $> 10^{12}\,{\rm M}_\odot$ (black) and blue galaxies in the same mass range (blue). 63

4.1 Star formation efficiency and merger rate in the halo mass - redshift plane. 67

4.2 The star formation histories of galaxie with different present-day halo masses. 70

4.3 The star formation history of dwarf central galaxies in halos with masses $\approx 3 \times 10^{10}\, h^{-1}\mathrm{M}_{\odot}$. 71

4.4 The fraction of stars formed *in situ* in present day central galaxies of halos with different masses. 73

4.5 The distribution of the redshift at which 50% ($z_{0.5}$) and 10% ($z_{0.1}$) of the stars in a central galaxy have formed (green contours) and have assembled (red contours). 74

4.6 The average characteristic redshifts, $z_{0.5}$ and $z_{0.1}$, as a function of halo mass. 75

4.7 The bulge-to-total mass ratio of local central galaxies as a function of host halo mass. 77

4.8 The merger histories of present-day galaxies. 78

4.9 The redshift of the last major merger of present-day central galaxies as a function of host halo mass. 79

4.10 The "bulge" mass fraction of dwarf galaxies with total stellar masses in the range $10^8\, h^{-2}\mathrm{M}_{\odot}$ to $10^9\, h^{-2}\mathrm{M}_{\odot}$. 80

4.11 The stellar age as a function of stellar mass predicted by Model II (colored dashed lines) and Model III (colored solid lines). 82

4.12 The SMF of galaxies predicted by Model III (grey line) compared to the prediction of the same model but assuming no stripping of satellites, i.e. $f_{\mathrm{TS}} = 1$ (the thick black line). 84

4.13 The mass ratio between halo stars and the stellar mass of the central galaxy as a function of host halo mass at $z = 0$. 86

4.14 The solid contours show the formation redshift and location (i.e. the stellar mass of the galaxy) of the halo stars. 87

5.1 Relation between the mass of the classical bulge (formed via merger of galaxies) and the total stellar mass of the host galaxy, predicted by Model II (left) and Model III (right). 94

5.2 Relation between the SMBH mass and the total galaxy stellar mass predicted by Model II (left) and Model III (right). 95

5.3 SMBH mass function predicted by Model II (colored dashed lines) and Model III (colored solid lines) at $z = 0$ (blue) and $z = 3$ (red). 97

6.1 The dependence of SFR (left) and stellar mass to halo mass ratio (right) on halo mass and redshift. 110

6.2 Gas phase metallicity compiled by Maiolino et al. (2008). 110

6.3 Cold gas to stellar mass ratio as a function of stellar mass at different redshifts calculated using the Kennicutt-Schmidt Law (left) and the Krumholz model (right). 113

6.4 Molecular gas to stellar mass ratio (left), molecular gas to total gas mass ratio (right) as a function of stellar mass, all predicted by the Krumholz model. 113

6.5 Baryon mass (in both stars and cold gas) to halo mass ratio as a function of halo mass calculated using the Kennicutt-Schmidt Law (left) and the Krumholz Law (right). 114

6.6 The net yield $\mathcal{Y}$, loading factor of gas mass loss rate $\epsilon_{\rm loss}$ and gas accretion efficiency $\epsilon_{\rm acc}$. 119

6.7 The same as Figure 6.6 but here the Krumholz star formation model is assumed. 120

6.8 The net yield at $z = 0$ calculated using Eq. (6.21) assuming $\epsilon_{\rm acc} = 1$. 124

6.9 The loading factor of metal loss rate. 127

6.10 Comparison between the upper limit of the accretion efficiency derived from our empirical model (red lines) with a number of physical models. 128

7.1 Fractions of quenched galaxies predicted by the model in which disk gas is consumed only by star formation and associated outflow. 143

7.2 The distribution of the sSFR of satellite galaxies predicted by the model in which disk gas is consumed only by star formation and associated outflow. 144

7.3 Fraction of quenched satellites as a function host halo mass, predicted by the model in which disk gas is consumed only by star formation and associated outflow. 146

7.4 The predicted fractions of quenched galaxies by the model with ram-pressure stripping. 147

7.5 The distribution of sSFR of satellite galaxies predicted by the model of ram-pressure stripping. 147

7.6 The evolution history of three satellites on three different initial orbits 148

7.7 The fraction of quenched satellites as a function host halo mass. 148

7.8 The predicted fractions of quenched galaxies by the model of tidally-triggered starburst. 149

7.9 The distribution of the sSFR of satellite galaxies predicted by the model of tidally-triggered starburst. 150

7.10 The fraction of quenched satellites as a function host halo mass. by the model of tidally-triggered starburst with $\epsilon_{\rm sb,w} = 5$. 151

7.11 The fraction of quenched satellites as a function host halo mass. 152

7.12 The distribution of sSFR predicted by the model in which both ram-pressure stripping and tidally-triggered starburst (with $\epsilon_{\rm sb,w} = 3$) are included. 153

7.13 The fraction of quenched satellites as a function of the redshift of accretion, predicted by the model in which both ram-pressure stripping and tidally-triggered starburst (with $\epsilon_{\rm sb,w} = 3$) are included. 153

CHAPTER 1

INTRODUCTION

1.1 The census of the galaxy population

Observational and theoretical studies suggest that galaxies are ecosystems of many different components, including stars, diffuse gas and presumably dark matter, which are involved in a network of interactions. Galaxies are diverse and it requires a number of quantities to fully describe a galaxy.

Luminosity or stellar mass is a measure of how many stars contained in a galaxy. Observations show that this quantity spans more than six decades. The largest galaxies are the brightest cluster galaxies (BCGs) located at the center of galaxy clusters, which can reach $10^{12}\,\mathrm{M}_{\odot}$. The faintest ones are the ultra-faint dwarfs found in the Milky Way neighborhood and are $\sim 10^{5}\,\mathrm{M}_{\odot}$.

Interstellar medium (ISM) is the gas contained in a galaxy and is multiphase. The temperature ranges from 10K to 10^{7}K. The coldest gas is in molecular phase and the distribution and the total mass can be determined from tracers, such as CO. The atomic gas is about $\sim$ 100K and can be traced using the $21cm$ spectral line of the neutral atomic hydrogen.

Metallicity is the fraction of metals in the atmosphere of the stars or the gas. Elements heavier than Helium are considered as metals in astronomy. Since most of the metals are produced by the nuclear reaction inside stars or during supernovae, younger generations of stars generally have higher metallicity.

Color of a galaxy is the ratio between the luminosity in two photometric passbands. A redder color could mean a stellar population with older age or higher metallicity, or abundant dust surrounding the stars, which preferentially absorbs the photons with lower wavelength.

Star formation rate (SFR) is the rate at which gas converts into stars. Tracers, such as the UV luminosity or the recombination lines, are used to measure the SFR. Such tracers actually only trace the formation of massive stars, and the initial mass function (IMF) is required to convert into the total SFR.

Active galactic nucleus (AGN) and Supermassive black holes (SMBHs) Some galaxies have spectral energy distribution (SED) that covers a full wavelength range from radio to X-ray. The SED is much broader than what is expected from stars, gas and dust all together. The emission is concentrated within parsecs in the central region and can surpass the total emission of other parts of the host galaxies by a large factor. They are called AGN. The different types of AGNs can be explained with a unified model, that is accretion of gas onto SMBH, with mass of $10^6 \sim 10^9\,\mathrm{M}_\odot$.

The major task in the study of extragalactic astronomy is to census the galaxy population, such as their spatial distribution, the abundance of galaxies with respect to their properties and the correlation between different properties. One of the most prominent discoveries of the census is the dichotomy: the galaxies fall into two distinct populations. One population is red in color, spheroidal in morphology and shows weak star formation activity and contains little neutral gas, while the other is blue, has apparent stellar and gas discs and is actively forming stars.

The galaxy SMF is the abundance of galaxies as a function of their stellar mass. The galaxy SMF of local galaxies has been well established. It can be described with a Schechter function, characterized by a power law at the low mass end and an exponential decay in the high mass end

$$\Phi\left(M_{\star}\right)=\Phi_{*}\left(\frac{M_{\star}}{\mathcal{M}_{*}}\right)^{\alpha}\exp\left(-\frac{M_{\star}}{\mathcal{M}_{*}}\right), \tag{1.1}$$

where $\mathcal{M}_{*}$ is the characteristic stellar mass and is about $10^{10}\,\mathrm{M}_{\odot}$ for the local galaxies. With the advent of multi-wave band deep surveys from the Hubble Space Telescope, galaxy SMFs/LFs can now be determined up to a redshift of $z\approx 8$ (e.g. Bradley et al., 2012). These observations have revealed a number of interesting findings. The amplitude of the SMFs (Φ_{*}) increases as the Universe evolves from high to low redshift, while the characteristic stellar mass ($\mathcal{M}_{*}$) does not change significantly (e.g. Pérez-González et al., 2008; Marchesini et al., 2009; Marchesini et al., 2012; Santini et al., 2012b). For local galaxies, large redshift surveys have now made it possible to determine the SMF of galaxies down to $10^{7}\,\mathrm{M}_{\odot}$ (e.g. Baldry et al., 2012). Such surveys have revealed that the Schechter function may not be sufficient, especially for galaxies in massive clusters. Instead there seems to be a marked upturn at the faint end ($M_r-5\log_{10}h>-17$) of the cluster galaxy luminosity function (CGLF) (Popesso et al., 2006; Milne et al., 2007; Jenkins et al., 2007; Barkhouse et al., 2007; Bañados et al., 2010; Wegner, 2011). This feature has also been found for galaxies in the general field regardless of their environments, which are usually referred to as field galaxies (Blanton et al., 2005; Baldry et al., 2012; Drory et al., 2009; Loveday et al., 2012), although the upturn appears shallower than that for cluster galaxies.

Tight correlations between different properties have been revealed. Here we list a couple of examples

- The galaxies follow a bimodal distribution on the SFR-$M_{\star}$ plane. The star forming galaxies fall on a tight sequence. The SFR and the stellar mass relation can be described using a simple power law SFR $\propto M_{\star}^{\alpha}$ with lognormal scatter $1\sigma\lesssim 0.3\,\mathrm{dex}$. The normalization and the power index α can be found in Figure 1 of Leitner (2012), which shows a compilation of data from a variety of sources. The normalization increases as $(1+z)^{3.4}$. The power index depends weakly on time but varies with

different data sources which is most likely caused by the different stellar mass ranges in the data sources. It roughly lies between 0.6 and 1.

- Tight correlation has been found between the gas phase metallicity and the stellar mass (Tremonti et al., 2004). Typically the gas phase metallicity is measured from the emission line of the HII regions. It increases with the stellar mass and becomes saturated at $10^{10}\,\mathrm{M}_\odot$. Also efforts have been made to measure the gas phase metallicity of high redshift galaxies. Although the data sample is small and the measurement is subject to considerable systematic uncertainties, different works have shown that the gas phase metallicity increased dramatically since $z \approx 3.5$ for galaxies of similar mass. It is also suggested that the stellar mass, SFR and gas phase metallicity form a fundamental plane, with the latter two anti-correlated with each other (Mannucci et al., 2010).

- Another prominent correlation is the relation between the mass of the central black holes and the velocity dispersion of the classic bulges and elliptical galaxies

$$\log_{10}\left(\frac{M_{\mathrm{BH}}}{10^9\,\mathrm{M}_\odot}\right) = -0.510 + 4.377\log_{10}\left(\frac{\sigma}{200\mathrm{km/s}}\right)\,. \tag{1.2}$$

  The intrinsic scatter in this relation is about 0.29 dex over 4 decades in M_{BH} (Kormendy and Ho, 2013, §6.6). The tight correlation is surprising because the black holes and the bulges are on completely different spatial scales. Tight correlation between M_{BH} and the mass of the classic bulges or elliptical galaxies also holds with similar scatter. No correlation was found between M_{BH} and the mass or velocity dispersion of peudo-bulges.

1.2 Theory of galaxy formation and evolution

1.2.1 Cosmology

The formation of galaxies is studies in a cosmological context, that is the large scale geometry and evolution of the Universe we live in. It is determined by the matter (energy) contents, including radiation, particles with non-zero rest mass (baryons and dark matter) and dark energy. The Friedmann equation,

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left(\Omega_{\mathrm{r},0}(1+z)^4 + \Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_0)(1+z)^2 + \Omega_{\Lambda,0}\right) \tag{1.3}$$

describes how the matter in different forms determines the fate of a universe. The left side of the equation is the growth rate of the scale factor a. On the right side, H_0 is the Hubble constant and $\Omega_{\mathrm{r},0}$, $\Omega_{\mathrm{m},0}$ and $\Omega_{\Lambda,0}$ are the mass fraction of radiation, matter and dark energy. In the past decades, these parameters have precisely measured using distant Type Ia Supernovae, the cosmic microwave background (CMB), and baryonic acoustic oscillation. The Universe is believed to be homogeneous and isotropic on large scales, but not on small scales. The small-scale perturbations in the early Universe result in the anisotropy in the present CMB, and also seed the galaxies and galaxy clusters. The density peaks in the perturbed density field of matter (dark matter and baryons) eventually collapse because of gravitational instability and trigger the formation of galaxies. Throughout the paper, we use a ΛCDM cosmology with $\Omega_{\mathrm{m},0} = 0.273$, $\Omega_{\Lambda,0} = 0.727$, $\Omega_{\mathrm{b},0} = 0.0455$, $h = 0.704$, $n = 0.967$ and $\sigma_8 = 0.811$. The last two parameters specify the power spectrum of the initial perturbation. This set of parameters is from the seven year WMAP observations (Komatsu et al., 2011).

1.2.2 Formation of dark matter halos

Dark matter halos are gravitationally bound clumps of dark matter particles. They are the cradles where galaxies form and evolve. The formation of dark matter halos

starts with the initial perturbations in the overdensity field $\delta(\mathbf{x})$. The amplitude of the overdensity on different scales can be characterized by the rms mass fluctuation $\sigma(M)$,

$$\sigma^2(M) = \left\langle \delta_{\rm S}^2(\mathbf{x}, R) \right\rangle \ . \tag{1.4}$$

M is total mass enclosed within a sphere of radius R. $\delta_{\rm S}(\mathbf{x}, R)$ is the density field smoothed with a top-hat window function of radius R. $\sigma(M)$ is the average amplitude of the overdensity on scale R. $\sigma(M)$ decreases monotonically with M, which means overdensity is generally larger on smaller scales. In the linear regime, the density field on different scales grow at the same pace, which is described by the linear growth factor $D(t)$. Beyond the linear regime the density peaks will collapse and form virialized dark halos. According to the simple spherical collapse model, a halo of mass M will form once a peak in the smoothed density field $\delta_{\rm S}(\mathbf{x}, M)_0 D(t) > \delta_{\rm c} \approx 1.69$. The larger amplitude in the fluctuation on smaller scales suggest that typically a less massive halo form earlier than more massive ones.

Press and Schechter (1974) proposed an analytic scheme based on the spherical collapse model and that initial overdensity field is a Guassian random field to calculate the number density of dark mass halos as a function of mass and time. The Press-Schechter formulum is written as

$$\frac{{\rm d}n}{{\rm d}\ln M} = -\sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \frac{\delta_{\rm c}(z)}{\sigma(M)} \frac{{\rm d}\ln\sigma(M)}{{\rm d}\ln M} \exp\left(-\frac{\delta_c^2(z)}{2\sigma^2(M)}\right) \ , \tag{1.5}$$

where ρ_0 is the matter density at present and $\delta_c(z) = \delta_c / D(z)$. A more accurate halo mass function is derived by Sheth et al. (2001) based on the ellipsoidal collapse model. The non-spherical feature in collpase can change the formation time of the halos and therefore the mass function. The new mass function writes as

$$\frac{{\rm d}n}{{\rm d}M} = -\frac{\rho_0}{M} \frac{\delta_{\rm c}(z)}{\sigma(M)} \frac{{\rm d}\ln\sigma(M)}{{\rm d}\ln M} A\sqrt{\frac{2a}{\pi}} \left(1 + \left(\frac{\sigma^2(M)}{a\delta_{\rm c}^2(z)}\right)^p\right) \exp\left(-a\frac{\delta_c^2(z)}{2\sigma^2(M)}\right) \ , \tag{1.6}$$

where $A = 0.3222$, $a = 0.707$ and $p = 0.3$.

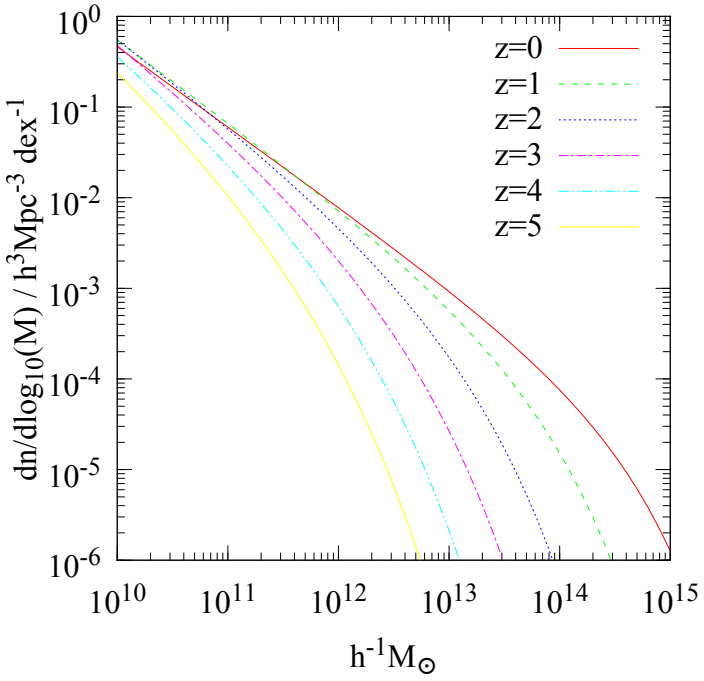

Figure 1.1. The evolution of halo mass functions.

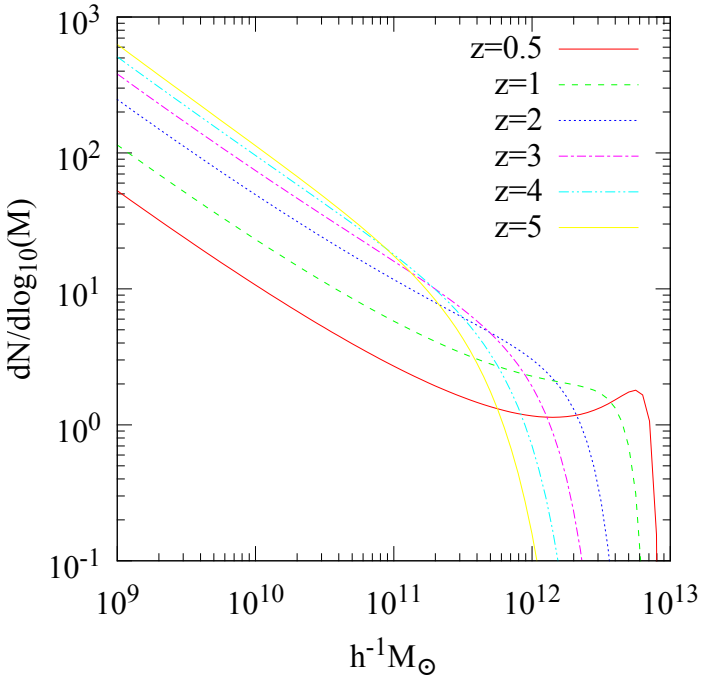

Figure 1.2. Conditional halo mass functions of the halos of mass $10^{13}\,h^{-1}\mathrm{M}_{\odot}$ at present.

That perturbations on small scales collapse earlier than large ones means that halos form in a hierarchical manner. In a sphere enclosing mass M that collapses at z_2, smaller halos have already formed at an earlier time. The halo mass function in this sphere at an earlier time z_1 has been derived using the extended Press-Schechter theory (Bond et al., 1991). This is the conditional halo mass function

$$\frac{\mathrm{d}N}{\mathrm{d}\ln M_1}_{M_1,z_1|M_2,z_2} = -\frac{M_2}{M_1}\sqrt{\frac{2}{\pi}}\frac{\sigma_1^2\left(\delta_1-\delta_2\right)}{\left(\sigma_1^2-\sigma_2^2\right)^{3/2}}\exp\left(-\frac{1}{2}\frac{\left(\delta_1-\delta_2\right)^2}{\left(\sigma_1^2-\sigma_2^2\right)^2}\right)\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M_1}\,. \tag{1.7}$$

One can sample this conditional halo mass function using Monte Carlo method recursively to build up the full merger history given a dark matter halo with mass M and formed at z.

1.2.3 Galaxy formation in dark matter halos

The baryonic matter also collpases with the dark matter. Without radiative cooling, or in the case that cooling time scale is significantly lower than the age of the system, a converging shock will heat up the inflowing gas to a temperature of $T_{\rm vir} = \frac{1}{2}\frac{\mu m_{\rm p}}{k_{rmb}}V_{\rm vir}^2$ at about the virial radius and a hydrostatic and hot gas halo will form in the potential well of the dark matter halos. Efficient radiative cooling can make the post-shock gas graviationally unstable (Birnboim and Dekel, 2003). In this case, the cold gas flows onto the galactic center directly on a dynamic time-scale. The converging shock may form at a inner radius or never form at all. These different ways of accretion are referred to as "hot mode" and "cold mode" respectively. The accretion modes strongly depend on the halo mass. It is shown that accretion in halos with mass $> 10^{11.5}\,\mathrm{M}_\odot$ is dominated by "hot mode" and dominated by "code mode" in halos with smaller mass (Birnboim and Dekel, 2003; Kereš et al., 2005).

Both the dark matter and the baryonic matter gain angular momentum from the tidal torque generated by the surrounding structures. For a virialized system, the overall

importance of the rotation relative to the random motion is characterized by the spin parameter

$$\lambda = \frac{J|E|^{1/2}}{GM^{5/2}}\,. \tag{1.8}$$

The spin paramter of dark matter halos follow a log-normal distribution with $\langle\lambda\rangle \approx 0.035$, which is nearly a constant as to different halo mass, redshift and cosmology (Barnes and Efstathiou, 1987). The cold gas flow or the gas that cools out of the hot halo condenses onto a disk because of the angular momentum. With a simple model that the gas disk follows an exponential profile, the host dark matter halo follows a singular isothermal sphere and the gas disk and the dark halos share the same distribution of the specific angular momentum, Mo et al. (1998) showed that the scale radius of the disk is

$$R_{\rm d} = \frac{\lambda}{\sqrt{2}} R_{\rm vir}\,. \tag{1.9}$$

In a more realistic model, the angular momentum distribution of the disk depends on how the gas cools out of the gaseous halo and therefore can be different from host dark halo (Lu et al., 2015a).

The most challenging part is the physics on sub-galactic scales, including star formation and stellar feedback and the formation of central black holes.

As to star formation, the two quantities relevant to galaxy evolution is the SFR and the initial mass function (IMF). A lot of effort has been made to understand how stars form out of the cold interstellar medium both empirically and theorectically. Kennicutt (1998) discovered that the star formation surface density is related to the HI surface density by a power law, that is $\Sigma_\star \propto \Sigma_{\rm HI}^{1.4}$. The surface densities are averaged over the whole galaxies. Tight linear relation has been found between the star formation surface density and the surface density of the molecular cloud by Bigiel et al. (2008), with a nearly constant molecular gas consumption time scale of 1.5Gyr. The IMF is the mass spectrum

that newly formed stars follow. There are different version of IMF in the literature. Two widely adopted ones include the Salpeter (1955) IMF

$$\frac{\mathrm{d}n}{\mathrm{d}\ln m} \propto m^{-1.35}\,, \tag{1.10}$$

with $0.4\,\mathrm{M}_\odot \leq m \leq 10\,\mathrm{M}_\odot$; and the Chabrier (2003) IMF

$$\frac{\mathrm{d}n}{\mathrm{d}\ln m} \propto \begin{cases} m^{-1.35} & \text{if } m > 1\,\mathrm{M}_\odot \\ \exp\left(-\left(\log_{10}(m/0.2\,\mathrm{M}_\odot)\right)^2/0.6\right) & \text{if } m < 1\,\mathrm{M}_\odot\,. \end{cases} \tag{1.11}$$

The shapes are quite different for dwarf stars ($m \leq 1\,\mathrm{M}_\odot$). The fraction of massive stars of a IMF determines how much energy and heavy elements are release into the interstellar medium.

Massive stars inject energy and momentum into the interstellar medium in different ways. The spectrum of massive star peaks at the UV band, and the energetic photons can be easily absorbed by the cold and dusty medium and transfer the momentum to the medium. The massive stars also die as type II supernovae (SNe) and the mechanic energy released can shock heat the interstellar medium to millions of Kelvin. It is believed the momentum and energy can power wind on galactic scales. However how much energy and mass is gone with the wind is still quite countravericial. As to the heating by SNe, the blast wave will sweep through the medium and at some point the heated gas will switch from adiabatic phase to radiative cooling phase. If the shock heated gas can not leak out of the gas disk before this transtion occurs, most of the mechanic energy will radiate away. There is no cooling problem for the radiation pressure scenario. However, in order for the radiation pressure to accelerate up the surrouding medium, the medium must efficiently trap the radiation. With sub-galactic simulation, Krumholz and Dekel (2012) showed that this is difficult because the accelerated gas will fragment due to radiative RT instability.

1.3 Outstanding problems and challenges

Because of the large dynamical range, it is impossible to build up a robust model from the basic laws of physics. As a compromise, two theoretical frameworks have been developed to study the statistical properties of the galaxy population, which are semi-analytic models and cosmological simulations with subgrid models.

A semi-analytic model starts with a set of pre-generated dark halo merger trees. A set of ordinary differential equations are used to describes all the baryonic physics such as the cooling rate, SFR and outflow rate. There are different ways to estimate these rates. For instance, the cooling rate is generally derived from simple analytic model of shocked isothermal atmosphere and the SFR is based on empirical calibrations. The pioneer works (White and Rees, 1978; White and Frenk, 1991; Cole, 1991) identified many impactful problems that are under intensive study until present day, such as the excessive amount of dwarf galaxies predicted by the CDM cosmology. Also they provided original solutions to such problems, such as the importance of feedback in shaping the galaxy population.

Cosmological simulation is to solve the hydrodynamic equations of the baryonic matter and dynamic equations of dark matter particles simultaneously, starting from the initial perturbations in the density field. Since first including radiative cooling in simulation of galaxy formation (Katz and Gunn, 1991; Katz, 1992; Katz et al., 1992), cosmological simulation has evolved into an important platform to study galaxy formation in a cosmological context. At present time, a typical cosmological simulation has a box size of $\sim$ 100Mpc and a resolution of $\sim$ kpc. This means physics that happens on scales smaller than the resolution, especially star formation and feedback, is implemented as subgrid prescriptions.

To match the great wealth of observational data, the galaxy formation models have been developed into systems with a large number of free parameters. Typically, a semi-analytic model has 15 to 20 tunable parameters (Lu et al., 2014b). As to cosmological simulations, the free parameters can be both explicit and implicit. There are different

ways to implement galactic wind, including "mechanic kick" scheme (Springel et al., 2005a) and "thermal heat" scheme (Stinson et al., 2006). In the former ISM particles near newly formed stars will be given a velocity $v_{\rm w}$ and hydrodynamics will be turned off on those wind particles until they get out of the host galaxies. In the latter, ISM near newly formed stars are heated to a certain temperature with radiative cooling shut off for a certain period. How much energy and momentum is injected into the ISM and how much ISM is affected are controlled by free parameters. Despite the large number of free parameters, the community still finds it difficult to reproduce some of the most basic statistics. Many attempts to explain the evolution of the stellar mass functions have failed. The problem seems to be common among different models. The models tend to overesitmate the number density of dwarf galaxies at high redshifts (Guo et al., 2011; Lu et al., 2014c; Mutch et al., 2013a). Another challenge is the evolution of the gas phase metallicity. It has been shown that since $z \approx 3.5$ the gas phase metallicity increased by a factor of ≈ 3 for the same stellar mass. In contrast, current galaxy evolution models are unable to explain such significant evolution (Lu et al., 2014c).

Different processes are degenerate, and this makes the models hard to comprehend. By tweaking the subgrid wind models, Oppenheimer et al. (2010) found that a wind model with $\eta \propto V_{vir}^{-1}$ is able to reproduce the low-mass end slope of the galaxy SMF at $z \approx 0$. As to semi-analytic models, various feedback recipes can be found in the current literature. In the semi-analytic model by Croton et al. (2006), the cold gas reheated by stellar feedback is

$$\frac{\Delta m_{\rm reheat}}{\Delta m_\star} = 3.5\,. \tag{1.12}$$

Bower et al. (2006) assumed that the reheating rate is

$$\frac{\Delta m_{\rm reheat}}{\Delta m_\star} = \left(\frac{485{\rm km/s}}{V_{\rm max}}\right)^{3.2}\,, \tag{1.13}$$

where $V_{\rm max}$ is the maximum circular velocity in a NFW halo. The heating rate is a strong function of the halo mass. In Somerville et al. (2008),

$$\frac{\Delta m_{\rm reheat}}{\Delta m_\star} = 1.3 \left(\frac{200{\rm km/s}}{V_{\rm max}} \right)^{2.0} . \tag{1.14}$$

Those stellar feedback recipes are quite different in their mathematical form and probably have different implications for the physics behind the stellar feedback. However, all of those models are claimed to be able to reproduce a wide range of observables, especially the present day stellar mass function. The wind recipes are embeded in a system with a large number of free parameters and are likely to conspire with some other processes and produce the same results. Yet due to the large parameter space, it is difficult to identify those degeneracies.

1.4 Outline of the dissertation

The usual approach using cosmological hydrodynamic simulations or semi-analytic models is to make *ab initio* models of galaxy formation including all the physical processes that one thinks are important. One then makes predictions from these models and compares them with observations. If the model does not compare well with the observations then one changes the model, typically either by changing the parametrisations of the previously physical processes or by adding new physical processes. These approaches have gained a lot of success, especially in understanding of the different forms of feedback processes (Croton et al., 2006; Oppenheimer et al., 2010). However, because of the increasing complexity in the theorectical models and a lack of robust knowledge on the fundamental physics of the cosmic gas, it does not seem to be an easy task to identify the missing physical processes or to test a particular prescription using this "forward" approach. A useful supplementary to the traditional modeling methods would be a "backward" approach. In this paper, what we do is to start form the data and attmept to let the data speak for themselves in a way that is as independent as possible of any model assumptions.

First we start with the recent measurement of SMFs at multiple redshifts and the luminosity function of cluster galaxies to infer how the SFR and the stellar mass are

related to mass of the host halos and how such relation evolves with time. We parameterize the SFR as a function of halo mass and time. The parameterization is constructed using Bayesian model selection technique to achieve good match to the observational constraints and avoid overfitting. The model predictons on different statistics of the galaxy population are detailed in §3. The predictions include i) the SMFs of galaxies of different types (star forming and quenched), in different environments and at different redshifts; ii) the two point correlation function of present-day galaxies; iii) the cosmic star formation history. Some of these predictions are tested against observations not included in the inference and others can be tested against future observations. The star formation history and stellar mass assembly history of present galaxies are presented in §4. Specifically the dependence on stellar (halo) mass and the contribution of *in situ* star formation and galaxy mergers are characterized in details.

The star formation - halo mass relations can be combined with other observational constraints to further infer the physics that regulates the galaxy evolution. The mass of classic bulges, which are believed to form through galaxy mergers, can be estimated from the galaxy merger histories. With the empirical relation between black hole mass and the classic bulge mass, we study the coevolution of galaxies and their central black holes and discuss how the formation of central black holes in dwarf galaxies may impact on the evolution of their hosts (§5). In §6, we develop a simple model for the galactic ecosystem, which includes the gas content (both atomic and molecular), inflow and outflow of gas and the enrichment of metals. Using current observational data on the gas mass fraction of local galaxies (Peeples and Shankar, 2011; Papastergis et al., 2012) and the evolution of gas phase metallicity-stellar mass relation (Maiolino et al., 2008), together with the empirically constrained star formation histories, we estimate the rates of gas inflow, gas outflow and metal outflow. In §7, we study the quenching of satellites by intra-halo environments. The stellar masses, SFRs and gas contents of the progenitors of satellite galaxies at the time of accretion are set according to the results in previous chapters,

which are obtained from a broad range of observations. This allows us to isolate out the problem of satellite galaxies from the whole galaxy evolution problem. The subsequent evolution of satellite galaxies is modeled with physically motivated processes and tested by comparing the prediction with results obtained from groups of galaxies.

CHAPTER 2

THE EMPIRICAL MODEL

2.1 Introduction

Observationally stellar mass is one of the best determined properties of galaxies. Theorectically the formation of stars links the physics occurs on super-galactic scales, such as formation of dark matter halos and accretion of baryonic matter, and sub-galactic scales such as origin of the galactic wind. However, physically motivated models have lagged behind observations, which are pushing to higher redshift and fainter galaxies. As discussed in the first chapter, it is difficult to explain the evolution of SMFs with simulations and semi-analytic models.

In recent years, much effort has been made to establish the statistical connection between galaxy stellar mass (luminosity) and dark matter halos via the conditional luminosity function (e.g. Yang et al., 2003; van den Bosch et al., 2003) or the halo occupation distribution (e.g. Jing et al., 1998; Peacock and Smith, 2000; Seljak, 2000; Scoccimarro et al., 2001; Berlind and Weinberg, 2002). Such empirically established galaxy-dark matter halo connections describe how galaxies with different properties occupy halos of different masses and, therefore,provide important insights into how galaxies form and evolve in dark matter halos (Mo et al., 2010). With data obtained from deep, multi-wavelength surveys, attempts have been made to establish the relation between galaxies and their dark matter halos out to high z using Abundance Matching (AM). This technique links galaxies of a given luminosity or stellar mass to dark matter halos of a given mass by matching the observed abundance of the galaxies to the halo abundance obtained from the halo mass function, typically also accounting for subhalos. This approach was first used by Mo

and White (1996) to model the number density and clustering of Lyman-break galaxies. More recently, several studies have used this abundance matching technique to probe the galaxy-dark matter connection out to $z \approx 5$ (e.g. Vale and Ostriker, 2004; Conroy et al., 2006; Shankar et al., 2006; Conroy and Wechsler, 2009; Moster et al., 2010; Guo et al., 2010; Behroozi et al., 2010; Yang et al., 2012; Béthermin et al., 2013; Behroozi et al., 2013a). One can infer the average SFR for halos of different masses at different redshifts from the stellar mass-halo mass relation (e.g. Behroozi et al., 2013b; Moster et al., 2013; Yang et al., 2013; Wang et al., 2013).

In this chapter, we present an empirical scheme to infer the connection between the SFR and halo mass and it evolution with time. The model is built upon dark matter halo merger trees. The SFR of the central galaxies is parameterized as a function of the halo mass and the redshift. A galaxy becomes a satellite its star formation is quenched exponentially and it can eventually merge with the central galaxy on a dynamical friction timescale. Compared with previous works in the literature, the model self-consistently takes into account the *in situ* star formation, galaxy mergers and the evolution of satellites. We put our model on a firm statistical footing using Bayesian inference. We start with a very simple model to describe the SFR of central galaxies and only increase its complexity if the data requires it, assessed using Bayes ratios. We build up our model in a stepwise manner, increasing the complexity as we add more constraining data, instead of using a single complicated model, so that we can clearly see how each SFR model is constrained by the different observations and to see whether or not a more complex model is required.

This chapter is organised as follows. The generic form of our empirical model is described in §2.2. In §2.3 the observational constraints, including the SMFs at different redshifts and the CGLF. In §2.4, we show how we increase the complexity of our model in a series of steps. This results in a series of nested model families that can reproduce more and more of the observational constraints: Model I is able to match the SMF of local galaxies, Model II reproduces the evolution of the SMFs, and Model III is the minimum

model that can also explain the CGLF. Finally, we discuss and summarise our results in §2.5.

2.2 The empirical of galaxy evolution

2.2.1 Dark halo merger histories

Our empirical model is built upon dark halo merger histories generated using the algorithm developed by Parkinson et al. (2008). The algorithm is based on the Extended Press-Schechter (EPS) formalism and it is tuned to agree with the conditional mass functions of merger trees (Cole et al., 2008) constructed from the Millennium Simulation (MS, Springel et al., 2005b). As shown by Jiang and van den Bosch (2014), this algorithm is in good agreement with simulations in terms of many other properties, such as mass assembly history, merger rate and unevolved subhalo mass function. For this work, the merger trees span a redshift range $0 \leq z \leq 15$ with 100 snapshots evenly distributed in $\ln(1+z)$ space. The mass resolution is $2 \times 10^9\, h^{-1}\mathrm{M}_\odot$. The sets of merger trees used to predict the observational constraints are described in §2.3 in more detail.

2.2.2 Star formation in central galaxies

We assume that the SFR of the central galaxy of a halo at a given redshift z is completely determined by the virial mass of the host halo, $M_\mathrm{h}(z)$, and z. Then the SFR can be written as

$$\mathrm{SFR} \equiv \mathrm{SFR}\left(M_\mathrm{h}(z), z\right) \ . \tag{2.1}$$

The above equation describes an average among halos of a given mass at a given z. It ignores any variations in the SFR that owe to variations in the formation histories of individual halos of a given mass and any large-scale environmental effects. Note though that the $\dot{M}_\star[M_\mathrm{h}(z)]$ can still result in halos of the same $M_\mathrm{h}(z)$ having different $M_\star$ simply because of scatter in the halo assembly histories.

Guided by the observational demand that the star formation efficiency must be suppressed in both low and high mass halos (Yang et al., 2003), we assume the following form for the dependence of the star formation rate on redshift and halo mass:

$$\mathrm{SFR} = \mathcal{E}\frac{f_B M_\mathrm{h}}{\tau_0}\,(1+z)^\kappa\,(X+1)^\alpha\left(\frac{X+\mathcal{R}}{X+1}\right)^\beta\left(\frac{X}{X+\mathcal{R}}\right)^\gamma\,, \tag{2.2}$$

where $\mathcal{E}$ is an overall efficiency; f_B is the cosmic baryonic mass fraction; τ_0 is a dynamic timescale of the halos at the present day, set to be $\tau_0 \equiv 1/(10H_0)$; and κ is fixed to be $3/2$ so that $\tau_0/(1+z)^{3/2}$ is roughly the dynamical timescale at redshift z. We define the quantity X to be $X \equiv M_\mathrm{h}/M_\mathrm{c}$, where M_c is a characteristic mass and $\mathcal{R}$ is a positive number that is smaller than 1. Hence, in our model the SFR of a galaxy depends on its dark matter halo mass through a piecewise power law, with α, β, and γ being the three power indices in the three different mass ranges separated by the two characteristic masses M_c and $\mathcal{R}M_\mathrm{c}$:

$$\dot{M}_\star \propto \frac{M_\mathrm{h}}{\tau_0}\begin{cases} M_\mathrm{h}^\alpha & \text{if } M_\mathrm{h} \gg M_\mathrm{c} \\ M_\mathrm{h}^\beta & \text{if } M_\mathrm{c} > M_\mathrm{h} > \mathcal{R}M_\mathrm{c} \\ M_\mathrm{h}^\gamma & \text{if } M_\mathrm{h} \ll \mathcal{R}M_\mathrm{c}\,. \end{cases} \tag{2.3}$$

The simplest model is the one where all the model parameters, $\mathcal{E}$, α, β, γ, M_c and $\mathcal{R}$ are redshift-independent. In what follows, we will make more parameters redshift-dependent whenever the observational data demands it.

2.2.3 Star formation in satellite galaxies

For satellites, the SFR has to be modelled differently. As a dark matter halo gets accreted by a larger one, it becomes a subhalo and experiences environmental effects such as tidal stripping, galaxy harassment (Moore et al., 1996), and tidal disruption. The satellite galaxy associated with the subhalo may also be affected as it orbits in the

host halo. For example, the diffuse gas initially in the subhalo and the cold gas disk may get stripped by the ram pressure or tidal forces of the host halo. Consequently, the star formation in the satellite can be suppressed or even quenched. Indeed, observations clearly show that satellite galaxies have larger quenched fractions than centrals of the same stellar mass (e.g., Balogh et al., 2000; van den Bosch et al., 2008; Wetzel et al., 2012).

We use a simple τ model to describe the suppression of star formation in a satellite after it is accreted:

$$\dot{M}_{\star,st}(t) = \dot{M}_{\star}(t_a) \exp\left(-\frac{t-t_a}{\tau_{st}}\right) , \tag{2.4}$$

where t_a is the time when the satellite is accreted into its host, $\dot{M}_{\star}(t_a)$ is the SFR of the satellite galaxy at $t = t_a$, and τ_{st} is a time scale characterising the decline of the star formation. We adopt the following model for the characteristic time

$$\tau_{st} = \tau_{st,0} \exp\left(-\frac{M_{\star}}{M_{\star,c}}\right) , \tag{2.5}$$

where $\tau_{st,0}$ is the exponential decay time for a galaxy with a stellar mass of $M_{\star,c}$, with both of these being free parameters in our model. The choice of τ_{st} is motivated by the fact that massive galaxies tend to be more quenched in star formation than low mass galaxies, independent of environment (Peng et al., 2010; Wetzel et al., 2013). For central galaxies, this trend is naturally reproduced by assuming that in massive halos the star formation efficiency decreases with halo mass (as long as $\alpha < 0$). For satellites, however, such a halo-mass dependence of star formation efficiency does not work, because the host subhalo mass is expected to decrease with time owing to stripping.

2.2.4 Merging and stripping of satellite galaxies

A halo with mass $M_{\rm sat}$ accreted by a larger (primary) halo with mass $M_{\rm pry}$ will gradually sink towards the center of the primary halo owing to dynamical friction. Using N-body simulations, Boylan-Kolchin et al. (2008) found that the dynamical friction time

scale, $\tau_{\rm merger}$, depends on the mass ratio between the satellite halo and the host halo, as well as on the orbital parameters of the satellite halo η at the time of accretion:

$$\tau_{\rm merger} = 0.216 \frac{(M_{\rm pry}/M_{\rm snd})^{1.3}}{\ln(1+M_{\rm pry}/M_{\rm snd})} \exp\left(1.9\eta\right) \tau_{\rm dyn}\,, \tag{2.6}$$

where $\tau_{\rm dyn} = r_{\rm vir}/V_{\rm vir}$ is the dynamical time of the halo, with $r_{\rm vir}$ and $V_{\rm vir}$ being the virial radius and virial velocity of the halo, respectively; η is the orbital circularity, which is the ratio between the orbital angular momentum and the orbital angular momentum of a circular orbit with the same energy. Following Zentner et al. (2005), we assume η to have the following distribution,

$$P(\eta) \propto \eta^{1.2}(1-\eta)^{1.2}\,. \tag{2.7}$$

For each satellite galaxy at the time of accretion, we draw a value of η from this distribution and use it in Equation (2.6) to estimate a dynamical friction time scale.

We assume that a satellite galaxy merges with the central galaxy of the primary halo in a time $\tau_{\rm merger}$ after accretion. We further assume that only a fraction $f_{\rm TS}$ of the stellar mass of the satellite is added onto the central galaxy and the rest of its stellar mass becomes halo stars. In our model we treat $f_{\rm TS}$ as a free parameter.

2.2.5 Spectral synthesis and metal enrichment

Given a star formation model, we use the procedure described above, together with a halo merger tree, to predict the star formation and stellar mass assembly history of a given galaxy. For each galaxy, we track its star formation $\mathrm{SFR}\,(t)$ and mass assembly history $M_\star\,(t)$. At a time step t_i, we obtain the stellar mass following

$$\begin{aligned} M_\star(t_i) &= \int_0^{t_i} \mathrm{SFR}(t)\mathcal{R}\,(t_i - t)\,dt \\ &\quad + \sum_a f_{\rm TS} M_{{\rm s},\star}(t_i)\,, \end{aligned} \tag{2.8}$$

which takes into account both the passive evolution and accretion of satellite galaxies denoted by the subscript 's'. The summation is over all satellites. $\mathcal{R}\,(t_i - t)$ is the

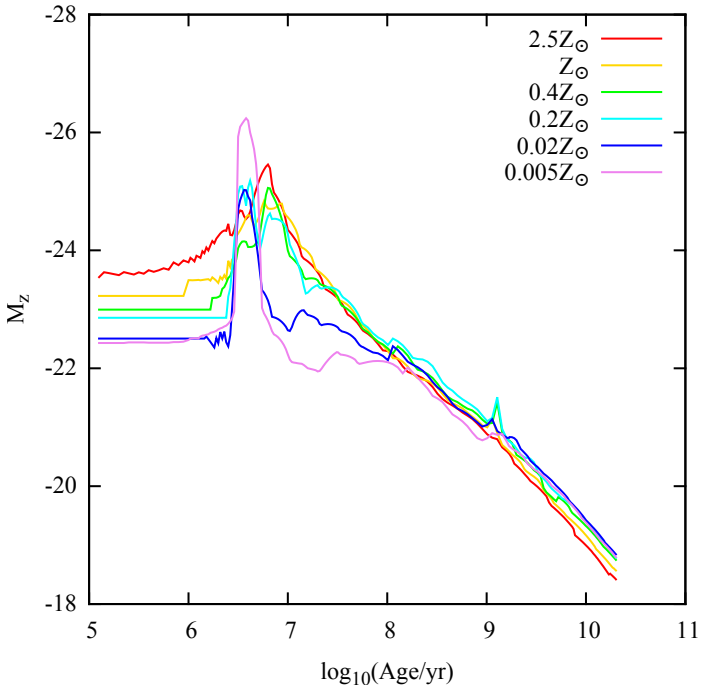

Figure 2.1. The z-band magnitude of a simple stellar population with initial mass $1\,\mathrm{M}_\odot$ as a function of age and metallicity.

remaining fraction of stellar mass for a stellar population of age $t_i - t$, and is adopted from the stellar evolution model of Bruzual and Charlot (2003) assuming a Chabrier IMF (Chabrier, 2003). The total SFR is simply the sum of all the progenitors

$$\mathrm{SFR}(t_i) = \mathrm{SFR}_\mathrm{c}(t_i) + \sum f_\mathrm{TS}\mathrm{SFR}_\mathrm{s}(t_i)\,, \tag{2.9}$$

where the central is denoted by a subscript 'c' and satellites by 's'.

The luminosity can be calculated in the same way as the stellar mass by replacing $\mathcal{R}(t_i - t)$ with the luminosity of a simple stellar population $\mathcal{L}(t_i - t)$ starting with $1\,\mathrm{M}_\odot$. $\mathcal{L}$ depends not only on the age of the stellar population but also on the metallicity, which can only be obtained by solving a full chemical evolution model. Since our model does not include the gas component in galaxies, we cannot trace the chemical evolution of stars directly. Instead, we use the metallicity - stellar mass relation observed for local galaxies

at all redshifts. We adopt the mean relation based on the data of Gallazzi et al. (2005), which can roughly be described as

$$\log_{10} Z = \log_{10} Z_{\odot} + \frac{1}{\pi} \tan \left[\frac{\log_{10}(M_{\star}/10^{10} M_{\odot})}{0.4} \right] - 0.3 \,. \tag{2.10}$$

This observational relation extends down to a stellar mass of $10^9 M_{\odot}$ and has a scatter of 0.2 dex at the massive end and of 0.5 dex at the low mass end. Fortunately, the z-band luminosity, which we use to compare with observations, depends only weakly on metallicity as shown in Figure 2.1. Hence, our results are expected to be insensitive to the exact chemical evolution.

2.3 Observational Constraints and Likelihood Functions

2.3.1 The SMFs of field galaxies

The SMF of the local Universe ($z \approx 0$) is from Baldry et al. (2012). The galaxies are selected from the SDSS DR6 down to a magnitude limit of $m_r \approx 19.8$ with a sky coverage of 143 deg^2. Redshift measurements for galaxies fainter than the SDSS redshift survey limit are made with the Galaxy And Mass Assembly (GAMA) survey. Stellar masses for individual galaxies are estimated from their *ugriz* photometry using a spectral synthesis model with the assumption of a Chabrier (2003) IMF. Owing to surface brightness incompleteness, the measured number density for galaxies with masses below $10^8\, h^{-2}\mathrm{M}_{\odot}$ can only be considered as lower limits. In our analysis we, therefore, only use data points above $10^8\, h^{-2}\mathrm{M}_{\odot}$.

In this paper we use two sets of SMFs recently published. The first is that of Santini et al. (2012b) based on the Early Release Science (ERS) data of the Wide Field Camera 3 (WFC3) in the GOODS-S Field, and the SMFs for galaxies in the redshift range between 0.6 and 4.5 are estimated to a stellar mass limit of a few times $10^9\,\mathrm{M}_{\odot}$. The second set is that of Tomczak et al. (2014), who estimated the galaxy SMFs in the redshift range 0.2 - 3.0 using data from the FourStar Galaxy Evolution Survey (ZFOURGE) and the Cosmic

Assembly Near-IR Deep Extragalactic Legacy Survey (CANDELS). Unfortunately, the two sets are not completely consistent with each other. While at $z < 2.0$ the results obtained by both are similar, at $z > 2$ the SMFs at the low-mass ends obtained by Tomczak et al. (2014) are lower than those obtained by Santini et al. (2012b) by a factor of 2. We use the two sets of data separately to construct the models, and to check the reliability of our results against uncertainties in the observational data. In the following sections, the results shown are based on the data of Santini et al. (2012b) unless otherwise indicated.

To make predictions for the field galaxy SMF, we use 5,000 dark matter halos drawn from a power law distribution $f'(M_{\rm h}) \propto [M_{\rm h}/(\,h^{-1}{\rm M}_{\odot})]^{-1.5}$ in the mass range $5{\times}10^{9}h^{-1}{\rm M}_{\odot} < M_{\rm h} < 5 \times 10^{14}h^{-1}{\rm M}_{\odot}$. Including halos outside this mass range does not change our results significantly. The fact that our merger trees do not resolve progenitor halos with $M_{\rm h}(z) < 2 \times 10^{9}\,h^{-1}{\rm M}_{\odot}$, implies that we implicitly assume that star formation is suppressed in halos with masses below this 'threshold'. In other words, our star formation model (Eq. [2.2]) should be augmented with $\dot{M}_{\star} = 0$ for $M_{\rm h}(z) < 2{\times}10^{9}\,h^{-1}{\rm M}_{\odot}$. When we calculate the stellar mass/luminosity function of galaxies, the predicted number of galaxies associated with a halo of mass $M_{\rm h}$ is assigned a weight $\omega = f_{\rm SMT}(M_{\rm h})/f'(M_{\rm h})$ where $f_{\rm SMT}(M_{\rm h})$ is the halo mass function from Sheth et al. (2001) for our adopted cosmology.

2.3.2 The composite luminosity function of cluster galaxies

As an additional constraint, we also use the z-band composite luminosity function of cluster galaxies obtained by Popesso et al. (2006). The z-band luminosities are less affected by dust extinction than their bluer counterparts, making the comparison between our model and the data less affected by dust corrections. Furthermore, for a stellar population older than 100 Myr the predicted z-band luminosity depends only weakly on metallicity, making our results less sensitive to the chemical evolution model described above (see Figure 2.1). Finally,the use of the composite luminosity function, a weighted

average over a number of individual clusters, reduces statistical uncertainties arising from variances in the formation history of clusters.

The composite luminosity function of Popesso et al. (2006) is based on 69 clusters selected from ROSAT X-ray data using cross identifications with galaxies in the SDSS. Once the luminosity functions of individual clusters are known the composite luminosity function can be evaluated using

$$N_{\rm cj} = \frac{N_{\rm c0}}{m_{\rm j}} \sum_i \frac{N_{\rm ij}}{N_{\rm i0}} \,, \tag{2.11}$$

where $N_{\rm cj}$ is the number of galaxies in the jth bin of the composite luminosity function, $N_{\rm ij}$ is the number of galaxies in the jth bin contributed by the ith cluster, $N_{\rm i0}$ is the normalisation factor for the ith cluster, which is the total number of the member galaxies brighter than the magnitude limit $m_{\rm lim}$ at the cluster redshift, $m_{\rm j}$ is the number of clusters contributing to the jth bin and

$$N_{\rm c0} = \sum_i N_{\rm i0} \,. \tag{2.12}$$

We follow the same method to calculate the composite luminosity functions in our model predictions.

To use the CGLF as a constraint it is necessary to know the masses of those clusters accurately, since a systematic error in cluster mass can lead to an error in the amplitude of the predicted CGLF. Unfortunately, the mass estimates for the clusters in the Popesso et al. (2006) sample are uncertain, and so it is dangerous to use the overall amplitude of the CGLF to constrain the model. To bypass this problem, we treat the ratio between the real mass of a cluster $M_{\rm real}$ and the measured value $M_{\rm obs}$,

$$e_{\rm M} \equiv \frac{M_{\rm real}}{M_{\rm obs}} \,, \tag{2.13}$$

as a free parameter, and we renormalise the luminosity function of an individual cluster by $e_{\rm M}$ (i.e., we marginalise over potential systematic biases in the inferred cluster masses).

To make predictions for the CGLF, we generate 23 merger trees for halos with a present-day mass distribution similar to that observed. The mass resolution adopted for these merger trees is also $2 \times 10^9\, h^{-1}\mathrm{M}_\odot$. We confirmed that this number of halos is sufficiently large so that the variance between the different realisations is smaller than the uncertainties in the observational data.

2.3.3 The Likelihood Function and Sampling Algorithm

We explore a series of nested model families, based on the generic parametrisation given by Eq. 2.2. At each step we increase the complexity of the model and check whether the fit to the observational constraints is acceptable and whether any improvement to the fit is sufficient to justify the increased complexity by comparing the posterior predictions with the constraints. In addition, we also make use of the Bayes Factor to avoid developing an overcomplicated model. In Bayesian statistics, the Bayes Factor is

$$K = \frac{p(M_\mathrm{a}|D)}{p(M_\mathrm{b}|D)}\,, \tag{2.14}$$

where $p(M|D)$ is the probability of model M given data D. It is obtained by marginalising over all the model parameters of the posterior distribution. The advantage of the Bayes factor is that it automatically includes a penalty for too much complexity in the model. The values of $\ln\left[p(M|D)\right]$ for all the models discussed in the text are listed Table 2.1.

The posterior distribution of the model parameters is proportional to the likelihood function, the probability of the data given a set of parameters. A rigorous likelihood function is supposed to include all the processes in the data acquisition. The data acquisition process includes deriving a stellar mass from multiband photometry using a stellar population synthesis model, correcting for incompleteness, weighting each galaxy sample according to its corresponding survey volume, and so on. As a result, the uncertainties in individual stellar mass bins are not independent. As is demonstrated in Lu et al. (2011a), the covariance for binned data may change the posterior distribution substantially. Un-

fortunately, the covariance matrix in the data used here is not available and we have to assume that the SMF in different bins is independent and that the likelihood is approximated by a Gaussian function. with these assumptions the likelihood function can be written as

$$\ln L(\Phi_{\rm obs}|\theta) = C - \frac{1}{2}\sum_i \frac{\left[\Phi_{i,\rm obs} - \Phi_{i,\rm mod}(\theta)\right]^2}{\sigma^2_{i,\rm obs}}, \tag{2.15}$$

where C is the normalization factor.

To efficiently sample the high dimensional parameter space, we make use of the MULTINEST method developed by Feroz et al. (2009), which implements the nested sampling algorithm of Skilling (2006). We have compared this method with the MCMC implemented in Lu et al. (2011a), Tempered Differential Evolution. Both methods are designed to deal with probability distributions with multiple modes and strong degeneracies between model parameters. For the problem in this paper, we found the two give identical results, but that the number of likelihood evaluations required by MULTINEST is smaller by more than a factor of 10. A more detailed description of the method is given in Appendix A.

2.4 Construct the Parametric Models

2.4.1 The Simplest Model Family (Model I)

We start with the simplest case in which all the parameters in Equation (2.2) are assumed to be time-independent, so that the SFR is determined completely by the halo mass and by the dynamical time scale of the halo at the time in question. We call this Model I. First, we use only the SMF at $z \approx 0$ Baldry et al. (2012) to constrain the model parameters. Figure 2.2 shows that Model I fits the constraining data well (see the first panel). However, as one can see from the other panels of Figure 2.2, the posterior of the model predicts too few massive galaxies at high z compared to the observational data. Next, we use the SMFs at all redshift bins, to constrain the model parameters. The resulting fits to the data are shown in Figure 2.3. This model family still fails to

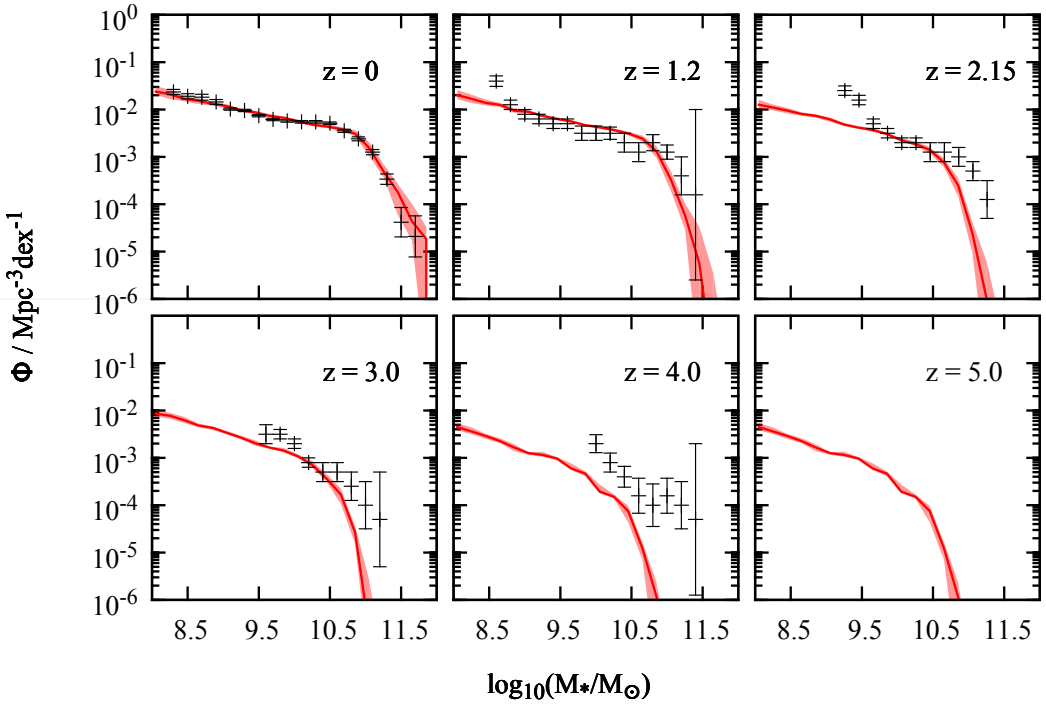

Figure 2.2. The posterior predicted SMFs of Model I constrained by the SMF at $z \approx 0$. The red bands encompass the 95% credible intervals, and the red solid lines are the medians.

Table 2.1. Summary of the posterior simulations, including the models, the observational constraints and the marginalised likelihood.

Model	Constraints	Marginalised Likelihood (natural log)
Model I	SMF($z = 0$)	-22.2
Model I	SMF	-120
Model II	SMF($z = 0$)	-21.0
Model II	SMF	-31.2
Model II	SMF + CGLF	-89.7
Model IIb	SMF	-55.5
Model III	SMF	-30.6
Model III	SMF + CGLF	-63.3
Model IV	SMF + CGLF	-58.4

reproduce the SMFs at the bright end: it underestimates the number density at $z > 1$. This suggests that massive galaxies form too late, and that the SFR at high-z in massive halos needs to be boosted.

2.4.2 Fixing the Bright-End Problem (Model II)

As an attempt to fix the bright-end problem identified above, we consider a second model family (Model II) that allows α to be redshift-dependent. We assume that the redshift-dependence be given by the following power law,

$$\alpha = \alpha_0 (1 + z)^{\alpha'} , \tag{2.16}$$

where both α_0 and α' are introduced as new free parameters. The SFR in massive halos with $M_{\rm h} > M_{\rm c}$ will be enhanced at high-z if α' is a negative number.

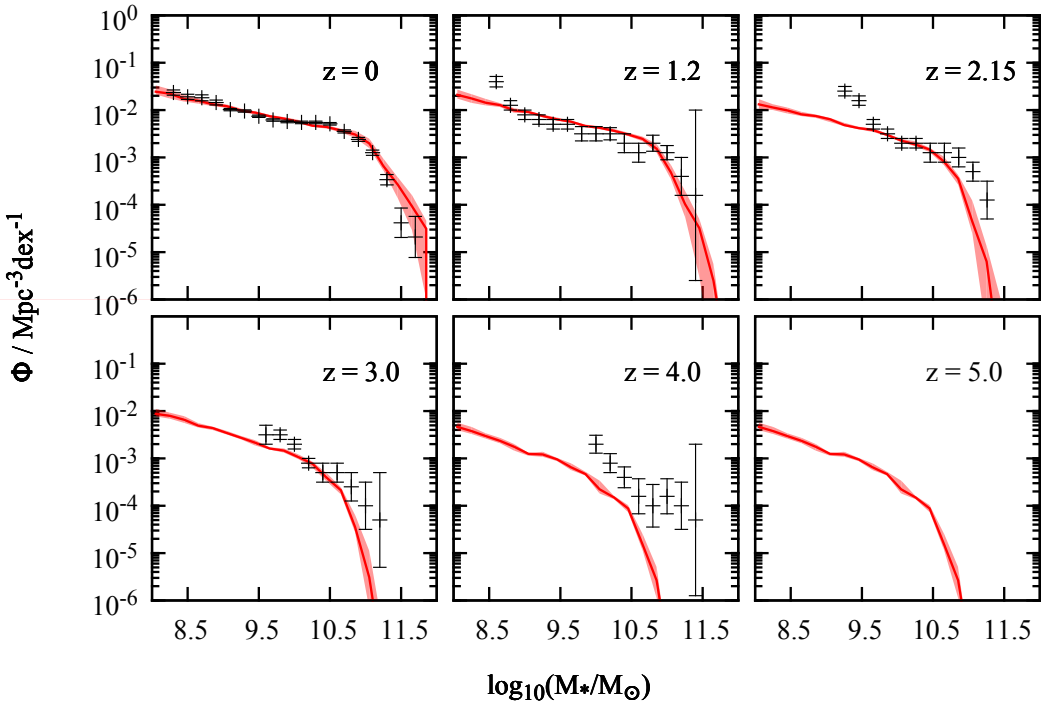

Figure 2.3. The same as Figure 2.2 but for Model I constrained using SMFs at all the redshift bins.

Table 2.2. The constrained model parameters of Model II, in terms of the means and the variances. The observational constraints used are listed in the first row. B12 is for Baldry et al. (2012), S12 for Santini et al. (2012a) and T14 for Tomczak et al. (2014). $M_{\rm c}$ is in units of $10^{10}\,h^{-1}{\rm M}_{\odot}$, and $M_{*,{\rm c}}$ is in units of $10^{10}\,h^{-2}{\rm M}_{\odot}$.

	SMF ($z \approx 0$,B12) SMF ($z \geq 1$, S12)	SMF ($z \approx 0$, B12) SMF ($z \geq 1$, T14)
Parameter	mean$\pm\sigma$	mean$\pm\sigma$
α_0	-3.7 ± 0.82	-3.4 ± 0.84
α'	-0.46 ± 0.11	-0.45 ± 0.11
β	3.4 ± 0.86	2.6 ± 0.99
γ	0.89 ± 0.63	1.3 ± 0.77
$\log_{10}(M_{\rm c})$	1.7 ± 0.13	1.8 ± 0.18
$\log_{10}(\mathcal{R})$	-1.1 ± 0.34	-1.1 ± 0.45
$\log_{10}(\mathcal{E})$	0.30 ± 0.27	0.017 ± 0.27
$\log_{10}(H_0\tau_{\rm sat,0})$	-0.98 ± 0.17	-0.85 ± 0.13
$\log_{10}(M_{*,{\rm c}})$	0.81 ± 0.42	0.84 ± 0.38
$f_{\rm TS}$	0.36 ± 0.17	0.38 ± 0.15

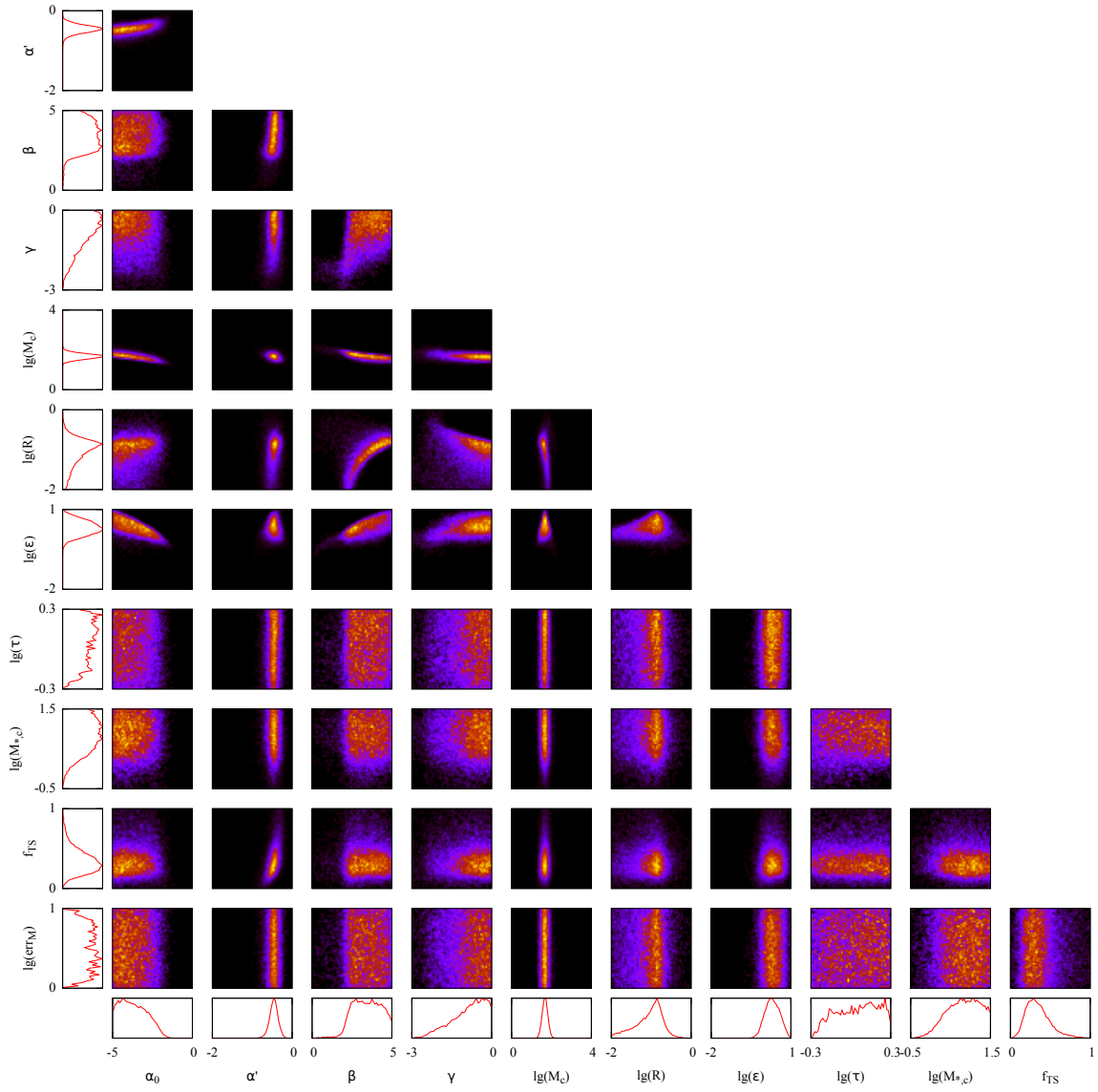

Figure 2.4. The posterior distribution of the parameters of Model II constrained using the SMFs at different redshift bins.

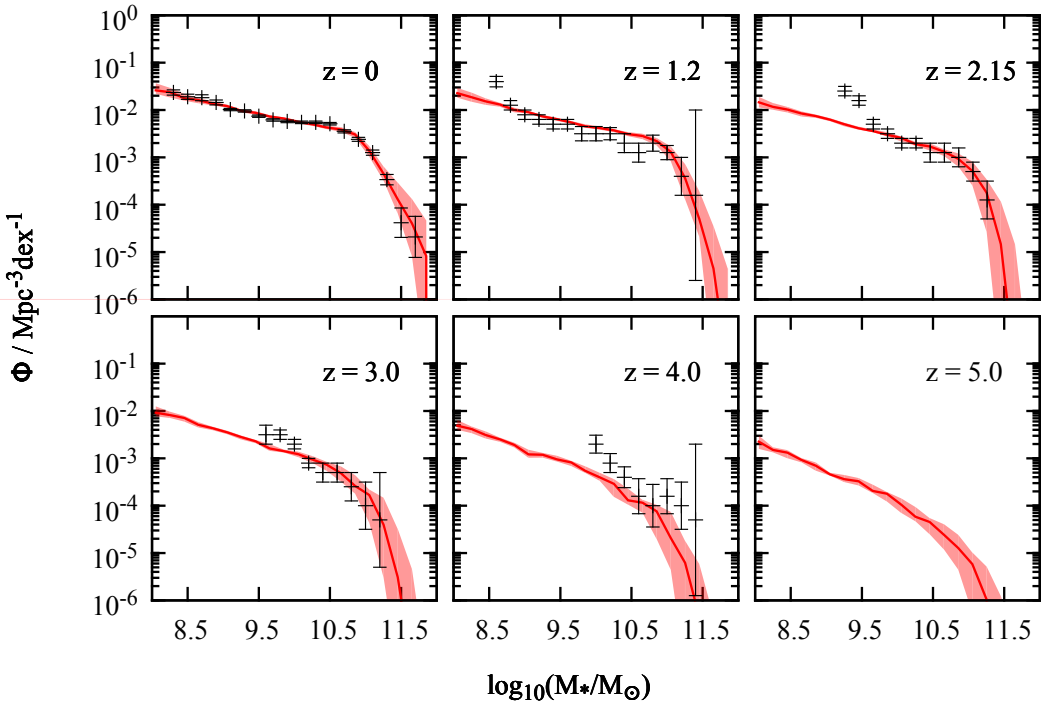

Figure 2.5. The same as Figure 2.2 but for Model II constrained using the SMFs at different redshift bins.

Once again we use the four SMFs as observational constraints. Figure 2.5 compares the posterior predictions with the constraining data. This model family matches the observational data over the entire redshift range from $z = 0$ to $z = 4$. The natural log of the Bayes ratio between Model II over Model I (constrained by the same data sets) is $-31.2 - (-120) = 88.8$, making the odds of preferring Model II over Model I given the data a whopping $e^{88.8}$ to one (Table 2.1). The credible intervals for the parameters are listed in Table 2.2. α' is -0.54 ± 0.11, which means significant evolution in the SFR of massive central galaxies.

Figure 2.6 shows the SFR of constrained Model II The SFR peaks at around $10^{12}\,h^{-1}\mathrm{M}_{\odot}$ at each redshift. In halos with low masses the SFR increases rapidly with halo mass, roughly as SFR $\propto M_{\mathrm{h}}^{2.5}$, quite independent of redshift. The SFR decreases with halo mass for halos above $10^{12}\,h^{-1}\mathrm{M}_{\odot}$. The decrease is more pronounced at low redshifts, becomes weaker as one goes to higher redshifts, and is quite weak by $z = 4$. It should be

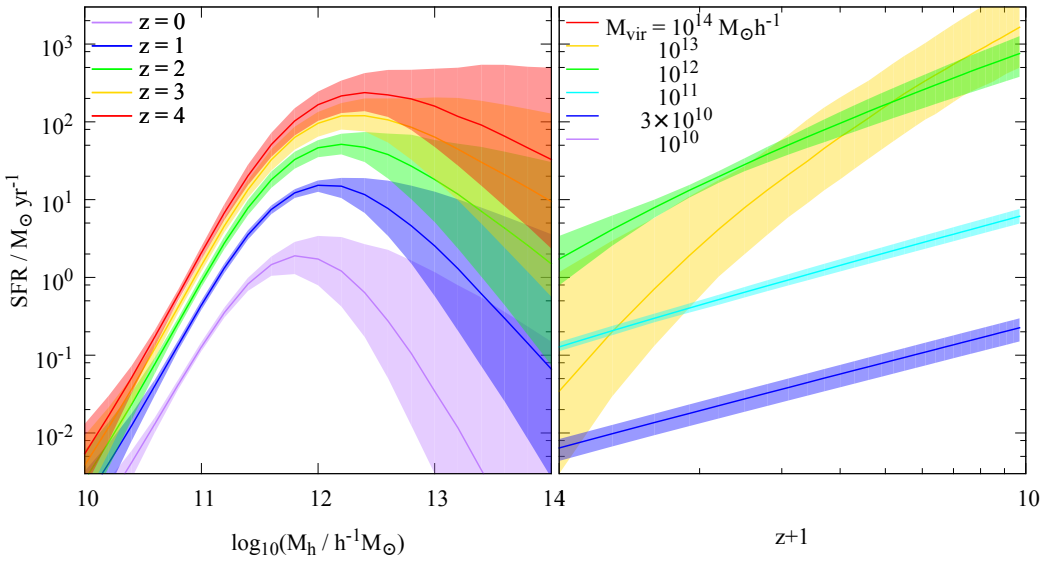

Figure 2.6. The SFR as a function of halo mass (left panel) and redshift (right panel) of Model II, constrained by the SMFs. The solid lines are the medians and the bands are the 95% credible intervals.

noted, however, that at $z \approx 4$ the SFR beyond $10^{13}\,h^{-1}\mathrm{M}_\odot$ can only be considered as an extrapolation because the number density of such high-mass halos is extremely small at such high redshifts. For halos with masses $\sim 10^{12}\,h^{-1}\mathrm{M}_\odot$, the SFR increases with redshift roughly as $(1+z)^{2.3}$. The increase with redshift is faster for more massive halos. For $M_\mathrm{h} < 10^{11}\,h^{-1}\mathrm{M}_\odot$ SFR $\propto (1+z)^{1.5}$.

In the following we examine how Model II matches the z-band CGLF. The posterior prediction is shown in the left panel of Figure 2.7. The normalisation of the prediction is adjusted to match the observations at $M_z - 5\log_{10}(h) \approx -20$ to account for a potential systematic bias in the cluster masses. The model prediction is consistent with the observational data at the bright end but it under-predicts the number of dwarf galaxies with $M_z - 5\log_{10}(h) > -17$. In particular, the predicted faint end of the luminosity function is roughly a power law, in contrast with the observational data that shows a significant upturn. To test whether the model family Model II can accommodate the observed luminosity function of cluster galaxies, we carry out a new inference, this time including the

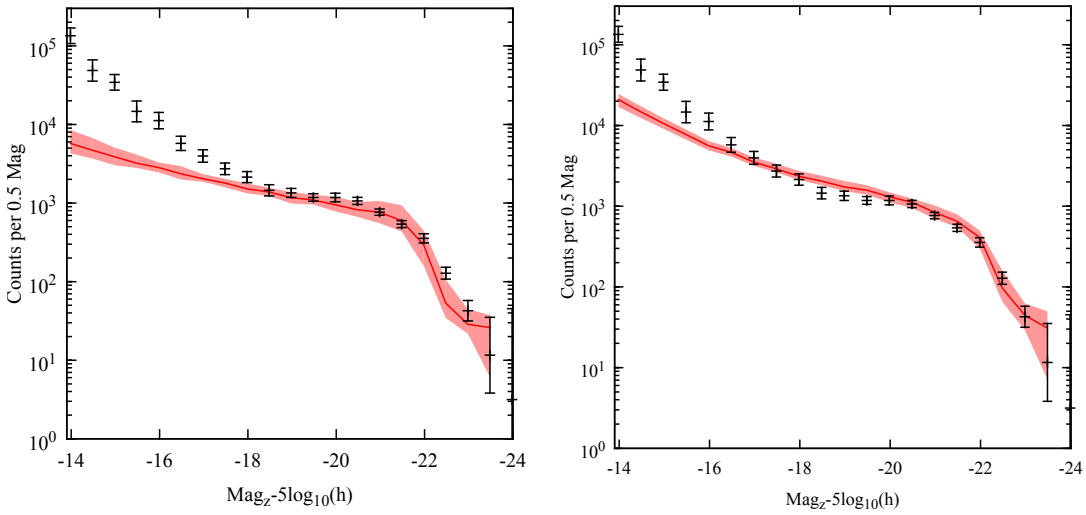

Figure 2.7. Posterior predicted CGLF of Model II. Left panel: The posterior prediction of Model II constrained by the SMFs compared with the CGLF of Popesso et al. (2006). The bands encompass 95% of the posterior distribution, and the red line is the median. The observational data are shown as data points with errorbars. Right panel: The same as the left panel but now Model II is constrained both by the SMFs and the observed z-band CGLF.

z-band luminosity function as an additional data constraint. The predicted CGLF still fails to match the observed one at the faint end, as shown in the right panel of Figure 2.7. Thus, it is unlikely to find a model in the parameter space of Model II to simultaneously match the SMFs and CGLF.

2.4.3 Fixing the faint end problem of the cluster galaxy luminosity function (Model III)

There are at least two possibilities that could lead to the differences between the SMF of field galaxies and the CGLF. Since the over-abundance of dwarf galaxies is more prominent in galaxy clusters, some environmental effects specific to high-density environments may cause the strong upturn. However, many processes such as tidal stripping and tidal disruption are destructive, which would suppress the number of satellite galaxies rather than enhance it. It is possible that relatively massive galaxies could have experienced significant mass loss, moving them to lower masses. However, since more massive galaxies

Table 2.3. The same as Table 2.2, but for Model III, and the cluster galaxy luminosity function (Popesso et al., 2006) is also used as a constraint.

	SMF ($z \approx 0$, B12) SMF ($z \geq 1$, S12)	SMF ($z \approx 0$,B12) SMF ($z \geq 1$, T14)
Parameter	mean$\pm\sigma$	mean$\pm\sigma$
α_0	-3.0 ± 1.0	-2.7 ± 0.85
α'	-0.36 ± 0.16	-0.37 ± 0.10
β	3.7 ± 0.73	3.9 ± 0.69
γ_a	2.0 ± 0.55	0.58 ± 0.39
γ_b	-0.84 ± 0.14	-0.90 ± 0.08
γ'	-4.4 ± 0.52	-4.2 ± 0.62
$z_{\rm c}$	1.8 ± 0.31	2.0 ± 0.38
$\log_{10}(M_{\rm c})$	1.6 ± 0.15	1.6 ± 0.13
$\log_{10}(\mathcal{R})$	-0.86 ± 0.18	-0.88 ± 0.17
$\log_{10}(\mathcal{E})$	0.20 ± 0.29	0.07 ± 0.27
$\log_{10}(H_0\tau_{\rm sat,0})$	-0.90 ± 0.16	-0.74 ± 0.04
$\log_{10}(M_{*,\rm c})$	0.34 ± 0.28	0.36 ± 0.17
$f_{\rm TS}$	0.44 ± 0.22	0.34 ± 0.19
$\log_{10}(e_{\rm M})$	0.15 ± 0.04	0.16 ± 0.03

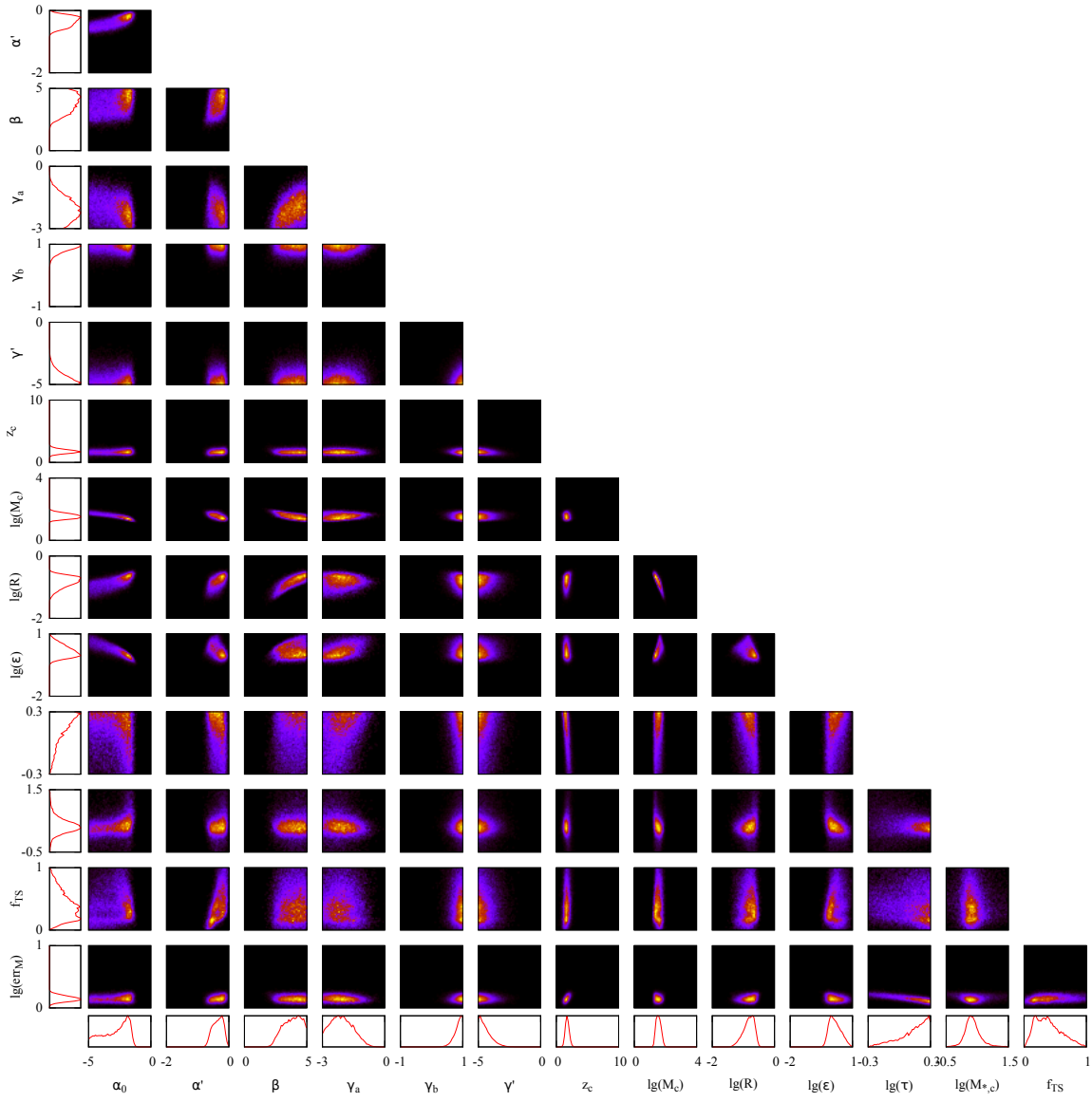

Figure 2.8. The posterior distribution of the parameters of Model III constrained by the SMFs and the CGLF.

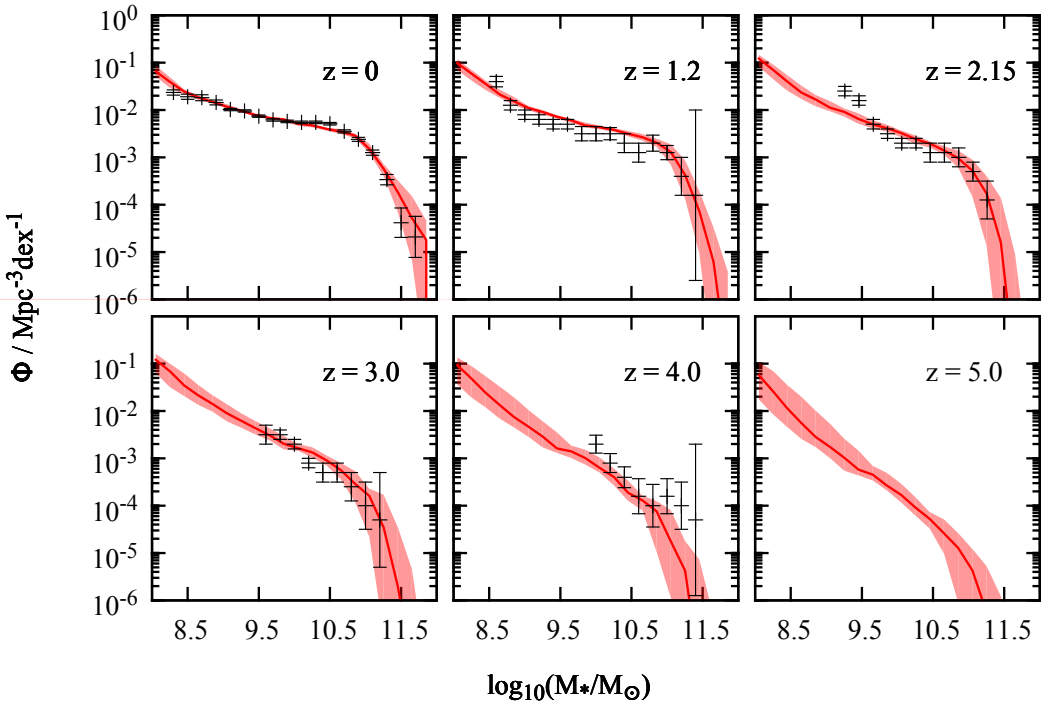

Figure 2.9. The same as Figure 2.2 but predicted by the posterior of Model III constrained by both the SMFs and the observed z-band CGLF.

are less abundant, it is difficult to make such a scenario work in detail. Thus, unless environmental effects in clusters operate in a way very different from what is generally believed, it is difficult to explain the over-abundance of dwarf galaxies in clusters with such effects.

Another possibility is that the environmental dependence of the luminosity function may be the result of some time-dependent processes. Indeed, galaxies in a cluster are expected to form early in their progenitor halos. Thus, if star formation in dark matter halos were different at high redshift when the majority of cluster galaxies formed, then the luminosity functions of cluster galaxies and field galaxies could show different behaviours owing to their systematically different formation times. One concrete example is the preheating model proposed in Mo and Mao (2002), where the intergalactic medium (IGM) is assumed to be preheated at some high redshift, so that the star formation in dark matter halos proceeds differently before and after the preheating epoch. Such preheating may

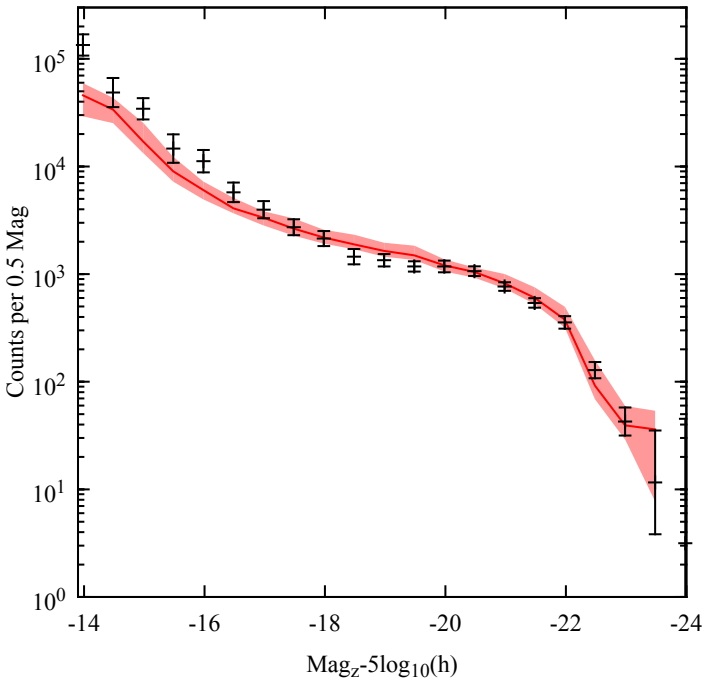

Figure 2.10. The CGLF predicted by the posterior of Model III constrained by both the SMFs and the observed z-band CGLF. The symbols have the same meaning as those in Figure 2.2.

owe to the formation of pancakes, as envisaged in Mo et al. (2005), owe to an episode of starbursts and AGN activity (Mo and Mao, 2002), or owe to heating by high-energy gamma rays generated by blasars, as envisaged in Chang et al. (2012). In all these cases, the preheating is expected to occur around $z \approx 2$ and the preheated entropy of the IGM is a few times $10\,\mathrm{KeV\,cm^2}$. In what follows, we consider a generic model family (Model III) inspired by the physical processes discussed above. We allow γ, which controls star formation in low-mass halos, to be time-dependent in such a way that it changes from γ_b at high-z to γ_a at low-z, with a transition redshift z_c. Specifically we assume that

$$\gamma = \begin{cases} \gamma_\mathrm{a} & \text{if}\, z < z_c \\ (\gamma_\mathrm{a} - \gamma_\mathrm{b}) \left(\frac{z+1}{z_c+1}\right)^{\gamma'} + \gamma_\mathrm{b} & \text{otherwise}\,. \end{cases} \tag{2.17}$$

Note that Model II is a special case of Model III, with $\gamma' = 0$. If $z_c = 0$ and $\gamma_\mathrm{b} = 0$ then γ is a simple power law of $(1 + z)$ with an index of γ'.

Before we add this to Model II, thereby creating Model III, we want to explore the possibility that adding this behaviour to Model I and not allowing the massive end slope α to depend on redshift might also provide a viable fit to the SMFs at the four redshifts. We refer to this as Model IIb. We perform an inference with Model IIb using just the SMFs as data constraints. As one can see from Table 2.1, the SMFs prefer Model II over Model IIb by a probability of $e^{24.3}$ to one, but they still prefer Model IIb over Model I.

Figures 2.9 and 2.10 compare the posterior predictions with the constraining data, which are the SMFs and the CGLF. We see that Model III can accommodate both observational data sets. In particular, the upturn in the faint end of the CGLF is well reproduced (see Figure 2.10). The two data sets prefer Model III over Model II by a probability of $e^{26.4}$ to one (Table 2.1).

The posterior model parameters obtained for Model III are listed in Table 2.3. As one can see, except for γ, the values of all the other parameters obtained from Model III are quite similar to those obtained from Model II. For the new model parameters introduced

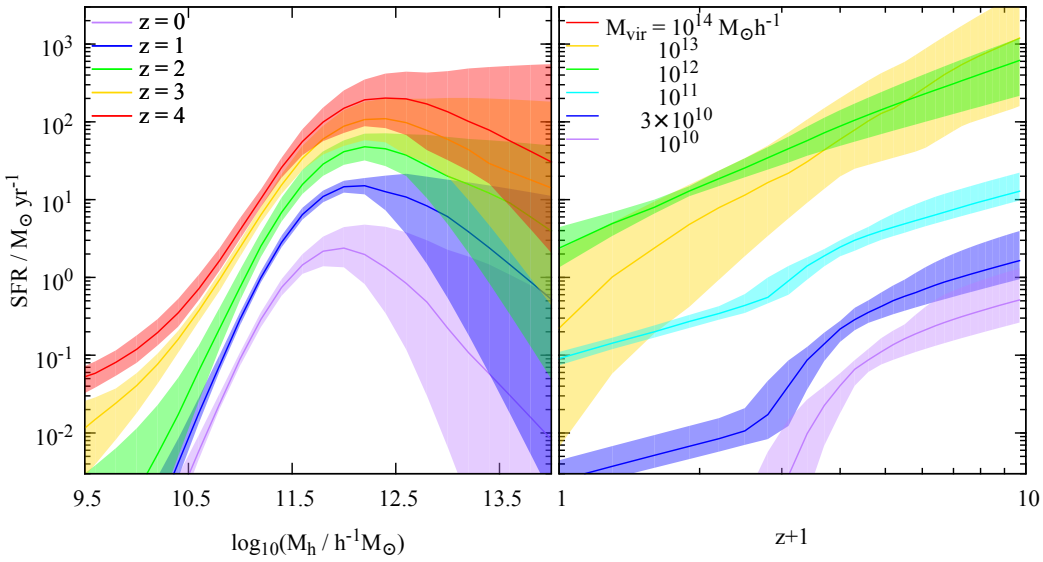

Figure 2.11. The SFR as a function of halo mass (left panel) and redshift (right panel) predicted by Model III. The solid lines are the medians of the posterior predictions and the bands are the 95% credible intervals.

in Model III, we have $\gamma' = -4.4$, $z_c = 1.8$, $\gamma_a = 2.0$ and $\gamma_b = -0.84$. The significant difference of γ' from zero implies that a redshift-dependent γ is preferred by the data, and the fact that γ_a is much larger than γ_b indicates that the SFR increases with halo mass much faster at low redshift ($z \ll z_c$) than at $z \gg z_c$. The value of z_c is constrained to 1.8±0.3 and, interestingly, is very close to the value expected from the preheating scenarios mentioned above. We will come back to discuss further the implications of these results. lso listed in the table are the parameters contrained by the data of Tomczak et al. (2014). The newly introduced parameters, including α', γ' and z_c are well constrained and the values are close to those constrained by the data ofSantini et al. (2012b). This suggests that our major conclusions are robust as to the uncertainties in the current data.

Figure 2.11 shows the model prediction of Model III for the SFR as a function of halo mass and redshift. For halos more massive than $10^{11}M_\odot$, the results are almost identical to those given by Model II. For less massive halos, however, Model III predicts a clear transition at $z_c \approx 2$ from a phase of elevated star formation at higher z to a phase of

reduced star formation at low z, as can be seen in the right panel of Figure 2.11. This behaviour owes to the adding of the cluster galaxy luminosity function data instead of the use of the more extended Model III. When using only the four SMFs as constraints, the less restrictive model III is only preferred over Model II by a factor $e^{0.6} \approx 1.8$, and the posterior for the SFR is similar to that of Model II shown in Figure 2.6. However, when including data constraints from the cluster luminosity function, the Bayes factor increases to $e^{26.4}$.

2.4.4 Are more general model families necessary?

As shown above, Model I successfully matches the observed $z \approx 0$ SMF, Model II successfully matches the observed galaxy SMFs over the redshift range from $z = 0$ to $z = 4$, and Model III successfully matches not only the observed galaxy SMFs but also the z-band CGLF at $z \approx 0$. Perhaps the observational data could be fit even better with a more complex model? Is the resulting SFR as a function of halo mass and redshift unique or are there other models that could match the observational data equally well but predict a $SFR(M_{\rm h}, z)$ that is very different from that predicted by the previous models? To investigate these questions we performed three more inferences. First, using only the $z \approx 0$ SMF as data constraint for Model II, Table 2.1 shows that this data only marginally prefers the more complex Model II over Model I and hence Model I is sufficient to describe the $z \approx 0$ SMF. Next, we performed an inference with Model III using the four observed galaxy SMFs as data constraints only. Again, Table 2.1 shows that these data only modestly prefer the more complex Model III over Model II, making Model II sufficient to describe the SMFs over the range $z = 0$ to 4.

Finally, to test the robustness of the Model III predictions using both the observed galaxy SMFs and also the CGLF, we consider another model family, Model IV, which allows even more model parameters to be redshift-dependent. Specifically, we write

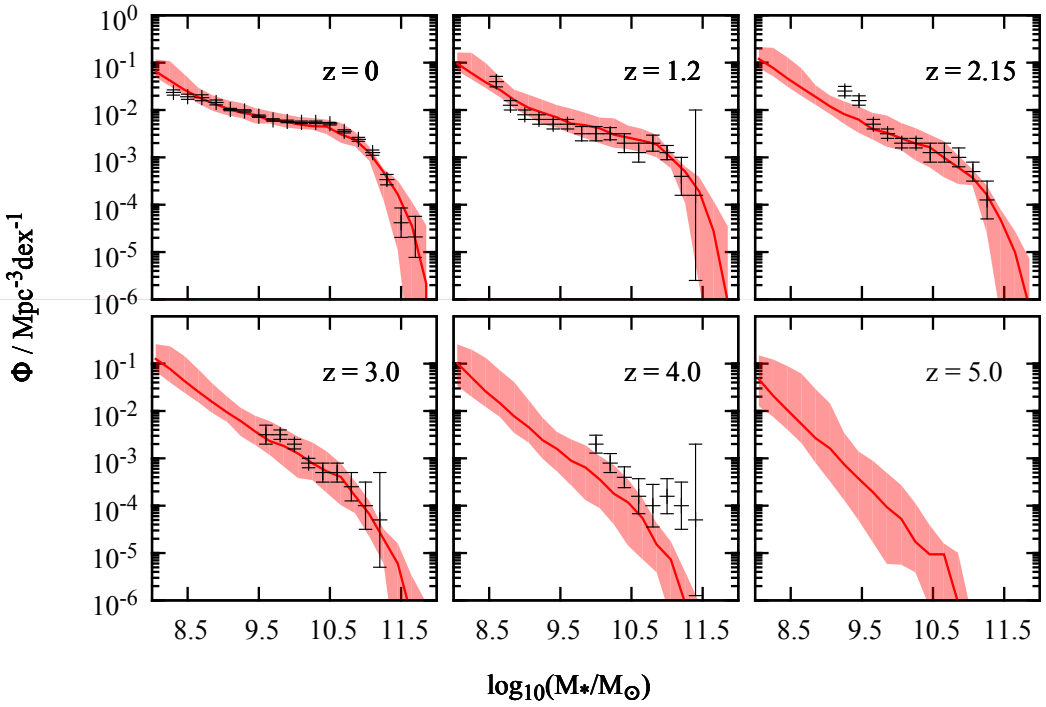

Figure 2.12. The galaxy SMFs predicted by the posterior of Model IV constrained by both the SMFs and the observed z-band CGLF. The red bands encompass the 95% credible intervals, and the red solid lines are the medians.

$$\gamma = \begin{cases} \gamma_{\rm a} \left(\frac{z+1}{z_c+1}\right)^{\gamma''} & \text{if } z < z_c \\ (\gamma_{\rm a} - \gamma_{\rm b}) \left(\frac{z+1}{z_c+1}\right)^{\gamma'} + \gamma_{\rm b} & \text{otherwise}\,; \end{cases} \tag{2.18}$$

$$M_{\rm c} = M_{\rm c,0}(z+1)^{\mu}\,; \tag{2.19}$$

$$\mathcal{R} = \mathcal{R}_0(z+1)^{\rho}\,; \tag{2.20}$$

$$\kappa = \frac{3}{2} + \kappa'\,, \tag{2.21}$$

with γ'', μ, ρ, κ' introduced as four new free parameters. Model IV reduces to Model III if these four parameters are set to be zero.

In Figures 2.12 and 2.13 we show the posterior predictions of Model IV for the galaxy SMFs and CGLF respectively, compared with the corresponding constraining data. As expected, the larger parameter number Model IV fits the constraining data better. In terms of the Bayes Factor, the ratio between Models IV and III is $e^{4.9} \approx 134$, which

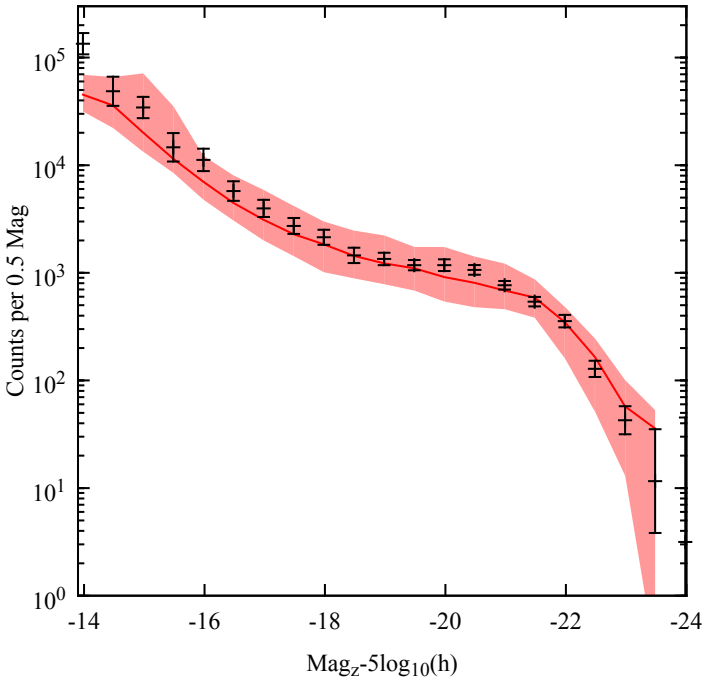

Figure 2.13. The CGLF predicted by the posterior of Model IV constrained by both the SMFs and the z-band CGLF.

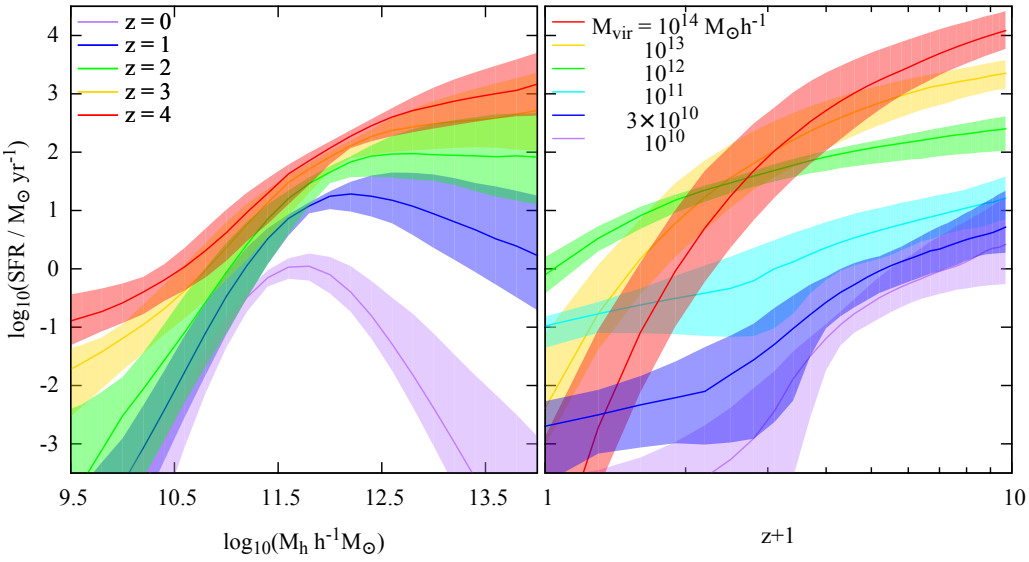

Figure 2.14. The SFR as a function of halo mass (left panel) and redshift (right panel) predicted by Model IV. The solid lines are the medians of the posterior prediction and the bands are the 95% credible intervals.

represents a marginally significant improvement. An improvement in the fit of the $z \approx 0$ SMF near the knee is evident, but no other significant improvements are noticeable. Furthermore, the SFR as a function of halo mass and redshift predicted by Model IV is qualitatively similar to that predicted by Model III, as shown in Figure 2.14. The only significant difference is that the star formation in massive halos with $M_{\rm h} > 10^{12.5}\,h^{-1}{\rm M}_{\odot}$ predicted by Model IV increases faster with increasing redshift than that predicted by Model III. Thus, the SFR as a function of halo mass and redshift predicted by Model III does not seem to owe to the particular parameterisations adopted but rather reflects requirements of the observational data. Unfortunately, such tests can never be exhaustive; there is always the possibility that some other model could match the data constraints and yet give a different SFR($M_{\rm h}, z$). In this sense, all our conclusions and predictions are restricted to the model families that we actually explore.

Since Model IV and Model III make similar predictions for the star formation histories for halos with different masses, we will base our following presentation on Model III.

2.5 Summary and Discussion

We use the observed SMFs of galaxies in the redshift range from $z \approx 0$ to $z \approx 4$ and the luminosity function of cluster galaxies at $z \approx 0$ to constrain the star formation histories of galaxies hosted by dark matter halos of different masses. To this end, we parametrise the SFR as a function of halo mass and redshift using piecewise power laws. We combine this empirical model for star formation with halo merger trees to follow the evolution of the stellar masses of galaxies and to make model predictions to be compared with the data constraints. We use the MULTINEST method developed by Feroz et al. (2009) to obtain the posterior distribution of the model parameters and the marginal likelihood. A series of nested model families with increasing complexity are explored to understand how the model parameters are constrained by the different observational data sets. Our main results are summarised as follows:

- To match the observed SMFs at different redshifts, the SFR in central galaxies residing in halos with masses above $10^{12}\,h^{-1}\mathrm{M}_{\odot}$ has to be boosted at high redshift relative to the increase that arises naturally from the fact that the dynamical time scale is shorter at higher z.

- To reproduce the faint end of the cluster and field galaxy luminosity functions ($M_z - 5\log_{10}(h) > -18$) simultaneously, we require a characteristic redshift $z_c \approx 2$ above which the SFR in low mass halos with masses $< 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ must be enhanced relative to that at lower z.

Our findings have important implications for the physical processes that regulate star formation and feedback. In general the star formation rate in a halo depends on the amounts of cold gas that can be accreted into the halo center, and on the time scale with which the cold gas converts into stars. In current theories of galaxy formation the amount of cold gas in a halo is determined by radiative cooling and feedback effects.

It is well known that radiative cooling introduces a characteristic halo mass, $M_{\rm cool} \approx 6 \times 10^{11}\,\mathrm{M}_{\odot}$, which separates cooling limited 'hot mode' and free-fall limited 'cold mode' in the accretion of cold gas into galaxies (Birnboim and Dekel, 2003; Kereš et al., 2005; Kereš et al., 2009). For halos below this characteristic mass, gas is never heated during accretion and so the amount of cold gas is limited by the free fall time of the gas. For halos with larger masses, on the other hand, the accreted gas first heats by accretion shocks and then cools radiatively before it sinks into the central galaxy, so that cold gas accretion by the central galaxy is limited by the radiative cooling time scale. The characteristic mass scales we find in the star formation efficiency shown in Figures 3.1 and 3.2 are very similar to $M_{\rm cool}$, suggesting that radiative cooling plays an important role in star formation. Furthermore, since the cooling time scale decreases faster than the free-fall time scale as redshift increases see §8.4 in Mo et al., 2010, cooling may also have played a role in the enhanced SFR in massive halos at high z (see §2.4.2).

However, radiative cooling alone cannot explain why the star formation efficiency in lower mass halos is suppressed. Even for massive halos, numerical simulations have shown that the suppression in radiative cooling at low z is not sufficient to explain the observed low SFRs, and some heating sources are needed to quench the star formation in massive galaxies at low z. One popular mechanism is AGN feedback. Observations show that AGN activity peaks at $z \approx 2$ and declines towards both higher and lower redshift (e.g. Hopkins et al. 2007), indicating that super-massive black holes may have already formed in massive galaxies by $z \approx 2$. Thus, the quenching of star formation in massive galaxies at $z < 2$ may owe to a combination of effective AGN feedback (e.g. in low-accretion radio mode) and inefficient radiative cooling owing to the reduced gas density. Similarly, the high star formation rate in high mass halos at high z may arise from an increased radiative cooling efficiency combined with reduced AGN feedback owing to the reduced number of super-massive black holes that have formed or the presence of cold, filamentary accretion in these massive halos at high redshift (Kereš et al., 2009). Our results for the star formation in massive galaxies are in quantitative agreement with these expectations, and a detailed comparison between these empirical results and theoretical predictions will provide important insights into the underlying physical processes.

Our results for the star formation in low-mass halos poses a number of challenges to standard theory. Since radiative cooling is expected to be effective at all redshifts in low-mass halos, some feedback processes must be invoked to suppress the star formation efficiency in these halos. A popular assumption is that galactic winds driven by supernova explosions may suppress the star formation in such halos. In many models considered thus far, the mass loading factor, which is defined to be the mass loss rate through winds divided by the SFR, is assumed to be some power of the circular velocity of the host halos. Such feedback models is applied to galaxies at all redshifts. The existence of a transition at $z = z_c \approx 2$ separating active from quiescent star forming phases, which is required to explain the faint-end upturn in the CGLF, is not expected in the conventional

supernova feedback model. Instead, our results lend support to a scenario in which the IGM is preheated at $z \approx 2$ and hence the accretion of baryons into low mass halos is delayed until they become sufficiently massive to allow significant accretion from the preheated IGM to form stars at lower redshift. The exact mechanism for preheating is still unclear. Possibilities that have been proposed the formation of pancakes (Mo et al., 2005), an episode of starburst and AGN activity at $z \gtrsim 2$ (Mo and Mao, 2002), and heating by high-energy gamma rays generated by blasars (Chang et al., 2012). In all these preheating scenarios, preheating is expected to be at $z \approx 2$, in excellent agreement with the value of z_c that we find. The preheated entropy of the IGM is about a few times $10\,\mathrm{KeV\,cm^2}$.

CHAPTER 3

MODEL PREDICTIONS

In this chapter, we use our posterior distributions to investigate different statistics of the modeled galaxy population, both to explore some observational consequences of our models and to better understand the physics that may give rise to our model. We present results of Model III, constrained using both the SMFs and the z-band CGLF, as well as Model II, constrained using just the SMFs. The latter is presented for comparison because this model is similar to the results of other past work (Yang et al., 2012; Behroozi et al., 2013b; Behroozi et al., 2013a; Moster et al., 2013; Mutch et al., 2013b) that do not use the CGLF as a constraint.

3.1 The correlations between galaxies and dark matter halos

3.1.1 The stellar mass - halo mass relation of central galaxies

The left panels in Figure 3.1 show the stellar mass to halo mass ratio as a function of halo mass at different redshifts as predicted by Model II. The dispersion owing to the inferential uncertainties are shown as the solid bands. These results are similar to those obtained from earlier investigations (e.g., Conroy and Wechsler, 2009; Yang et al., 2012; Behroozi et al., 2013b; Leauthaud et al., 2012; Moster et al., 2013) The ratio shows a broad peak around $10^{12}\,h^{-1}\mathrm{M}_{\odot}$ with a gradual shift towards larger halo masses at higher redshifts. The right panels show the prediction of Model III. Compared with the predictions of Model II, we see a different evolution in halos with masses below $2 \times 10^{11}\,h^{-1}\mathrm{M}_{\odot}$, where the stellar mass to halo mass ratio is independent of halo mass at high redshifts, owing to an almost constant star formation efficiency.

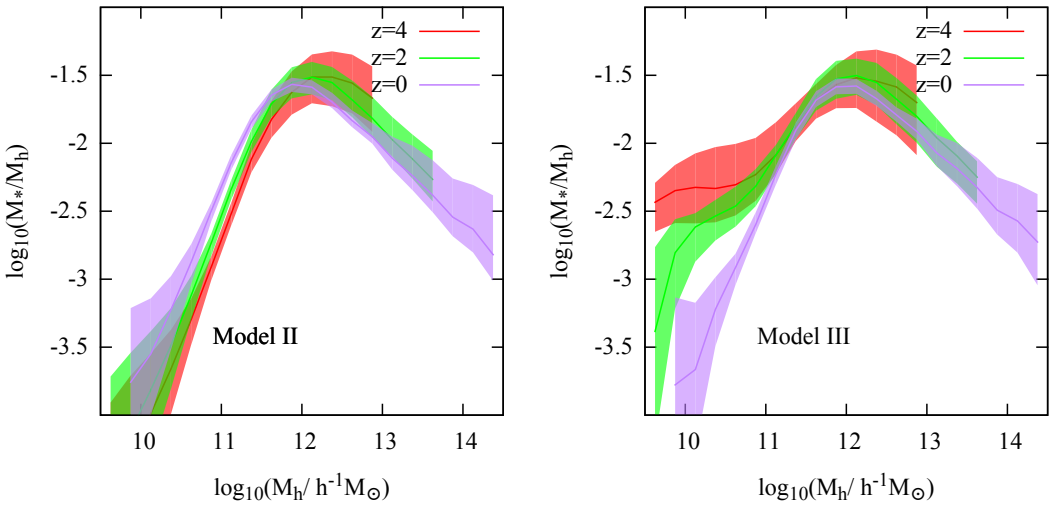

Figure 3.1. The stellar mass to halo mass ratio as a function of halo mass for Model II (left panels) and Model III (right panels). It shows both the medians of the posterior prediction as well as the 95% incredible intervals (bands).

3.1.2 SFR and halo mass accretion rate

One way to understand the star formation history in a halo is to examine how the SFR correlates with the mass accretion rate of the host halo. We follow Behroozi et al. (2013a) and define a star formation efficiency as the ratio between the SFR in the central galaxy and the mean halo mass accretion rate $\left\langle \dot{M}_{\rm h}(z) \right\rangle$ multiplied by the universal baryon fraction f_B:

$$\epsilon_{\rm SFR}(z) \equiv \frac{\rm SFR(z)}{f_B \left\langle \dot{M}_{\rm h}(z) \right\rangle} \,. \tag{3.1}$$

The SFR(z), defined in Eqs 2.1 and 2.2, is implicitly assumed to be an average over galaxies of similar halo mass. For a given halo mass $M_{\rm h}(t)$, the average mass accretion rate is calculated using

$$\langle \dot{M}_{\rm h}(t) \rangle \equiv [M_{\rm h}(t) - \langle M_{\rm prim}(t - {\rm d}t) \rangle]/{\rm d}t \,, \tag{3.2}$$

where $\langle M_{\rm prim} \rangle$ is the average mass of the primary progenitors. The primary progenitors are sampled using the same algorithm adopted to generate the merger trees.

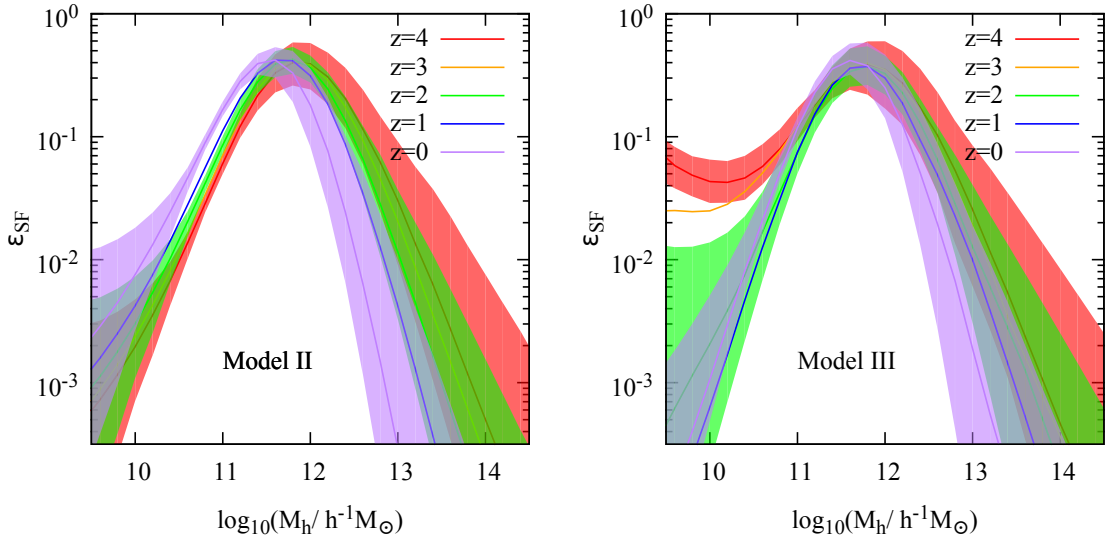

Figure 3.2. The median star formation efficiency as a function halo mass $M_{\rm h}$ for both Model II and Model III. For clarity, the 95% credible range of the posterior prediction is shown only for $z = 0$, 2 and 4.

For Model II, the star formation efficiency at $z \approx 4$ is strongly peaked at $M_{\rm h} \approx 10^{12}\,h^{-1}{\rm M}_\odot$, with a peak value $\epsilon_{\rm SFR} \approx 1/2$. The position and height of the peak depend mildly on redshift; at $z \approx 0$ it shifts to $M_{\rm h} \approx 4\times10^{11}\,h^{-1}{\rm M}_\odot$, with a peak value $\epsilon_{\rm SFR} \approx 0.8$. $\epsilon_{\rm SFR}$ increases (decreases) with halo mass as a steep power-law at the low (high) mass end. Such simple behaviors are similar to those obtained by Bouché et al. 2010; Behroozi et al. 2013a. In the toy model proposed by Bouché et al. (2010), the SFR in halos with masses between $3 \times 10^{11}\,h^{-1}{\rm M}_\odot$ and $2 \times 10^{12}\,h^{-1}{\rm M}_\odot$ follows the baryonic accretion rate, and is completely quenched in halos outside this range.

Model III predicts a different behaviour for halos with masses $< 10^{11}\,h^{-1}{\rm M}_\odot$ than Model II or the simple models proposed by Bouché et al. (2010), Behroozi et al. (2013a), Yang et al. (2013), and Behroozi et al. (2013a), which implies a simple but strong quenching of star formation. Beyond $z_{\rm c}$ at $z \approx 4$, the SFR is roughly about 1/20 of the baryonic accretion rate, and is independent of the host halo mass. The star formation is quenched abruptly at $z \approx z_{\rm c}$ if the halo is much smaller than $10^{11}\,h^{-1}{\rm M}_\odot$ and makes a mild recovery at $z \approx 0$.

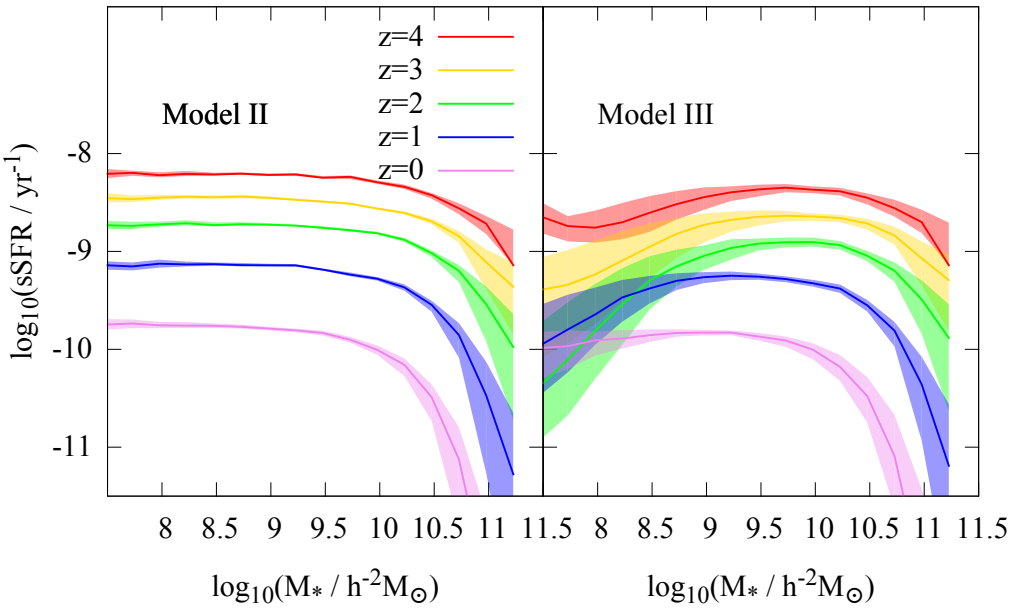

Figure 3.3. The specific star formation rate (sSFR) versus stellar mass at different redshifts for Model II (left panel) and Model III (right panel). The solid lines are the median and the bands are the 95% incredible interval.

3.1.3 The specific star formation rate

In Figure 3.3, we show the specific star formation rate (sSFR; defined to be the SFR divided by the stellar mass) as a function of stellar mass at different redshifts. Here again we compare between the predictions of Model II (left panels) and Model III (right panels). For a given stellar mass, the sSFR increases with redshift. On the other hand, for a given redshift, the sSFR declines rapidly with galaxy mass as the mass goes beyond a critical mass, which increases from $\approx 10^{10}\,h^{-2}\mathrm{M}_{\odot}$ at $z = 0$ to $\approx 5 \times 10^{10}\,h^{-2}\mathrm{M}_{\odot}$ at $z = 4$. For galaxies between $10^{9}\,h^{-2}\mathrm{M}_{\odot}$ and the critical mass, the sSFR is almost independent of stellar mass, in qualitative agreement with the observations (e.g. Daddi et al., 2007; Noeske et al., 2007b). Model II and Model III differ in their predictions for low-mass galaxies with stellar masses $< 10^{9}\,h^{-2}\mathrm{M}_{\odot}$. For Model II, the weak correlation between sSFR and stellar mass extends all the way down to such galaxies. For Model III, however, the correlation is much more complicated: the sSFR and stellar mass show no significant

correlation at $z = 0$, show a strong positive correlation between $z = 1$ and $z = 2$, and show a weak positive correlation at higher redshifts. For low-mass galaxies at high z, Model III predicts a lower sSFR compared to Model II, because these galaxies in Model III form their stars earlier than galaxies of the same mass in Model II. In Model III the SFR in low-mass galaxies drops dramatically at the critical redshift $z_c \approx 2$ after a significant amount of stars have already formed at higher z, making the sSFR in dwarf galaxies much lower than that in more massive galaxies at the same epoch. At $z \approx 0$ the sSFR in dwarf galaxies catches up with those in more massive galaxies, because of the growth of their halos and because of the strong mass dependence of the SFR at the low-mass end at low z.

Observations indicate that there is a negative correlation between sSFR and stellar mass for dwarf star forming galaxies (Noeske et al., 2007b), a trend that appears contrary to the predictions of Model III shown in the right panel. However, the existence of such a correlation in the observational data is still uncertain owing to sample incompleteness. Indeed, empirical models based on such correlations (Noeske et al., 2007a; Leitner, 2012) suggest that low-mass galaxies form most of their stars in the past few billion years, in apparent contradiction with the star formation histories inferred directly from the colour-magnitude diagrams of stars (e.g. Weisz et al., 2011), which seem to support Model III (see §4) More data on the detailed star formation histories of isolated dwarf galaxies is required to better discriminate between these different models.

3.2 The Stellar Mass Functions

3.2.1 The evolution of galaxy SMFs

As shown in the previous chapter, the luminosity function of local cluster galaxies put tight constraint on the star formation in the high-z Universe. The active star formation in the dwarf halos ($\lesssim 10^{11}\,h^{-1}\mathrm{M}_{\odot}$) in the high-$z$ Universe give rise to the steep low-mass end of the SMFs (Figure 3.4). At high mass end ($M_{\star} > 10^{9}\,\mathrm{M}_{\odot}$) the two models give

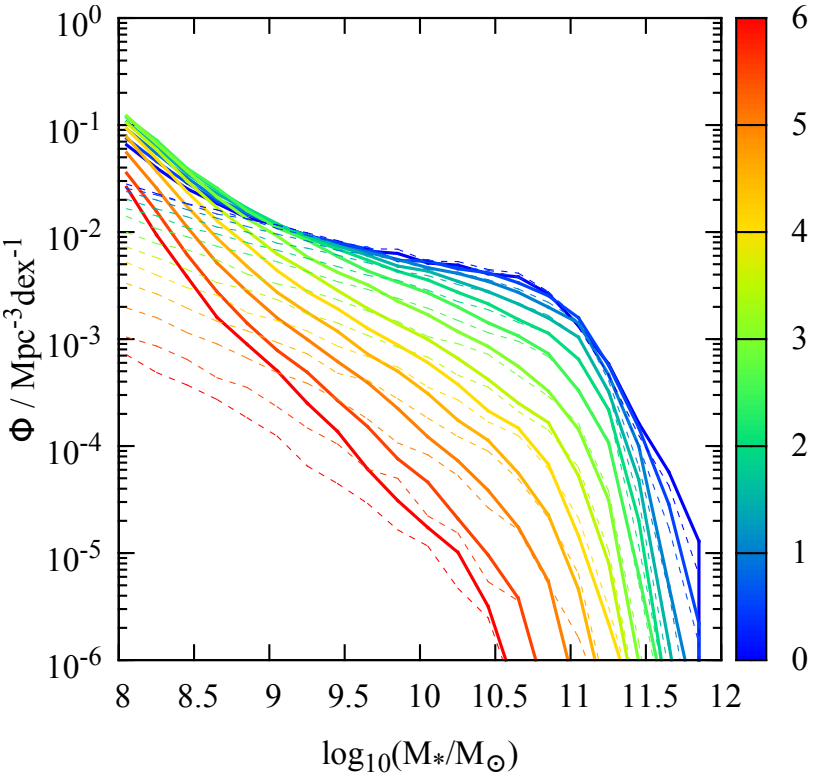

Figure 3.4. The evolution of galaxy SMFs. The solid lines are prediction of Model III and the dashed lines are prediction of Model II. The redshifts are coded in different colors as indicated in the figure.

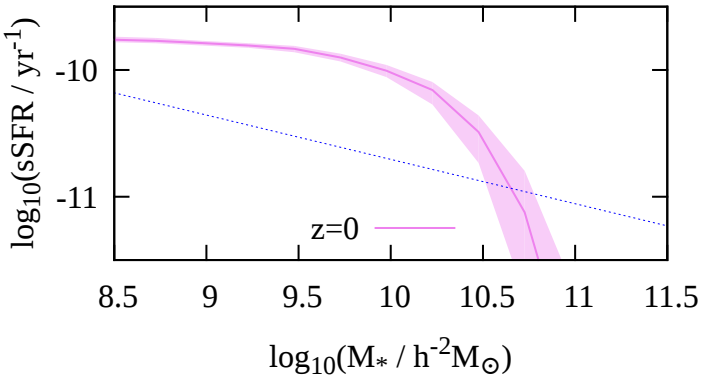

Figure 3.5. The sSFR stellar mass relation predicted by Model III. The blue line is the minimal sSFR of star forming galaxies (Moustakas et al., 2013).

similar prediction because of the same observational constraints. The number density of high mass galaxies has increased significantly with time. At the low mass end, especially at $10^8\,\mathrm{M}_\odot$, Model III suggests that the number density of such galaxies has changed little since $z \approx 5$. Therefore, accurate observations of the SMFs at $z > 3$ down to a stellar mass limit $M_\star \approx 10^8\,\mathrm{M}_\odot$ are crucial to discriminate the two models. However, the detection limit of current observations is roughly $\sim 10^9\,\mathrm{M}_\odot$.

3.2.2 The SMFs of star forming and quenched galaxies

The galaxies can be divided into star forming and quenched populations using the criterion suggested by (Moustakas et al., 2013), as shown in Figure 3.5, which writes as

$$\log_{10}\left(\mathrm{SFR}_{min}\right) = -0.49 + 0.65\log_{10}\left(\frac{M}{10^{10}\,\mathrm{M}_\odot}\right) + 1.07\,(z - 0.1)\ . \tag{3.3}$$

The cut indicates a critical stellar mass $10^{11}\,\mathrm{M}_\odot$. There are almost no star forming galaxies bigger than this mass, while too few quenched galaxies below this mass (Figure 3.6).

The problem is caused by the lack of bimodality in the distribution of central galaxies. Galaxies at $z \approx 0$ can be divided into two distinct populations, that is star forming and quenched. The bimodality seems to exist in different environments, span a large range

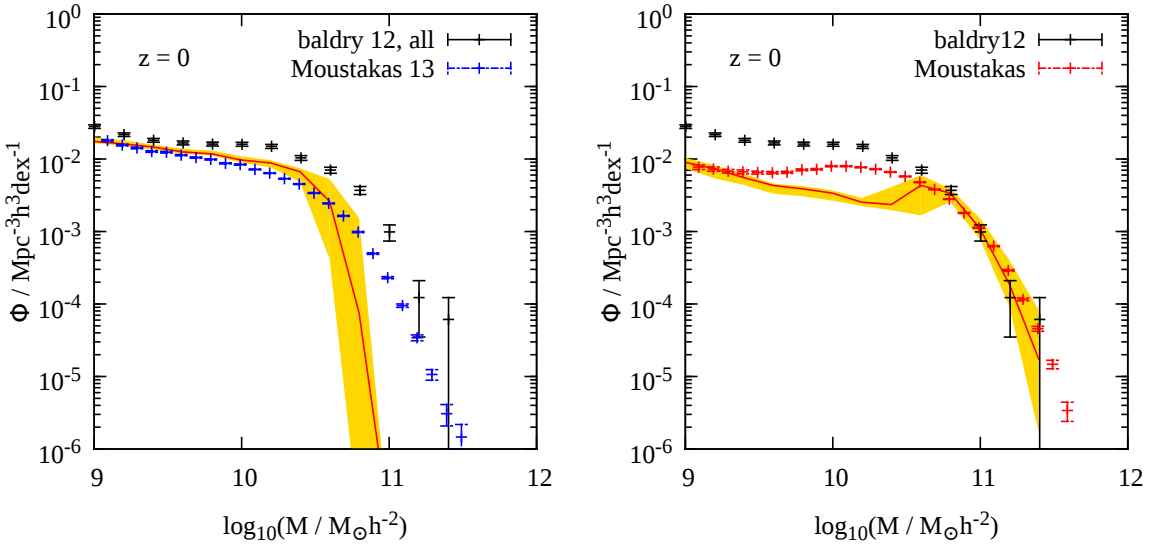

Figure 3.6. The SMF of star forming galaxies (left) and quenched galaxies (right) predicted by Model III.

of galaxy mass and has been recently confirmed up to $z \approx 1$ (Moustakas et al., 2013). In the model presented here the SFR in the central galaxies can only be considered as the average over all type of galaxies in the same halo mass bin. The sSFR decreases monotonically with stellar mass (Figure 3.5).

3.2.3 The SMF of satellite galaxies

The SMFs of satellite galaxies at $z = 0$ are shown in Figure 3.7. The contribution of satellites at high stellar mass is lower than that of the central galaxies, with a satellite to central ratio of about 1/3. The results are in good agreement with those obtained by Yang et al. (2008) based on galaxy groups selected from the SDSS (see also Mandelbaum et al., 2006; Cacciato et al., 2013). At the low-mass end ($M_* \approx 10^8 h^{-2}\,\mathrm{M}_\odot$) the predictions of the two models are in contrast: the satellite fraction predicted by Model III is higher than that predicted by Model II, and eventually overtakes the fraction of central galaxies of similar masses. The overabundance in the overall population of dwarf satellites has the same orgin as that of the cluster galaxies, that is the global enhancement in the star

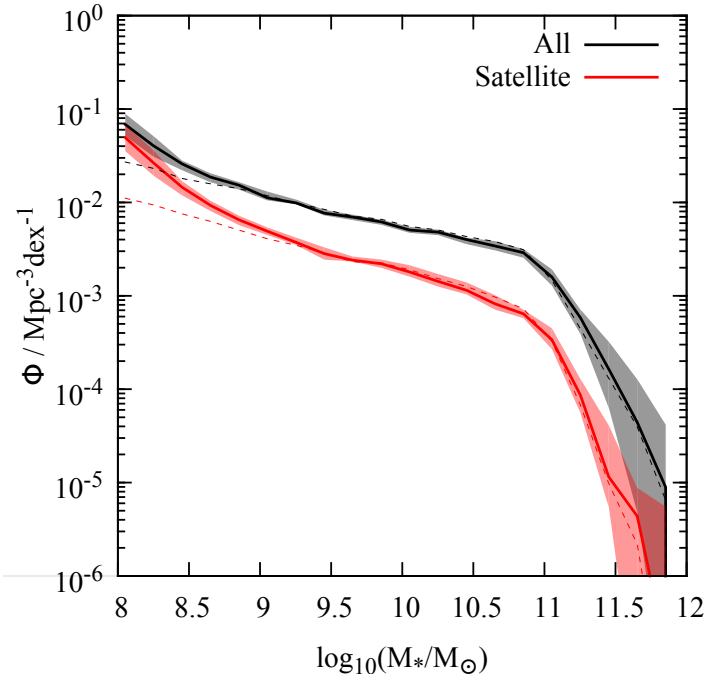

Figure 3.7. The SMF of satellite galaxies. The solid lines with bands are prediction of Model III and the dashed lines are prediction of Model II.

formation activity at high redshift. This could be checked by studying galaxy groups in the future deeper surveys.

3.2.4 The conditional stellar mass functions (CSMFs)

We also make predictions for the Conditional Stellar Mass Function (hereafter CSMF), which is the SMF of galaxies hosted by halos of a given mass. We show them in Figure 3.8 together with the observational results of Yang et al. (2008). Owing to the detection limits, the CSMF below $10^9\,h^{-2}\mathrm{M}_\odot$ is either noisy or unavailable. The predictions of both Model II and Model III above the detection limit are consistent with the observational data. (Note that a quantitative comparison requires the prediction be convolved with the effects of the group finder (Reddick et al., 2013), which is beyond the scope of this work. Compared with Model II, Model III predicts more dwarf galaxies, with masses below the current detection limit, not only in massive clusters but also in low-mass groups. One

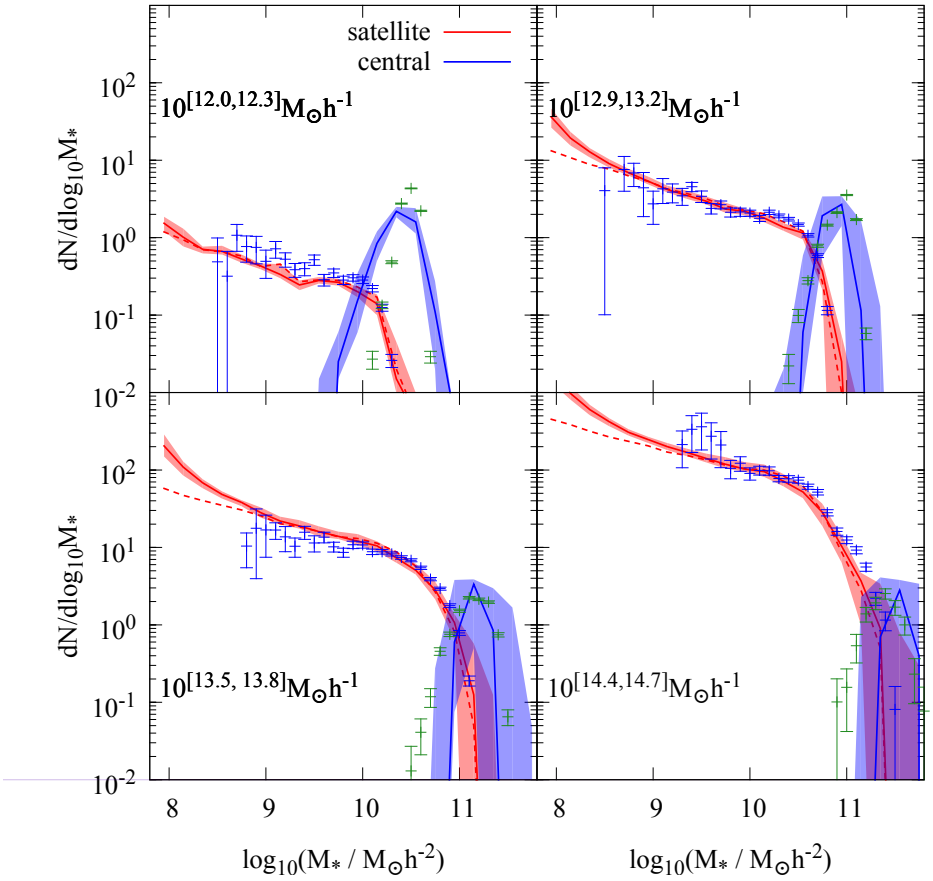

Figure 3.8. The conditional SMFs for different halo mass bins as indicated. The solid lines with bands are the predictions of Model III, while the dashed lines are the predictions of Model II. The distribution of centrals and satellites are in blue and red, respectively. The data points are from Yang et al. (2008).

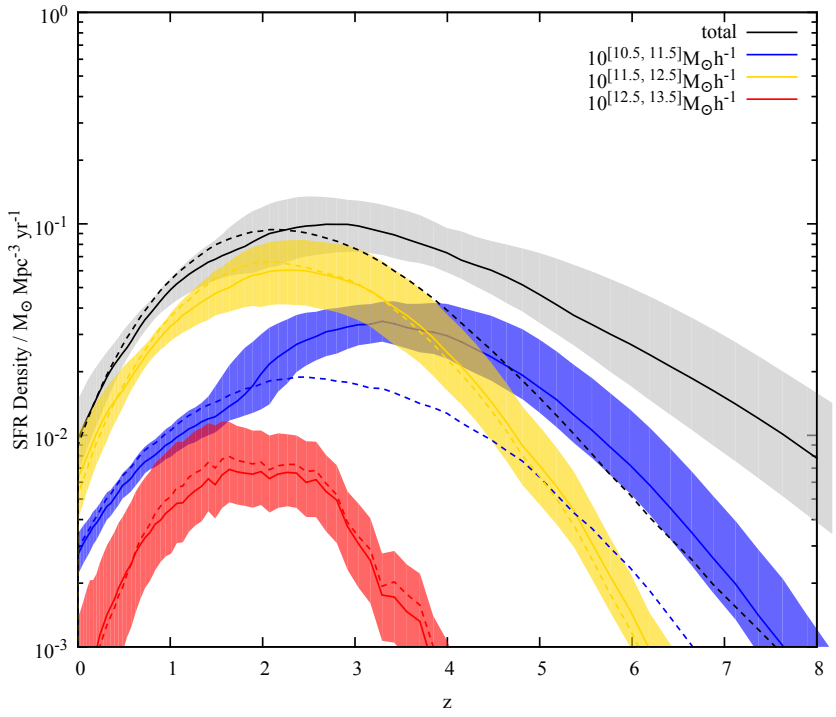

Figure 3.9. The cosmic star formation rate density (SFRD) as a function of redshift predicted by Model II (dashed lines) and Model III (solid lines and bands). The lines and bands are the medians and 95% credible intervals, respectively. The total star formation rate density is shown as the grey band.

expects this behaviour because in Model III dwarf galaxies in massive clusters are fossils of a relative global enhancement of star formation activity in dwarf halos in the high-z Universe. This boosted star formation at high-z also leaves an imprint in present-day galaxy systems of lower halo masses, not just in rich clusters.

3.3 The cosmic star formation history

3.3.1 The cosmic star formation rate density

Figure 3.9 show the star formation rate density (SFRD) predicted by Model III (solid lines and bands) and Model II (dashed lines), respectively. These results take into account star formation in halos down to a mass of $2 \times 10^9\, h^{-1}\mathrm{M}_\odot$, the resolution of our merger trees. The total SFRD is decomposed into contributions by halos of different masses. The total SFRD declines by an order of magnitude towards low-z beginning at $z = 2$. In both

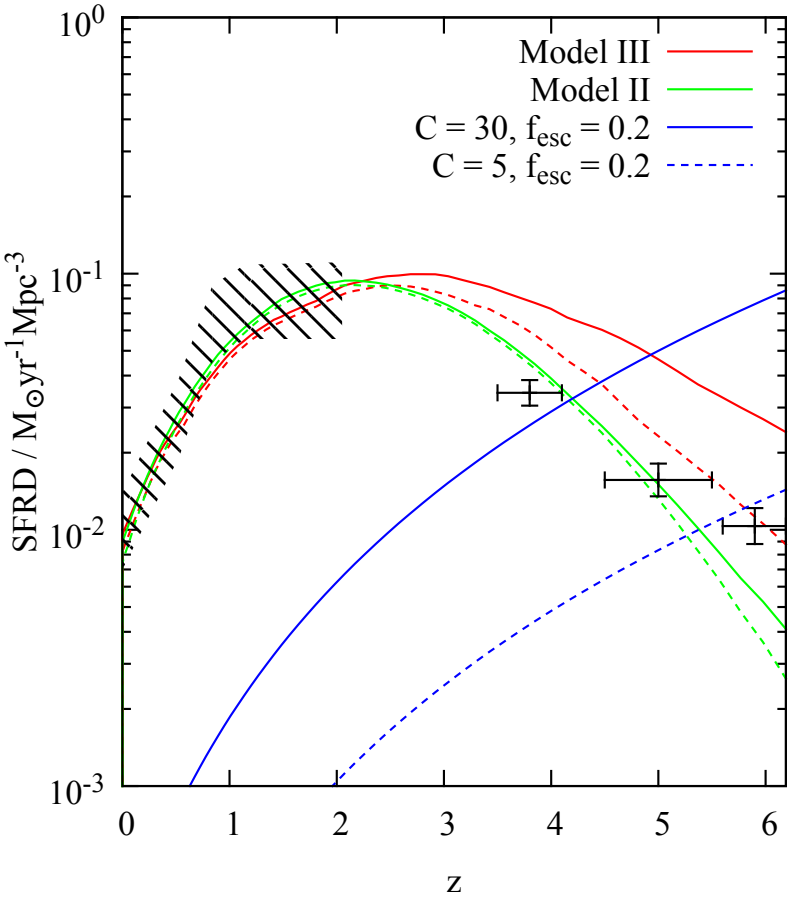

Figure 3.10. The predicted star formation rate density (red and green lines) and the minimal star formation rate density required to keep the Universe ionised (blue lines). The hedged band and the data points are from Bouwens et al. (2012). For the model predictions, the solid lines are obtained by integrating down to the merger tree resolution, which is $2 \times 10^9\, h^{-1}\mathrm{M}_\odot$, and the dashed lines are obtained by integrating down to the detection limit in Bouwens et al. (2012), which is $M_{\mathrm{UV,AB}} = -17.7$ or $0.36 M_\odot \mathrm{yr}^{-1}$.

Model II and Model III, the predicted total SFRD in this redshift range owes mainly to star formation in halos with masses between $3 \times 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ and $3 \times 10^{12}\,h^{-1}\mathrm{M}_{\odot}$, about Milky Way mass. The SFR in such halos scales with the halo mass accretion rate, which is proportional to $(1+z)^{2.3}$. However, at $z > 3$ the two models behave differently. Model II predicts a rapid decline in the total SFRD towards high z, as the abundance of $10^{12}\,h^{-1}\mathrm{M}_{\odot}$ halos decreases towards higher redshift while in dwarf halos, which are abundant at high-z, the star formation is strongly suppressed. The SFRD at $z > 3$ predicted by Model III is substantially higher, mainly because the SFR in low-mass halos ($< 3 \times 10^{11}\,h^{-1}\mathrm{M}_{\odot}$) or dwarf galaxies ($< 10^{9}\,h^{-2}\mathrm{M}_{\odot}$) is boosted at $z > z_c$ in this model.

The dwarf galaxies, which dominate the cosmic SFRD at high redshift, are missed in the current observations, and so the discrepancy between the prediction of Model III and the observational results at $z > 3$ (data points with error bars in Figure 3.10) probably owes to incompleteness in the data. Indeed, if we use the same lower limit of UV magnitude adopted in Bouwens et al. (2012), which is $M_{\mathrm{UV,AB}} = -17.7$, to predict the SFRD, we get the results shown by the dashed lines, which brings the prediction of Model III into much better agreement with the data. The change in the prediction of Model II is small, because in this model galaxies below the limit do not make a significant contribution. This demonstrates clearly the importance in observing and modelling very faint galaxies to understand the SFRD at high z.

If one assumes that the Universe is kept ionised by the UV photons from young stars, a minimal SFRD required can be estimated using the equation given by Madau et al. (1999):

$$\frac{\mathrm{SFRD}_{\mathrm{min}}(z)}{\mathrm{M}_{\odot}\mathrm{yr}^{-1}\mathrm{Mpc}^{-3}} \approx 3 \times 10^{-4} \left(\frac{C}{f_{\mathrm{esc}}}\right) \left(\frac{1+z}{6}\right)^{3}, \tag{3.4}$$

where C is the clumpiness factor of the IGM and f_{esc} is the average escape fraction of UV photons from star forming galaxies. The normalisation factor is consistent with the Chabrier IMF (Chabrier, 2003). The clumpiness factor is expected to evolve with redshift and its value is quite uncertain. Madau et al. (1999) adopted $C = 30$ at $z = 5$ based on

the cosmological simulation of Gnedin and Ostriker (1997). More recently, Bolton and Haehnelt (2007) claimed that C at $z = 7$ may be as small as 5. For comparison we show in Figure 3.10 the minimum SFRD as a function of z assuming $C = 30$ (solid blue line) and $C = 5$ (dashed blue line), with $f_{\rm esc} = 0.2$ in both cases. If the clumpiness factor is as high as 30, Model III is able to keep the Universe ionised at $z \approx 5$, while SFRD of Model II is too low. However, without better constraints on C and $f_{\rm esc}$, no stronger statement can be made about the two models.

3.3.2 The star formation rate function

In Figure 3.11 we show the predicted star formation rate functions for galaxies at different redshifts. Note that none of these functions can be well fit with a Schechter function. For $z < 3$, there is a sharp cutoff following a bump at the high-SFR end. This owes to the existence of peaks in the SFR-halo mass relations (see Figure 2.11). However, this feature should not be taken too seriously because what we show here is based on the average SFR - halo mass relation, ignoring any dispersion in the relation. Despite this, the characteristic star formation rate clearly decreases with decreasing redshift, by a factor of almost 100 from $z = 4$ to $z = 0$. Assuming that the faint part of the distribution can be fit by a power law, the power law index predicted by Model III changes significantly from roughly -2.0 at $z = 4$ to roughly -1.2 at $z = 0$. For Model II the change in the faint end slope is much more moderate, from roughly -1.5 at $z = 4$ to -1.2 at $z = 0$. For comparison we also show the SFR function derived by Smit et al. (2012) from the UV luminosity function of galaxies. We see that Model II significantly underpredicts the number density of galaxies at the low-SFR end, while Model III matches the data much better. Note that the observed SFR functions at the high-SFR ends are lower than the predictions of both Model II and Model III. One possible reason for this discrepancy is that the highest SFR galaxies are dusty and could be missed in UV observations.

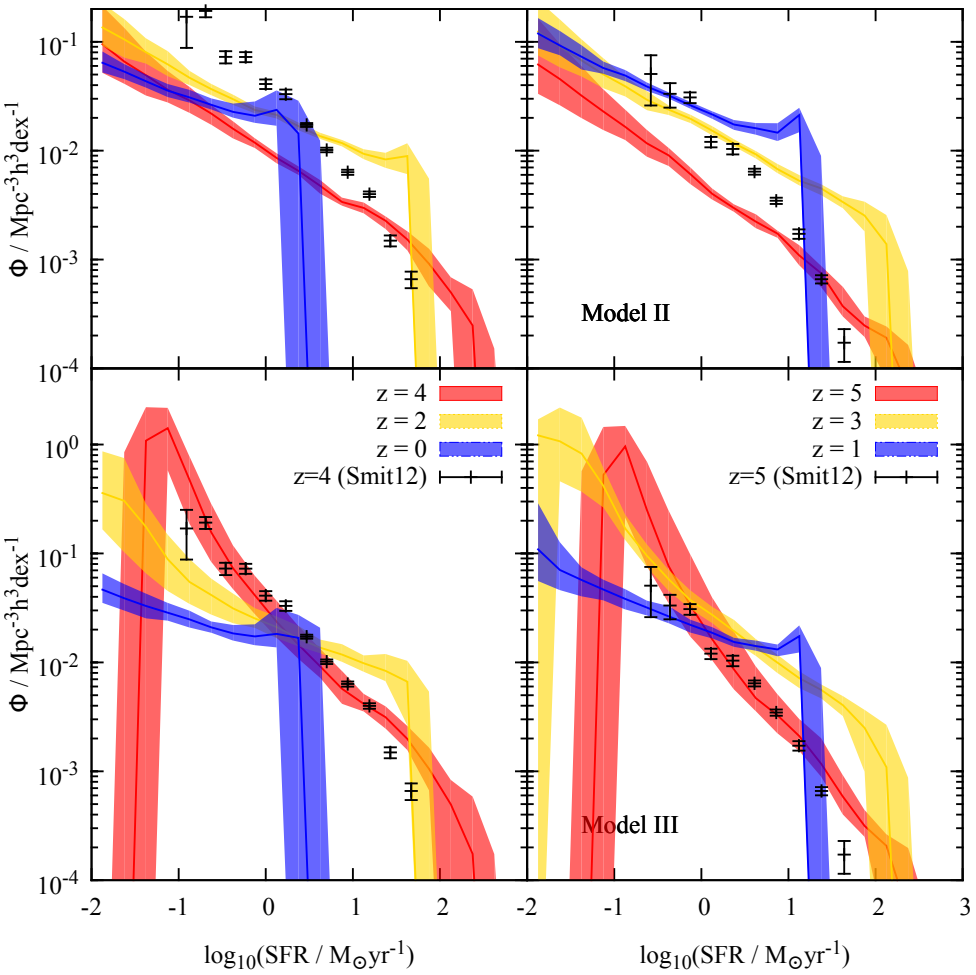

Figure 3.11. The star formation rate functions at different redshifts predicted by Model II (upper panels) and Model III (lower panels). The solid lines are the medians and the bands encompass the 95% credible intervals. The data points are the observational results from Smit et al. (2012).

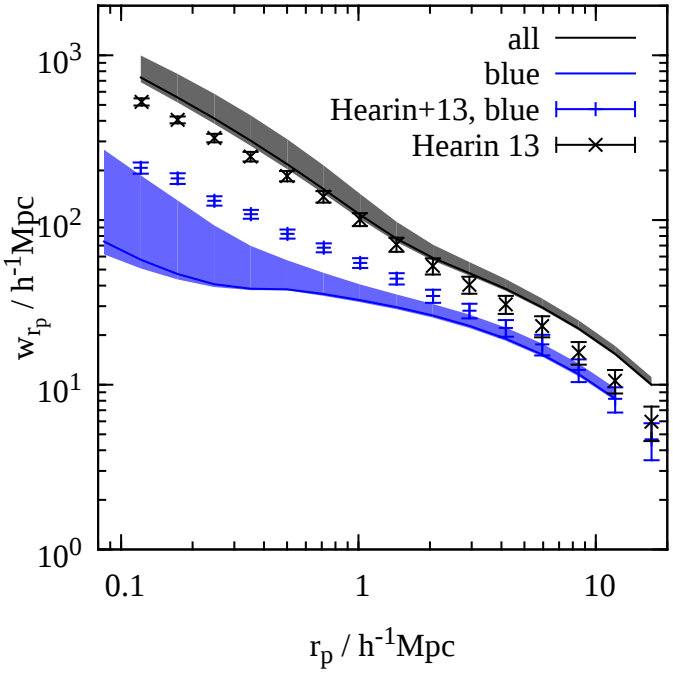

Figure 3.12. Projected 2 point correlation function of the galaxies with mass $> 10^{12}\,M_\odot$ (black) and blue galaxies in the same mass range (blue). The bands are predictions of Model III and the data points are from (Hearin and Watson, 2013).

3.4 The projected two point correlation function

Figure 3.12 shows the projected two point correlation function (2PCF hereafter) of all galaxies with masses larger than $10^{10.2}\,\mathrm{M}_{\odot}$ (black), in comparison with that for star forming galaxies in the same mass range (blue). The colored bands are the predictions of Model III. The predictions are calculated using the analytic model presented in Yang et al. (2012). Theoretically, the 2PCF can split into "one-halo" term and "two-halo" term. The "one-halo" term is correlation contributed by galaxy pairs that live in the same host halos, including the satellite-satellite pairs and central-satellite pairs. The "two-halo" term is due to galaxies that occupy different halos. On small scale ($r_{\mathrm{p}} < 1\,h^{-1}\mathrm{Mpc}$), the 2PCF is supposed to be dominated by the "one-halo" term.

The prediction on the 2PCF of all type of galaxies matches the observation reasonably well on all scales (note that the model is not tuned to match the clustering data). However, on small scales ($r_{\mathrm{p}} < 1\,h^{-1}\mathrm{Mpc}$), the 2PCF of star forming galaxies predicted by the model has a large variance and the median is lower than the observations. On such small scales, the 2PCF is dominated by the population of star forming satellites. The large variance in the prediction means that the star formation activity in satellites is not well constrained in our model. The observational constraints used in this work are not particularly sensitive to the star formation after infall of the satellites.

3.5 Summary

The fiducial model (Model III) successfully matches the recently measured SMFs and the faint upturn of the z-band cluster galaxy luminosity function. The model makes the following predictions that can be tested against future observations:

1. The star formation efficiency, that is the SFR divided by the baryonic mass accretion rate of the host halo, peaks in halos with masses between $3 \times 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ and $10^{12}\,h^{-1}\mathrm{M}_{\odot}$. In lower mass halos, the star formation efficiency is about 1/10 at $z > z_c$ and is strongly quenched at lower z and roughly scales as $M_{\mathrm{h}}^{3/2}$. While in

higher mass halos, the star formation tends to be quenched and the quenching is stronger with decreasing redshift.

2. The stellar mass to halo mass ratio, $M_*/M_{\rm h}$, for central galaxies peaks at a halo mass of $\approx 10^{12}\,h^{-1}{\rm M}_\odot$ with a value of $\approx 1/30$, quite independent of redshift. For halos with masses below $2\times10^{11}\,h^{-1}{\rm M}_\odot$ our model predicts that $M_*/M_{\rm h}\approx 1/100$ at $z=4$ quite independent of halo mass, but this ratio decreases rapidly with decreasing halo mass at $z=0$.

3. Because of the active star formation in low mass halos at high redshift the low-mass end slopes of the SMFs and the SFR function steepen toward high redshift. This enhanced star formation activity also leaves an imprint in the present-day satellite population. The SMF of the whole satellite population and the conditional SMFs of groups and rich clusters have an upturn at the low-mass end ($M_\star < 10^9\,{\rm M}_\odot$). Central galaxies dominate the present-day SMF at $M_\star > 10^9\,{\rm M}_\odot$ but satellite galaxies begin to dominate at $M_\star\sim 10^8\,{\rm M}_\odot$.

4. Halos with $M_{\rm h}\sim 10^{12}\,h^{-1}{\rm M}_\odot$, hosting centrals with $M_\star\sim 10^{10}\,h^{-2}{\rm M}_\odot$, dominate the SFRD of the Universe at $z<3$ while at higher z star formation in lower mass halos takes over. The model suggests that a lot of star forming dwarf galaxies are beyond the current detection limit at this high z. Those galaxies may contribute most of the SFRD and help keep the Universe reionized at $z>4$. Star formation in halos more massive than $10^{12.5}\,h^{-1}{\rm M}_\odot$ never significantly contribute to the total SFRD.

CHAPTER 4

STAR FORMATION AND STELLAR MASS ASSEMBLY IN DARK MATTER HALOS

In this chapter, we use the best fitting models of Model II and Model III to characterise in detail the star formation and merger histories of galaxies across cosmic time. The overall trends in the star formation and assembly histories are described in §4.1. In §4.2, we describe the total star formation and assembly histories as a function of halo (stellar) mass, paying particular attention to *in situ* star formation versus accretion and the breakdown of downsizing. In §4.3, we focus on galaxy merger rates and their implications for the morphological transformation of galaxies. In §4.4 we address the stellar populations of galaxies as a function of their stellar mass, paying particular attention to the properties of halo stars. Finally, we summarise our results in §4.5.

4.1 General Trends and Characteristic Scales

Galaxies can grow their stellar mass either via *in situ* star formation or mergers. In this section, we identify critical halo masses and redshifts to characterise the different stages of stellar mass acquisition as their host halos grow.

I use the star formation efficiency defined as Eq.3.1 to characterise the *in situ* star formation. The plots in Figure 4.1 show the star formation efficiency $\epsilon_{\rm SFR}$ as a function of $M_{\rm h}(z)$ and z. As one can see, Model II predicts that the *in situ* star formation is most efficient (with $\epsilon_{\rm SFR} > 0.3$) in a narrow band between $10^{11}\,h^{-1}{\rm M}_{\odot}$ and $10^{12}\,h^{-1}{\rm M}_{\odot}$, with a tilt towards lower mass halos at lower redshifts. Outside the band, the star formation efficiency drops rapidly towards both higher and lower masses without depending strongly on redshift. This trend agrees with the results obtained by Behroozi et al. (2013a) and

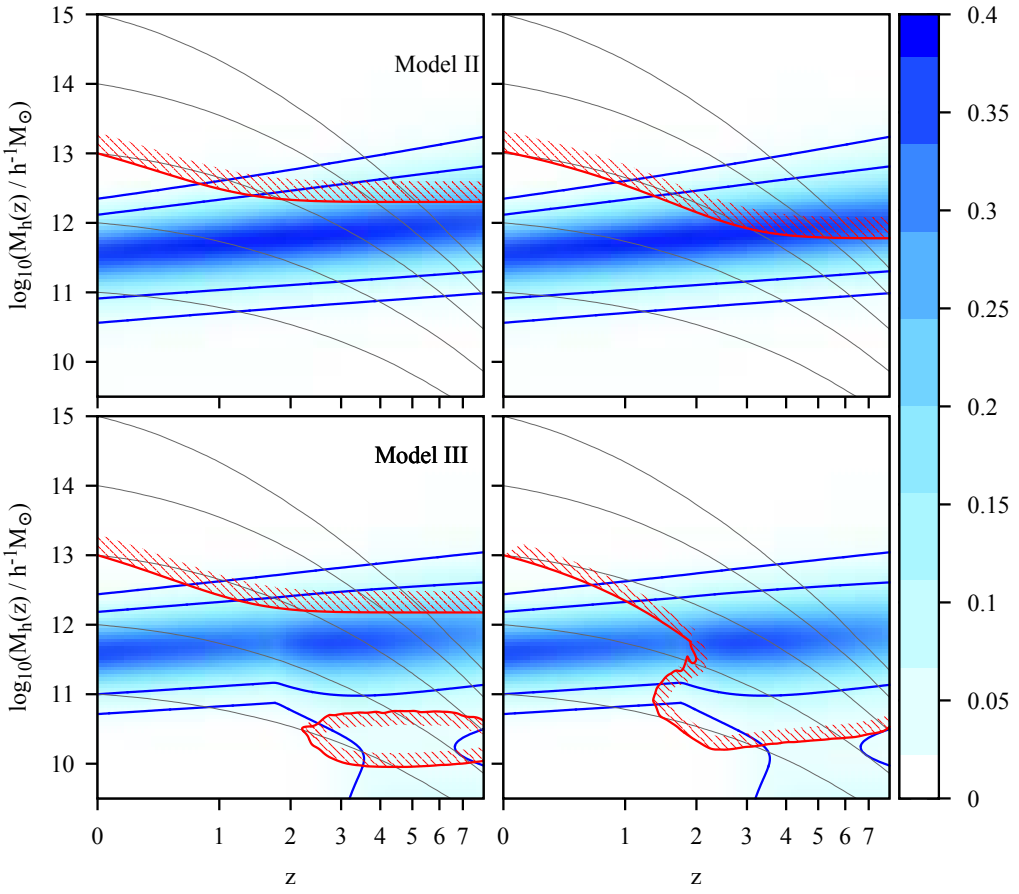

Figure 4.1. Star formation efficiency and merger rate in the halo mass - redshift plane. The blue shading is the star formation efficiency as indicated in the colour bar on the right. The blue solid lines correspond to star formation efficiency 0.1 (inner two lines) and 0.03 (outer lines). The solid red lines are the loci of one merger event per Hubble time, with the hedged sides corresponding to a higher merger frequency. The left panels are for major mergers (with mass ratios larger than 1/3) and the right panels are for minor mergers with mass ratios between 1/10 and 1/3. The solid grey lines are the mean mass assembly histories of present-day dark halos, obtained by averaging over the main branches of the merger trees with the same present-day halo mass.

Yang et al. (2013). Model III predicts a similar trend except for halos with masses $< 10^{11}\,h^{-1}\mathrm{M}_{\odot}$. Instead of a strong suppression of star formation with decreasing halo mass, the star formation efficiency remains $\sim 1/30$ at $z > 3$ for all $M_{\rm h} < 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ halos.

For central galaxies, another potentially important process that can affect their stellar mass and perhaps the size and morphology is the accretion of satellites. To characterise this process, we calculate the mean galaxy merger rate for halos of a given mass at a given redshift. We distinguish two different types of mergers: (i) major mergers for which the *stellar mass* ratio between the merging satellite and the central galaxy is $\geq 1/3$ and (ii) minor mergers for which the ratio is between 1/10 and 1/3. When calculating the mass ratios, we use the original stellar mass of the satellites before they deposit a fraction $1 - f_{\rm TS}$ into the stellar halo component. We discuss this choice in the last paragraph of §4.3. We define the merger rate as the number of mergers per unit time multiplied by the Hubble time $t_H(z) \equiv H(z)^{-1}$. In Figure 4.1, the red lines are the loci of one merger event per $t_H(z)$ in the halo mass - redshift plane; halos on the hedged sides of the loci on average have more than one merger per $t_H(z)$. We show results for both major (left panels) and minor (right panels) mergers, and separately for Model II (upper panels) and Model III (lower panels).

For Model II, central galaxies at $z = 0$ have experienced at least one major merger per Hubble time if they are hosted by halos with masses larger than $\approx 10^{13}\,h^{-1}\mathrm{M}_{\odot}$. At $z > 1$, frequent major mergers only occur for centrals hosted by halos with masses higher than $\sim 3 \times 10^{12}\,h^{-1}\mathrm{M}_{\odot}$. The predictions of Model III are quite similar for massive halos, but an additional branch of high major merger rate is also predicted for low mass halos ($10^{10}\,h^{-1}\mathrm{M}_{\odot} < M_{\rm h} < 10^{11}\,h^{-1}\mathrm{M}_{\odot}$) at $z > 3$. Comparing the red hedged lines indicating mergers with the star formation efficiency, one can see that most star forming galaxies are not associated with major mergers, except for central galaxies in low-mass halos with $M_{\rm h} = 10^{10}$ - $10^{11}\,h^{-1}\mathrm{M}_{\odot}$ at $z > 3$, where galaxies can experience major mergers

while actively forming stars. Minor mergers are more common. In particular, Model III predicts that active star forming central galaxies in all halos with $M_{\rm h} > 10^{10}\,h^{-1}{\rm M}_{\odot}$ may have experienced at least on minor merger at $z > 2$.

The following set of functions and scales characterise the star formation efficiency and merger frequency:

- The ridge of the highest star formation efficiency is well described by

$$M_{\rm h}(z) \approx 3 \times 10^{11}\,h^{-1}{\rm M}_{\odot}(1+z)^{0.3}\,, \tag{4.1}$$

  with a height $\epsilon_{\rm mx} \sim 0.5$ and a FWHM $\Delta \log_{10}(M_{\rm h}) \approx 1.0$.

- The line separating frequent from infrequent major mergers for massive halos can be approximated by

$$M_{\rm h}(z) \approx 10^{13}\,h^{-1}{\rm M}_{\odot}\left[0.7\exp(-z/0.6) + 0.3\right]\,. \tag{4.2}$$

- The line separating frequent from infrequent minor mergers in the entire redshift range for Model II, and at $z < 2$ for Model III, can be approximated by

$$M_{\rm h}(z) \approx 10^{13}\,h^{-1}{\rm M}_{\odot}\left[\exp(-z/0.8) + 0.06\right]\,. \tag{4.3}$$

- For Model-III, there is a characteristic redshift, $z_c \sim 2$, above which the star formation efficiency and the major merger frequency are boosted in low mass halos with $M_{\rm h}(z) \lesssim 10^{11}\,h^{-1}{\rm M}_{\odot}$, and the minor merger frequency is boosted in all halos with $M_{\rm h}(z) > 10^{10}\,h^{-1}{\rm M}_{\odot}$.

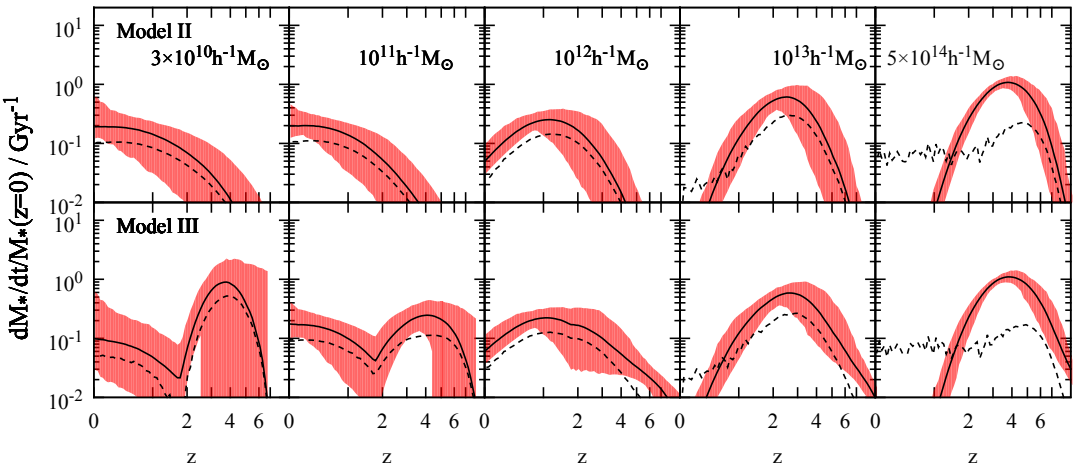

Figure 4.2. The star formation histories of galaxie with different present-day halo masses (as indicated in the upper panels) predicted by the best-fit parameters of Model II (upper panels) and Model III (lower panels). The thick solid lines are the averages; the shaded areas are the 95% ranges of variance owing to different halo merger histories. The dashed black lines are the average differential assembly histories.

4.2 Star formation and stellar mass assembly

4.2.1 Star formation histories

This total star formation history (Eq 2.9), which takes into account the history of the accreted stars and the stars formed *in situ*, is needed when modelling the stellar population of the galaxy. In Figure 4.2, the black solid lines show the total star formation histories of central galaxies in different dark halo mass ranges by averaging over a large number of halo merger histories using the best fit parameters. For comparison, the shaded band in each panel represents the variance among different merger trees of the final halo mass in question. We show results for both Model II (upper panels) and Model-III (lower panels) for five final halo masses, as indicated in each panel.

It is clear that halos of different present-day masses have different star formation histories. For centrals in massive clusters with $M_{\rm h}(z=0) \sim 5\times10^{14}\,h^{-1}{\rm M}_{\odot}$, the SFR peaks at $z \approx 3$ and the majority of the stars form in a narrow time range, which is about 10 Gyrs ago ($z > 2$). In contrast, for Milky-Way mass halos, the SFR reaches a maximum

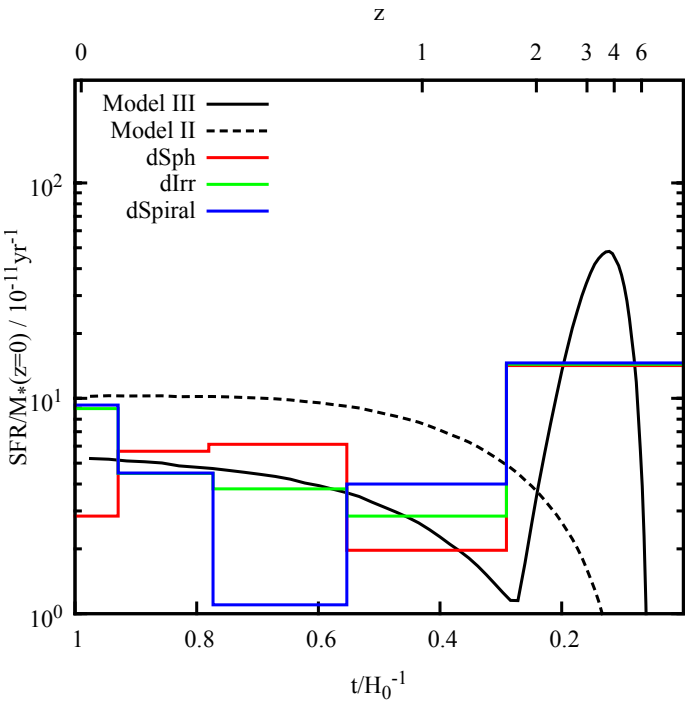

Figure 4.3. The star formation history of dwarf central galaxies in halos with masses $\approx 3 \times 10^{10}\, h^{-1}\mathrm{M}_{\odot}$ compared with the measurements of Weisz et al. (2011) for 60 dwarf galaxies of different types. The predictions of Model II are shown by dashed lines and the predictions of Model III by solid line.

between $z = 2$ and $z = 1$ and decreases only mildly to the present day. For dwarf galaxies in halos with $M_{\rm h}(z = 0) < 10^{11}\,h^{-1}{\rm M}_{\odot}$, the predictions of Model II and Model III are significantly different. For model II, most stars in these dwarf galaxies formed quite late (at $z \leq 1$) and the SFR is roughly a constant over this time interval. In contrast, Model III predicts star formation histories that are bimodal, with an initial star burst at $z > 2$, followed by a constant SFR. According to this model, this initial star burst will leave a significant amount of old stars in the present-day dwarfs.

Indeed, observations of nearby dwarf galaxies indicate that such an old stellar population is ubiquitous. With the use of deep HST imaging, individual stars of nearby dwarf galaxies can be resolved, and the colour-magnitude diagram (CMD) can be constructed to obtain the detailed star formation histories of these galaxies. Using this technique, Weisz et al. (2011) investigated 60 dwarf galaxies within a distance of 4 Mpc, many of which are field galaxies located outside the Local Group. These galaxies cover a wide range of morphological types, including dEs, dIrrs and dSpirals. They found that, on average, the dwarf galaxies formed 60% of their stars by $z \approx 2$ and 70% by $z \approx 1$, regardless of morphological type. Figure 4.3 compares our model predictions with the CMD-inferred star formation histories obtained by Weisz et al. (2011). While the predictions of Model III are in qualitative agreement with the data, Model II predicts an age distribution of stars that is clearly too skewed towards relatively young stars.

4.2.2 Stellar mass assembly histories

The average differential assembly histories are shown in Figure 4.2 as the dashed lines. The assembly history, defined in Eq. (2.8), takes into account *in situ* star formation, accretion of stars already formed, and mass loss due to stellar evolution. The average differential evolution is obtained by first averaging over the assembly histories of galaxies with the same halo mass, and then taking the time derivative of the mean assembly histories. For both dwarf and Milky-Way sized galaxies, the stellar mass assembly histories are almost parallel to the star formation histories except at the beginning of star formation.

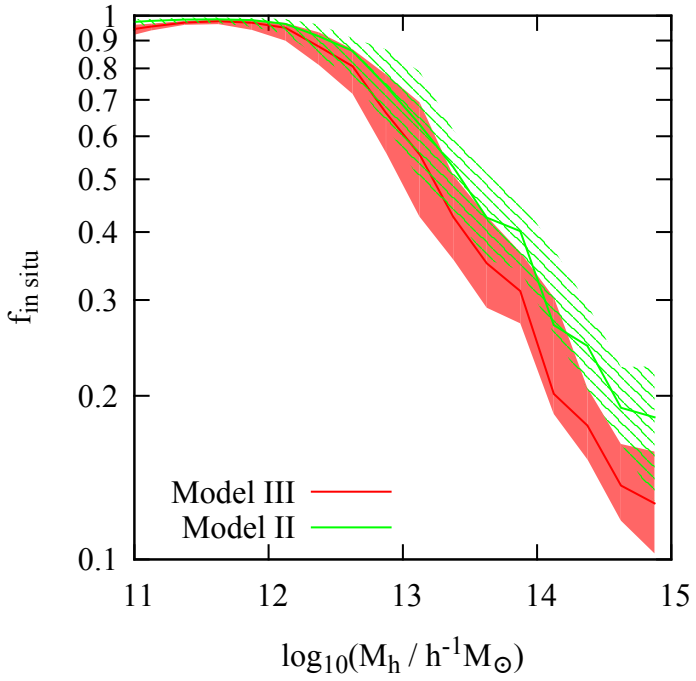

Figure 4.4. The fraction of stars formed *in situ* in present day central galaxies of halos with different masses. Shown are the predictions of Model II (green) and Model III (red). The curves are the averages obtained from an ensemble of halo merger trees, while the shaded areas represent the variances among the different halo merger trees.

The difference in amplitude, which is about a factor of 2, owes to the mass loss of evolved stars. This suggests that the assembly of such galaxies is dominated by *in situ* star formation, rather than by the accretion of stars formed in progenitors. We will have a more detailed discussion about this in the following subsection. For massive cluster galaxies, the assembly histories start with a strong episode of *in situ* star formation at $z > 2$, which is followed by a long period of mass accretion at roughly a constant rate.

4.2.3 In-situ star formation versus accretion

Figure 4.4 shows the fraction of stars formed *in situ* in present day central galaxies as a function of their host halo mass. The predictions of Model II are plotted as the green line (average) and the green shaded area (with the variance arising from different halo merger trees), while the predictions of Model III are plotted in red. The predictions of the

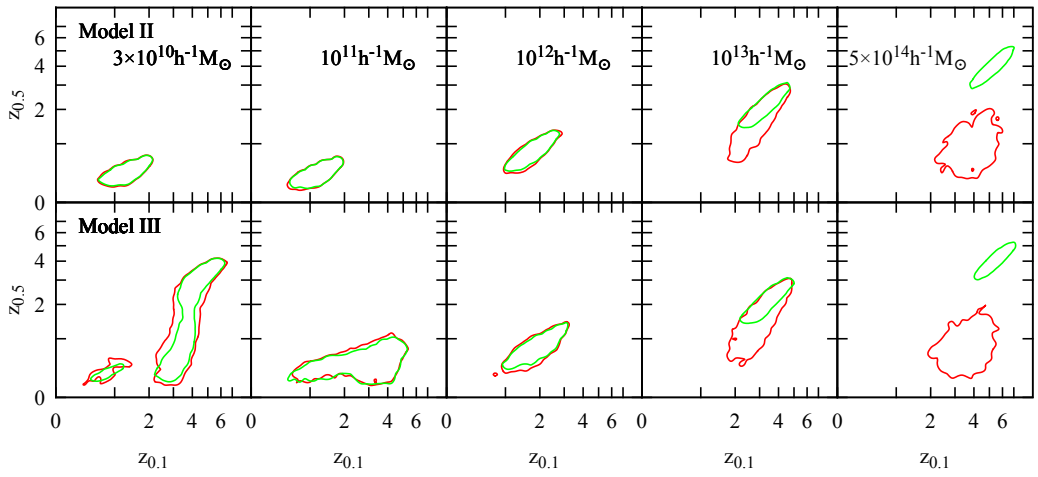

Figure 4.5. The distribution of the redshift at which 50% ($z_{0.5}$) and 10% ($z_{0.1}$) of the stars in a central galaxy have formed (green contours) and have assembled (red contours). The contours are the isodensity lines that enclose 90% of the halo merger trees. We show results for halos with five masses at $z = 0$, as indicated. The upper panels are the predictions of Model II, while the lower panels are for Model III.

two models are quite similar. For central galaxies in halos with masses below $10^{12}\,h^{-1}\mathrm{M}_{\odot}$, almost all the stars are formed *in situ*. This fraction decreases rapidly with increasing halo mass at $M_{\rm h} > 10^{12}\,h^{-1}\mathrm{M}_{\odot}$. About 70% of all stars in the central galaxy of a halo with $M_{\rm h} \sim 10^{13}\,h^{-1}\mathrm{M}_{\odot}$ are formed *in situ*; for cluster halos with $M_{\rm h} \sim 10^{15}\,h^{-1}\mathrm{M}_{\odot}$ this fraction is about 15%, so about 85% of the stellar mass is acquired through accretion.

4.2.4 Downsizing versus upsizing

Galaxies of different masses have different star formation (assembly) histories. To characterise these histories in a more quantitative way, we examine the characteristic redshift, z_f, by which a fraction f of the final stellar mass in a galaxy has formed (or assembled). Figure 4.5 shows the distribution of galaxies in the $z_{0.5}$ - $z_{0.1}$ plane for star formation (green contours) and stellar mass assembly (red contours), with the contours delineating the isodensity lines that contain 90% of all the galaxies. The results are again shown for halos with five different present-day masses, as indicated in the panels, predicted

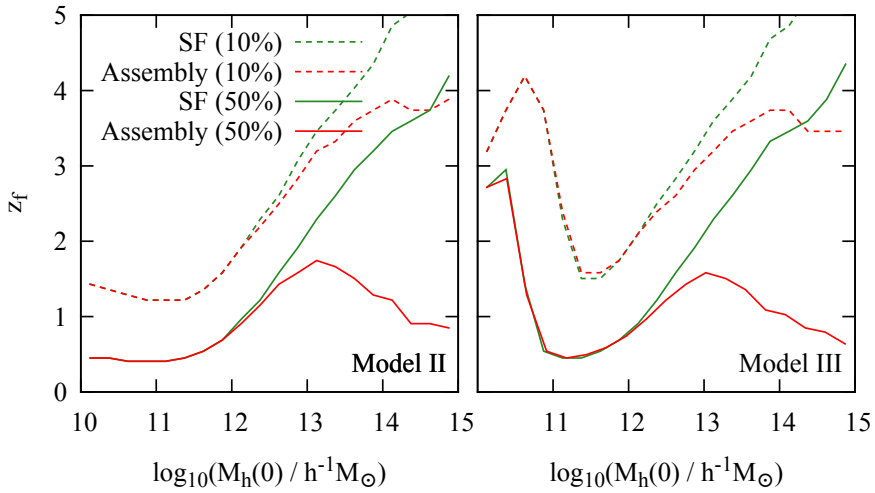

Figure 4.6. The average characteristic redshifts, $z_{0.5}$ and $z_{0.1}$, as a function of halo mass for star formation (green lines) and for stellar mass assembly (red lines).

by Model II (the upper 5 panels) and Model III (the lower 5 panels), respectively. For the most massive halos [$M_{\rm h}(z=0) > 10^{14}\,h^{-1}{\rm M}_\odot$], the star formation time and the assembly time differ considerably, especially in $z_{0.5}$. On average about 50% of the stellar mass in the central galaxies of such massive halos form before $z = 4$, but assemble much later at $z \approx 1$. For halos with masses lower than $10^{13}\,h^{-1}{\rm M}_\odot$, the star formation time and assembly time are almost identical, indicating that central galaxies in such halos acquire their stars mostly through *in situ* star formation, as we have already seen in the last subsection. A Milky Way mass galaxy [$M_{\rm h}(z=0) \sim 10^{12}\,h^{-1}{\rm M}_\odot$] on average formed about 10% of its stars by $z \approx 2$ and about 50% after $z = 1$. For dwarf galaxies residing in halos with $M_{\rm h}(z=0) < 10^{11}\,h^{-1}{\rm M}_\odot$, Model III predicts diverse formation redshifts. For example, the majority of galaxies residing in $10^{11}\,h^{-1}{\rm M}_\odot$ halos are predicted to form 10% of their stars by about $z = 4$, while a fraction is predicted to form their first 10% at much later times: $z \approx 1$. The diversity becomes larger for smaller galaxies: for galaxies in $3 \times 10^{10}\,h^{-1}{\rm M}_\odot$ halos, some show star formation as early as that in the centrals of galaxy clusters, while others formed most of their stars after $z \approx 1$. This diversity owes to the

transition in star formation efficiency at $z \approx 2$ and the variance in the halo accretion histories. halos that formed early generally have experienced an early burst phase of star formation, while younger halos that assembled most of their mass later than the transition redshift did not have such an early star burst.

In Figure 4.6 we show how the averages of $z_{0.1}$ and $z_{0.5}$ change with halo mass for both star formation (green lines) and stellar mass assembly (red lines). For halos with $M_{\rm h}(z = 0) > 10^{11}\,h^{-1}{\rm M}_{\odot}$, both Model II and Model III predict that the centrals of more massive halos on average form a fixed fraction of their stars earlier, a trend usually referred to as "downsizing" (Fontanot et al., 2009; Weinmann et al., 2012). A similar downsizing trend is also seen in the stellar mass assembly for halos with $10^{11}\,h^{-1}{\rm M}_{\odot} < M_{\rm h}(z = 0) < 10^{13}\,h^{-1}{\rm M}_{\odot}$. However, this trend breaks down in two ways. First, for the most massive halos [$M_{\rm h}(z = 0) > 10^{13}\,h^{-1}{\rm M}_{\odot}$], more massive centrals actually assemble their stars later. At late times, these galaxies tend to build up their mass hierarchically by accreting satellite galaxies, resembling the mass assembly history of the dark matter halos themselves. Second, although the downsizing trend holds for dwarf galaxies in Model II, the trend predicted by Model III is completely the opposite for these galaxies: on average smaller galaxies tend to be older. This again owes to the boost of star formation in low-mass halos at high z in Model III.

4.3 Merger history and the transformation of galaxies

In this section we characterise the galaxy merger histories in more details and use a simple model to study their implications for the morphological transformation of galaxies.

We make the simple assumption that stellar disks form only through *in situ* star formation, and that a major merger can transform a stellar disk into a spheroid. Here, major mergers are defined in the same way as in §4.1. The mass of the stellar disk is then simply the total mass of stars formed *in situ* after the last major merger. It should be cautioned, however, that a galactic bulge can be formed in other ways, such as the secular

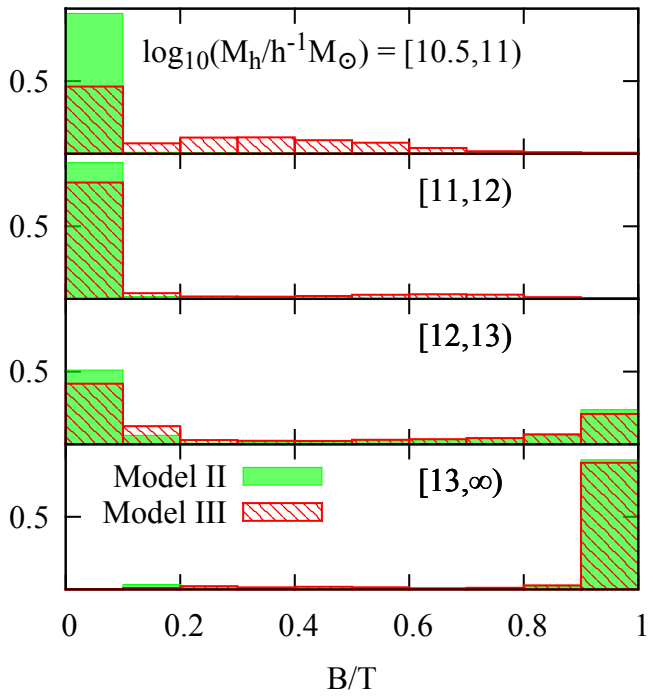

Figure 4.7. The bulge-to-total mass ratio of local central galaxies as a function of host halo mass. The green histograms are the predictions of Model II while the red hedged histograms are those of Model III.

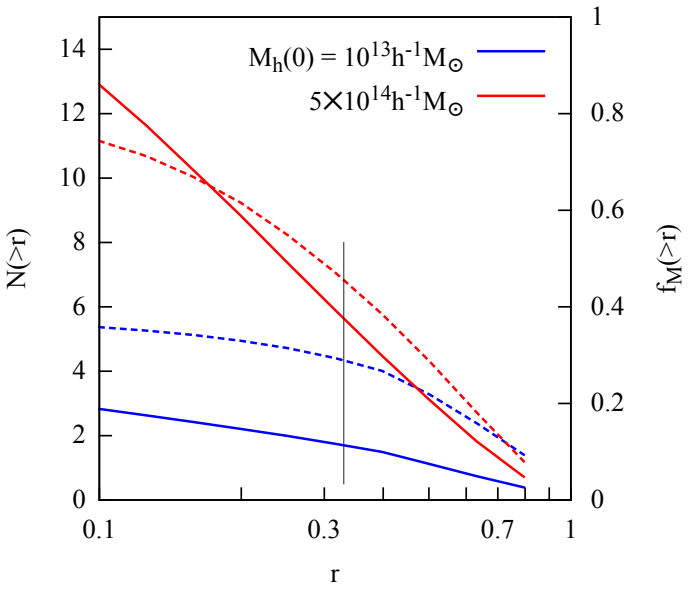

Figure 4.8. The merger histories of present-day galaxies. Solid curves: the number of merger events since the *in situ* SFR declines to 1/3 of the peak value with a mass ratio $> r$, where r is the stellar mass ratio between the secondary and primary merging progenitors. Dashed curves: the fraction of the total stellar mass acquired by mergers with a mass ratio $> r$. The vertical line marks $r = 1/3$, which separates major from minor mergers.

evolution of the disk. The bulge mass defined here, therefore, can only be taken as a lower limit and serves as a simple indicator as to how a central galaxy may be disturbed by infalling satellites. Figure 4.7 shows the distribution of 'bulge-to-total' ratio, B/T, for present-day central galaxies in four halo mass bins.

In cluster halos, the central galaxies are identified as BCGs in observations. They have been bombarded quite frequently by satellites of different masses. As shown in Figure 4.8, such galaxies on average have experienced about 5 major mergers and an even larger number of minor mergers during the last period of low *in situ* star formation (defined, quite arbitrarily, as the period after the *in-situ* SFR declines to be 1/3 of the peak value). Major mergers contribute about 50% of the total stellar mass. This casts doubt on the scenario in which the late growth and structure of BCGs are assumed to be determined by minor mergers (e.g. Bezanson et al., 2009; van Dokkum et al., 2010; Hilz et al., 2013). The last major mergers happened quite recently, at $z \sim 0.5$, as shown in

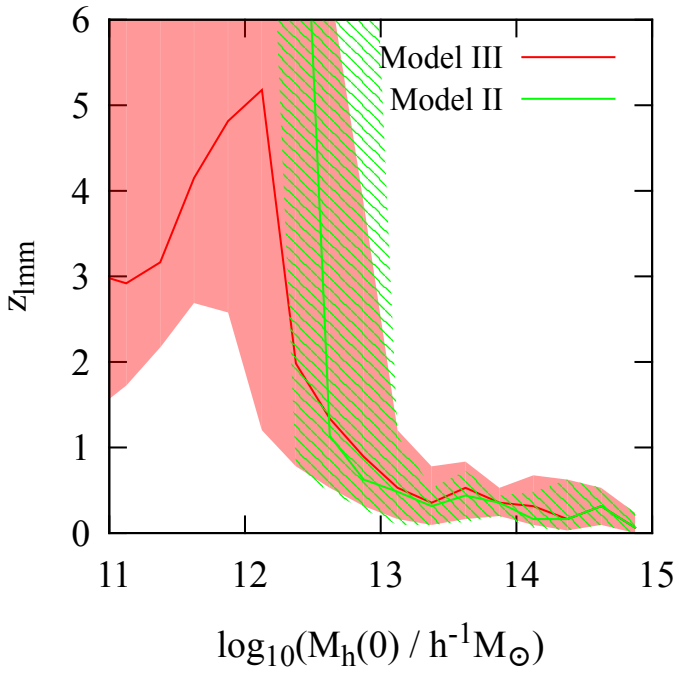

Figure 4.9. The redshift of the last major merger of present-day central galaxies as a function of host halo mass. The predictions of Model II and Model III are shown in green and red, respectively. Lines are the averages, while the shaded areas enclose the 95% range of variance owing to different merger histories.

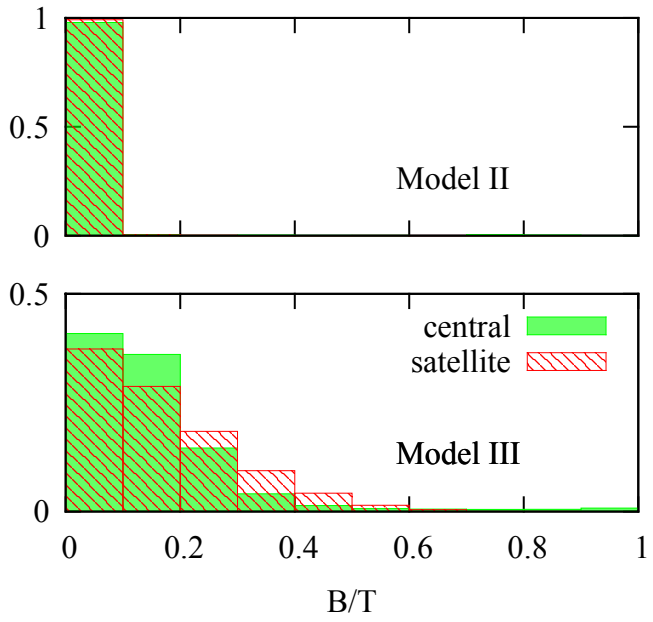

Figure 4.10. The "bulge" mass fraction of dwarf galaxies with total stellar masses in the range $10^8\,h^{-2}\mathrm{M}_\odot$ to $10^9\,h^{-2}\mathrm{M}_\odot$, predicted by Model II (upper panel) and Model III (lower panel) for centrals (green) and for satellites (red).

Figure 4.9. By inspecting images of BCGs at low z, McIntosh et al. (2008) and Jimmy et al. (2013) did find that a significant fraction of them indeed show signatures of recent major mergers, consistent with our predictions.

For halos in the mass range $10^{12}\,h^{-1}\mathrm{M}_\odot$ to $10^{13}\,h^{-1}\mathrm{M}_\odot$ there is strong bimodality in their B/T distribution. Significant mergers are sparse in their entire histories: the 'ellipticals' in such low mass groups on average only experienced ≈ 1.5 major mergers and no more than 3 mergers with a mass ratio > 0.1 (see Figure 4.8). The central galaxy is either dominated by bulge, if it has experienced a recent major merger, or remains disk dominated, if such a merger did not occur.

For a Milky Way sized galaxy with $M_\mathrm{h}(0) = 10^{11.5}$-$10^{12.5}\,h^{-1}\mathrm{M}_\odot$, *in situ* star formation has dominated, while major mergers have been rare since $z \sim 2$. These galaxies, therefore, remain disk dominated, free of any significant major merger - driven bulge components.

The predictions of Model II and Model III differ when it comes to dwarf galaxies. Model II, consistent with many other similar models in the literature, indicates major mergers between dwarf galaxies are extremely rare. Therefore, all the stars are expected to remain in a disk (see the upper panel of Figure 4.7). In contrast, Model III predicts that most of the galaxies experienced some major mergers during their initial star burst phases ($z > 2$). The exact time when a major merger occurs has large variations for such galaxies, as shown in Figure 4.9. Major mergers are very rare at $z < 2$ while *in situ* star formation continues, allowing the growth of new disks. The fraction of stars contained in the spheroid depends both on when the last major merger occurs and on the *in situ* star formation that follows. This complexity in star formation history results in diverse morphologies of present day dwarfs, as shown in Figure 4.7.

Major mergers at early stages of the evolution may shed light on the origin of dwarf ellipticals (dE) and dwarf spheroidal (dSph) galaxies. One popular scenario is galaxy harassment (Moore et al., 1996), in which high speed encounters of a dwarf disk with other galaxies in a dense environment heats up the disk and transform it into a dE or a dSph. However, the predicted kinematics, which shows significant rotation, is at odds with the the observational results (e.g. Geha et al., 2003; Toloba et al., 2013). Our result here suggests that some of the slow rotators could be the remnants of early major mergers between dwarf galaxies. The lower panel of Figure 4.10 shows B/T of both dwarf centrals and satellites predicted by Model III, with the latter skewed towards higher B/T. In contrast, Model II indicates that all dwarf galaxies are strongly disk dominated. However, the predicted fraction of bulge-dominated dwarfs by Model III may be too low to account for the total population of dE's and dSph's. It is likely that later evolution, such as harassment, also plays a role in transforming the disk component, making the bulge components more dominant.

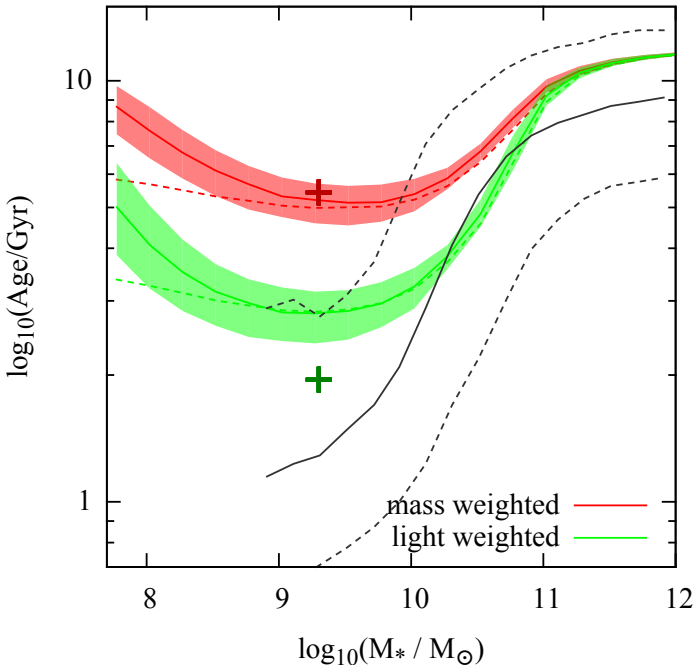

Figure 4.11. The stellar age as a function of stellar mass predicted by Model II (colored dashed lines) and Model III (colored solid lines). Both mass-weighted and light (luminosity) -weighted averages are shown, with the bands representing the 95% percentile of halo merger histories. The black lines show the light-weighted averages of stellar ages obtained by Gallazzi et al. (2005), with the solid line being the median and the two dashed lines representing the 16% and 84% percentile. The crosses are the mass-weighted (dark red) and light-weighted (green) ages of the LMC obtained from the star formation history given by Weisz et al. (2013).

4.4 Stellar populations

4.4.1 Stellar ages

Observationally, the star formation histories of individual galaxies can be estimated from their stellar populations. One popular way to do this is to use the spectra of galaxies, which contain information about the star formation history, IMF, chemical enrichment history and dust content of galaxies. In Figure 4.11, we compare the average stellar ages predicted by Model II and Model III with the observational results of SDSS galaxies obtained by Gallazzi et al. (2005). The observational stellar ages are determined by simultaneously fitting five spectral absorption features that break some degeneracies in

spectral synthesis parameters. The weighted age of a galaxy is defined as

$$\mathrm{Age} = \frac{\int_0^\infty \mathrm{SFR}(t) f(t) t \, dt}{\int_0^\infty \mathrm{SFR}(t) f(t) \, dt} \,, \tag{4.4}$$

where $f(t)$ is either the fraction of the remaining stellar mass or the luminosity of a simple stellar population with age t. In Figure 4.11 we plot the average ages weighted either by stellar mass (red bands) or by the r-band luminosity (green bands). The observational ages, with the median shown by the solid line and the 16 and 84 percentile by the dashed lines, are also weighted by the r-band luminosity, so a direct comparison between the model and the data can be made. The SDSS galaxies cover a stellar mass range between 10^9 and $10^{12}\,\mathrm{M}_\odot$, where the observed age increases with stellar mass generally, and increases sharply around $10^{10}\,\mathrm{M}_\odot$. The model predictions agree with the observational data only qualitatively: the predicted ages are systematically older than that observed between $10^9\,\mathrm{M}_\odot$ and $10^{10}\,\mathrm{M}_\odot$. Similar discrepancies have also been found by Lu et al. (2014b) using semi-analytic models.

This discrepancy may indicate an intrinsic deficiency in the approach adopted here or in similar approaches in the literature. By matching only SMFs of galaxies, our model may be insensitive to the star formation history in the recent past. For example, an enhancement in recent star formation may contribute little to the total stellar mass, and so is not well captured in our model, but can significantly increase the optical luminosity, thereby decreasing the light-weighted age. As a demonstration, the upper cross in Figure 4.11 shows the stellar mass weighted age of stars in the LMC based on the star formation history derived from the colour magnitude diagram of resolved stars (Weisz et al., 2013). Our model predictions match the observation well. For comparison, the lower cross shows the r-band luminosity-weighted age obtained from the same star formation history with our adopted spectral synthesis model. The age so obtained lies below the green band, because our average model underpredicts the current SFR of the LMC.

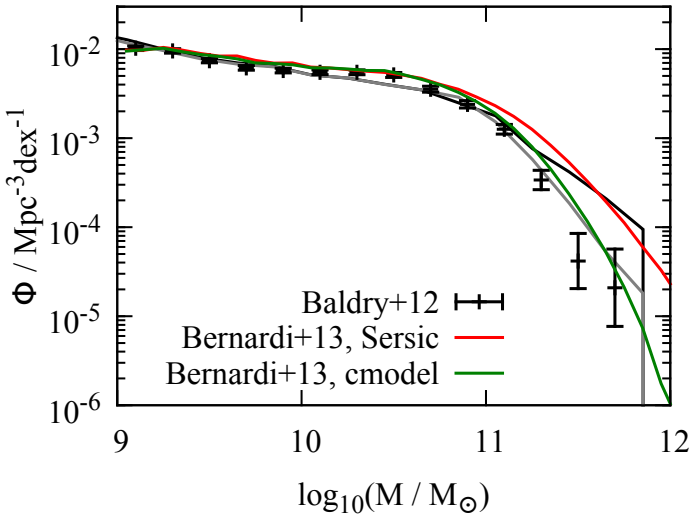

Figure 4.12. The SMF of galaxies predicted by Model III (grey line) compared to the prediction of the same model but assuming no stripping of satellites, i.e. $f_{\rm TS} = 1$ (the thick black line). The data points are the observational data of Baldry et al. (2012). The green line is the SMF obtained by Bernardi et al. (2013) using SDSS 'cmodel' magnitudes, while the red line is the result obtained by the same authors using the magnitudes within the entire Sersic profiles (i.e. integrated to infinity) obtained by Simard et al. (2011).

There are uncertainties in the observational data too. It is in general difficult to distinguish stars that formed about 8-10 Gyrs ago from those that formed 4-5 Gyrs ago from an analysis of the optical-NIR spectra (e.g. Bruzual and Charlot, 2003; Pacifici et al., 2012). Basically the UV light provides information about recent star formation, while strong Balmer absorption lines are sensitive only to intermediate-age (1-3 Gyrs old) stars, but one cannot distinguish stars older than 4 Gyrs. Thus, it is possible that the average stellar ages derived from the observational data have missed the contribution of such an old population.

4.4.2 Halo Stars

In our simple prescription for galaxy mergers, a constant fraction $f_{\rm TS}$ of the original stellar mass of the merging satellite is accreted by the central, and the rest is deposited in a diffuse component, referred to as intracluster stars in clusters or as halo stars in general.

$f_{\rm TS}$ is constrained to be about 40%, although the uncertainty is quite large (Tables 2.2 and 2.3). This low value of $f_{\rm TS}$ is driven predominantly by the observed SMFs at the massive end, i.e. with $M_\star \gtrsim 10^{11}\,{\rm M}_\odot$ corresponding to $M_{\rm h} > 10^{13}\,h^{-1}{\rm M}_\odot$. Had we set $f_{\rm TS} = 1$, i.e. if all the original mass in an accreted satellite were added to the central, we would get a SMF that is significantly higher than the observations of Baldry et al. (2012) at the very massive end (thick black line in Figure 4.12).

Recent analyses show that the most massive galaxies (with $M_\star > 10^{11}\,{\rm M}_\odot$) tend to have extended wings in their light profiles. This makes the luminosity (stellar mass) measurements of these galaxies quite uncertain. As shown by Bernardi et al. (2013) and He et al. (2013), using different methods and light profiles to fit galaxy images can lead to a factor of 2 difference in the estimated luminosity of a massive galaxy. The effect on the derived SMF can be seen in Figure 4.12 comparing the red line with the green line. It is interesting to note that our model prediction assuming $f_{\rm TS} = 1$ is consistent with the SMF obtained by Bernardi et al. (2013) using the magnitudes within the entire Sersic profiles (the red line) obtained by Simard et al. (2011). This suggests that the halo component defined in our model may simply be the extended profiles of massive galaxies that are missing in the SMFs we use as constraints.

Figure 4.13 shows the ratio between the total mass of halo stars and the stellar mass of the central galaxy as a function of host halo mass. In a cluster as massive as $10^{15}\,h^{-1}{\rm M}_\odot$, the mass in the diffuse component is about 2 - 3 times as high as that of the central galaxy and in a group sized halo of mass $\sim 10^{13}\,h^{-1}{\rm M}_\odot$, the ratio is about 1. It drops rapidly towards lower halo mass. In a Milky Way mass halo, the predicted ratio is only a few percent. The flattening at the even lower mass end predicted by Model III owes to the boosted star formation at high z in low-mass halos. The results for low-mass halos should be taken with caution. As mentioned above, our model assumes a constant $f_{\rm TS}$ and it is constrained primarily by the SMFs at the massive end. It is unclear whether the same number also applies to galaxies with lower masses.

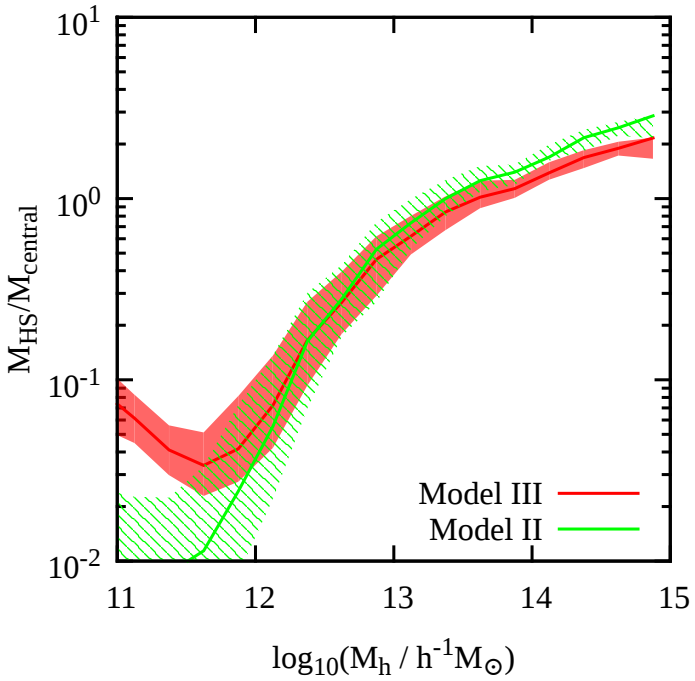

Figure 4.13. The mass ratio between halo stars and the stellar mass of the central galaxy as a function of host halo mass at $z = 0$. The green band is the prediction of Model II and the red band that of Model III. Here again the widths of the bands represent the variance among different halo merger trees.

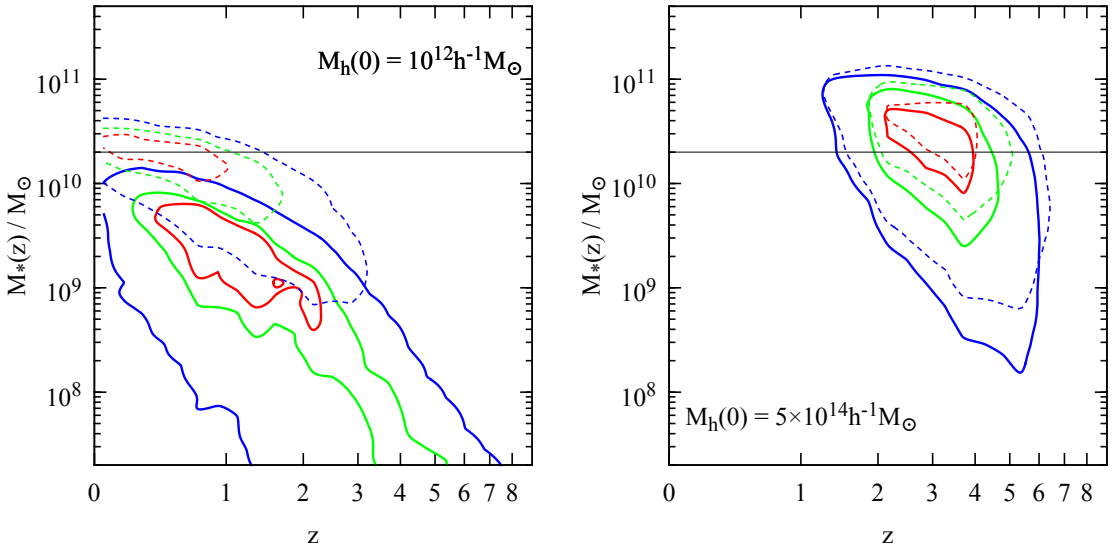

Figure 4.14. The solid contours show the formation redshift and location (i.e. the stellar mass of the galaxy) of the halo stars. Red, green and blue contours enclose 30%, 60% and 90% of the stars, respectively. The dashed contours are the same as the solid ones but for stars in central galaxies. We only show Model III (Model II is similar). Results are shown for Milky Way mass halos and massive cluster halos as indicated. The horizontal lines mark $M_\star = 2 \times 10^{10}\,\mathrm{M}_\odot$, above which the stellar metallicity begin to saturate (Gallazzi et al., 2005).

To understand the stellar population of the halo stars, we make a census of their formation time and location, in terms of redshift and the stellar mass of the host galaxy in which they formed, respectively. We show the result as the solid contours in Figure 4.14. For comparison, we also plot the results for stars in the central galaxies as the dashed contours. The stellar mass can be taken as crude proxy to stellar metallicity as the two are found to be correlated for local galaxies (Gallazzi et al., 2005). The stellar metallicity increases with stellar mass for galaxies with stellar masses below $2 \times 10^{10}\,\mathrm{M}_{\odot}$ and becomes saturated at about solar metallicity above this mass (indicated by the horizontal lines in Figure 4.14).

Compared to the stars in the central galaxy, the halo stars in a Milky Way sized halo form a distinct population: the mass weighted age of the halo stars is roughly about 9-10 Gyr, in contrast with the central galaxy, which is about 6 Gyr. These halo stars formed in progenitors with mass lower than the horizontal line, suggesting a metallicity much lower than the solar value. This is qualitatively consistent with the recent observations of M31 (Bernard et al., 2015). In massive clusters, on the other hand, the stellar populations in both the central galaxy and the halo component are quite homogeneous, with ages of ~ 10 Gyr and with nearly solar metallicity. This prediction can be checked by studying the stellar age and metallicity of halo stars in clusters.

4.5 Summary and Discussion

We use our constrained model parameters to characterise the star formation and stellar mass assembly histories as well as the merger histories of galaxies of different masses. The results are summarised in the following.

First, the evolution of the galaxy population is found to be characterised by a number of characteristic halo mass scales:

1. A mass scale of $10^{13}\,h^{-1}\mathrm{M}_{\odot}$ at $z = 0$, decreasing to $\sim 3 \times 10^{12}\,h^{-1}\mathrm{M}_{\odot}$ at $z > 2$, above which *in situ* star formation drops rapidly, and the central galaxies experience frequent major mergers most of which will be dry.

2. A mass scale of $3 \times 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ at $z = 0$, increasing to $10^{12}\,h^{-1}\mathrm{M}_{\odot}$ at high z, at which the efficiency of *in situ* star formation reaches a maximum, with a SFR as high as about half of the baryon accretion rate into the host halo. Major mergers are rare in this halo mass range.

3. A mass scale of $10^{11}\,h^{-1}\mathrm{M}_{\odot}$ at $z > 3$, below which *in situ* star formation has a rate about 0.1 times the baryon accretion rate into the host halo. On average, one or two major mergers are expected to occur for the central galaxies.

Second, galaxies hosted by halos of different masses follow distinct star formation and assembly histories. Based on the characteristic halo masses given above, central galaxies can be divided roughly into three different categories according to their formation and assembly histories:

1. For halos with $M_{\mathrm{h}} > 10^{13}\,h^{-1}\mathrm{M}_{\odot}$, a strong *in situ* star formation rate declines rapidly after reaching its peak value, and is followed by significant accretion of stars from satellites. For such massive systems, more massive galaxies tend to assemble their stellar mass later, contrary to the downsizing trends observed for lower mass galaxies.

2. For halos with masses $10^{11}\,h^{-1}\mathrm{M}_{\odot} < M_{\mathrm{h}} < 10^{13}\,h^{-1}\mathrm{M}_{\odot}$, mass assembly by accretion of satellites is not important. The star formation is delayed relative to the formation of the host halo. For example, the progenitors of present-day Milky-Way mass galaxies at $z \approx 2$ is about 1/10 of the final stellar mass, while the average halo mass at the same redshift is about 1/4 of the final halo mass.

3. For halos with masses below $10^{11}\,h^{-1}\mathrm{M}_{\odot}$, assembly by accretion is again unimportant. The star formation history is characterised by a burst at $z > 2$ and a nearly

constant star formation rate after $z = 1$. The relative importance of the early star formation increases with decreasing halo mass, and the 'downsizing' trend is reversed. The model, therefore, predicts the existence of an old stellar population formed at $z > 2$ in dwarf galaxies, which is consistent with the results obtained form direct observation of the stellar populations in such galaxies (Weisz et al., 2011).

Third, we use the merger history of the model galaxies to predict the bulge to total mass ratios of present galaxies. The average bulge mass fraction is found to depend strongly on halo mass:

1. In cluster sized halos with $M_{\rm h} > 3{\times}10^{13}\,h^{-1}{\rm M}_{\odot}$, almost all the centrals are ellipticals formed through frequent major mergers.

2. In group sized halos with masses between $3 \times 10^{12}\,h^{-1}{\rm M}_{\odot}$ and $3 \times 10^{13}\,h^{-1}{\rm M}_{\odot}$, the distribution of the bulge-to-total ratio of the central galaxies is strongly bimodal. Those galaxies that experienced a recent major merger are spheroid dominated, whereas the others are free of any significant merger-driven bulge.

3. For halos with masses $3 \times 10^{11}\,h^{-1}{\rm M}_{\odot} < M_{\rm h} < 3 \times 10^{12}\,h^{-1}{\rm M}_{\odot}$, central galaxies with a significant merger-driven bulge are extremely rare.

4. For dwarf galaxies, half of them have significant (with $B/T > 10\%$) spheroidal components formed during their early star burst phase ($z > z_c \approx 2$). Satellite galaxies of similar masses tend to have a larger bulge fraction than centrals.

We emphasise again that bulges can form in various other ways than through major mergers, and that our prediction only applies to major merger-driven bulges.

Finally, we have made predictions for the amount of halo stars, and when and where these stars form in comparison with stars in the corresponding central galaxies. The results are:

1. In a Milky Way mass halo, the total mass in halo stars is 2 to 5 percent of the mass of the central galaxy, and this number increases to $\sim$ 100% in halos with $M_{\rm h} \sim 10^{13}\,h^{-1}{\rm M}_{\odot}$, and to about 200% to 300% in massive clusters.

2. In a Milky Way mass halo, the stars in the central galaxy and halo stars form two distinct stellar populations, with the latter being older and poorer in metals. In contrast, these two components form a quite homogeneous population (old and with solar metallicity) in massive clusters.

All these results are obtained in an empirical way, independent of any detailed assumptions about the underlying physical processes that drive the evolution of the galaxy population. Clearly, our results should be compared with the predictions of numerical simulations and/or semi-analytical models to constrain theories of galaxy formation.

CHAPTER 5

CLASSICAL BULGES, SUPERMASSIVE BLACKHOLES AND AGN FEEDBACK: EXTENSION TO LOW-MASS GALAXIES

5.1 Introduction

The co-existence of supermassive black holes (hereafter SMBH) and early type galaxies, such as elliptical galaxies and the bulges of spiral galaxies, have been found ubiquitous (e.g. Kormendy and Richstone, 1995). Tight correlation has been observed between the SMBH mass and the stellar mass and velocity dispersion of the host spheroid component. The SMBH mass is roughly proportional to the spheroid mass, with a ratio between the two about 0.1% (McLure and Dunlop, 2002; Häring and Rix, 2004) to 0.5% (see Kormendy and Ho, 2013, for a review).

The observed scaling relation suggests that the growths of both the SMBH and the spheroid components are driven by similar physical processes. The most popular scenario is mergers of galaxies, in which a large amount of cold gas originally supported by a rotation disk can lose angular momentum and sink toward the center to fuel both the SMBH growth and the star formation. In addition, the energy feedback from the AGN associated with the SMBH accretion may heat and/or eject the cold gas, suppressing both star formation and SMBH growth (e.g. Springel et al., 2005a). Indeed, with plausible assumptions about AGN feedback and about how star formation and SMBH growth are related to the cold gas density, numerical simulations and semi-analytical models can reproduce the general trend in the observed scaling relation (e.g. Croton et al., 2006; Di Matteo et al., 2008). In addition to the merger scenario, other mechanisms have also been proposed, such as the direct accretion of low angular momentum cold flow (Di Matteo

et al., 2012; Dubois et al., 2012), the intense gravitational instability in clumpy disks (Bournaud et al., 2011), and some secular processes (Dubois et al., 2014). However, it is not clear how much these mechanisms contribute to the connection between SMBHs and their host galaxies.

Recent observations have extended the detections of SMBH and their mass measurements to late-type galaxies and to galaxies in the low-mass end (e.g. Filippenko and Ho, 2003; Greene and Ho, 2004; Greene and Ho, 2007; Greene et al., 2008; Jiang et al., 2011a; Jiang et al., 2011b). These observations indicate a number of interesting deviations from the scaling relation derived from massive, early-type galaxies. First, it was found that many late-type spiral galaxies, which do not have a significant classical bulge component but contain a pseudo-bulge formed through secular evolution of the disk, do contain central SMBH. At a fixed bulge mass (luminosity) the SMBH masses are about a factor of 5 to 10 lower than expected from the correlation seen for classical bulges, and the scatter is also significantly larger (see Kormendy and Ho, 2013). Or equivalently, the correlation between the central BH and the bulge is dramatically bent (Graham, 2012; Graham and Scott, 2013; Scott et al., 2013). Second, SMBH have also been observed in dwarf galaxies, such as dwarf spheroids and dwarf ellipticals (e.g. Barth et al., 2004; Reines et al., 2013). It is still unclear whether the SMBH in the late-type and low-mass galaxies have formed through the same processes as in classical bulges, or through some completely different mechanisms.

As shown in §4.3, all galaxies contain a classical bulge component produced by major galaxy-galaxy merger. The significance of such classical bulge depends strongly on the total stellar mass of the galaxy. In galaxies with sub-Galactic masses, this component is old in stellar population, small in stellar mass by an order of magnitude than pseudo-bulges (e.g. Fukugita et al., 1998), and so may be missed in observation owing to the presence of other dominating components. In this chapter, we demonstrate that the SMBH hosted by late-type galaxies can also be explained by the co-evolution with the

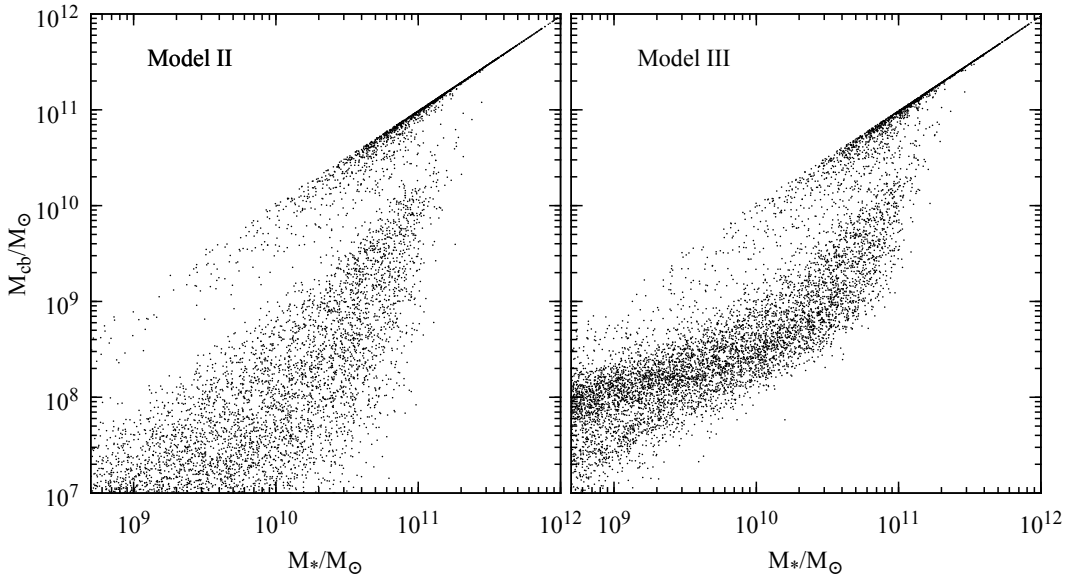

Figure 5.1. Relation between the mass of the classical bulge (formed via merger of galaxies) and the total stellar mass of the host galaxy, predicted by Model II (left) and Model III (right).

classical bulge components potentially hidden in these galaxies. In §5.2, we show the relation between the classical bulge mass and total stellar mass of galaxies predicted by the empirical models, and examine its implications for the SMBH mass - galaxy mass relation. We will see that applying the observed $M_{\rm BH}$ - $M_{\rm cb}$ relation to the modeled galaxies reproduces well the observed relation between $M_{\rm BH}$ and the total stellar mass of the host galaxies. This suggests that the co-evolution of SMBH is predominantly with classical bulges formed through mergers, while secular evolution does not play a major role in the growth of SMBH. We study the implications of our results for AGN feedback in §5.3, and summarize in §5.4.

5.2 Implications of galaxy merger histories for supermassive black holes

Figure 5.1 shows the relation between $M_{\rm cb}$ and $M_{\star}$ for central galaxies predicted by Model II (left) and Model III (right). Owing to frequent major mergers, the most massive

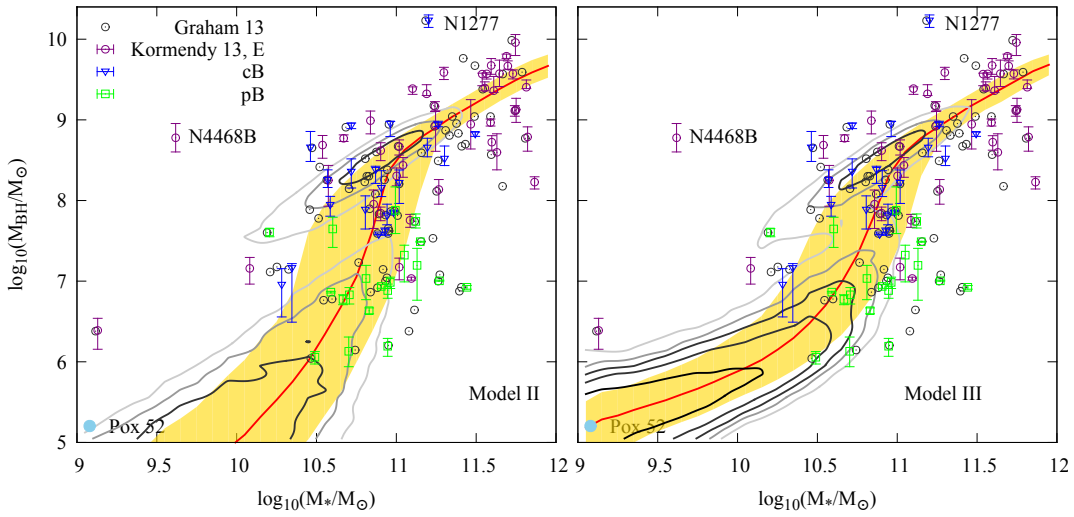

Figure 5.2. Relation between the SMBH mass and the total galaxy stellar mass predicted by Model II (left) and Model III (right). The red line is the median of the model prediction in each stellar mass bin and the yellow band encompasses 90% of the galaxies in the corresponding bins. For comparison, iso-density contours in the $M_{\rm BH}$ - $M_\star$ plane are also shown. The colored data points are compiled by Kormendy and Ho (2013). The two outliers, N4468B and N1277, probably have their stellar masses significantly reduced due to strong stripping. The black open circles are data from Graham and Scott (2013). The light blue point at the bottom left corner is the dwarf Seyfert 1 galaxy Pox 52 (Thornton et al., 2008).

galaxies are completely dominated by the classical bulge components. Both models predict a strong transition at $M_\star \sim 10^{11}\,{\rm M}_\odot$. For galaxies in the mass range $2 \times 10^{10}\,{\rm M}_\odot$ to $10^{11}\,{\rm M}_\odot$, *in-situ* star formation begins to quench at the present time while major mergers are sparse but not negligible. Consequently, these galaxies experience either one or no major merger in the recent past, so that they are either dominated by a classical bulge or remains disk dominated, producing the strong bi-modality in the $M_{\rm cb}$ - $M_\star$ relation shown in Figure 5.1. For dwarf galaxies with $M_\star \ll 10^{10}\,{\rm M}_\odot$, Model II predicts that the mass in the classical bulge component is much smaller than the total stellar mass of the galaxy. Model III, however, predicts much larger masses for the classical bulge components, because of the enhanced star formation in low-mass galaxies and the more frequent galaxy-galaxy major mergers at $z > 2$.

Given that all galaxies are expected to contain classical bulges formed through mergers of galaxies, we test the hypothesis that SMBHs are only due to classical bulges. The correlation between SMBHs and classical bulges from Kormendy and Ho (2013) is

$$\frac{M_{\rm BH}}{10^9\,{\rm M}_\odot} = \left(0.49^{+0.06}_{-0.05}\right)\left(\frac{M_{\rm cb}}{10^{11}\,{\rm M}_\odot}\right)^{1.16\pm0.08} . \tag{5.1}$$

The intrinsic scatter in $M_{\rm BH}$ is about 0.29 dex at fixed $M_{\rm cb}$. We apply this relation to the classic bulges to predict a SMBH mass for each model galaxy. The red curve in Figure 5.2 shows the mean of the predicted SMBH mass as a function of the total stellar mass, and the iso-density contours show the distribution of galaxies in the $M_\star - M_{\rm BH}$ plane. Results are presented for both Model II (left) and Model III (right). The predicted relations are compared with the observational data compiled by Kormendy and Ho (2013), shown as colored points with different styles representing different systems, as well as the catalog from Graham and Scott (2013), shown as black open circles. The dwarf Seyfert 1 galaxy Pox 52 (Thornton et al., 2008) is also plotted (the solid light blue points).

Both models reproduce the strong transition at $M_\star = 10^{10.5}$ - $10^{11}\,{\rm M}_\odot$, below which the value of $M_{\rm BH}$ drops rapidly with decreasing $M_\star$. In the observational data, the transition occurs at a place where pseudo-bulges start to take over classical bulges. In the model predictions, the transition is due to the rapid decline of classical bulge mass with decreasing $M_\star$, as shown in Figure 5.1. For galaxies with $M_\star < 10^{10.5}\,{\rm M}_\odot$, Model II predicts a rapid decline of SMBH mass with decreasing stellar mass, $M_{\rm BH} \propto M_\star^{2.5}$, with large scatter. In contrast, Model III predicts a new $M_{\rm BH} \propto M_\star$ sequence, but with the amplitude dropped by a factor of ~ 50 relative to that at the massive end. At the moment, the small number of observational data points appear to be better matched by Model III, but no reliable conclusion can be reached yet without more data and a better understanding of observational selection effects. It is clear, though, that Model III, which is favored by a large set of observational data, does predicts that many galaxies with sub-Galactic masses host SMBH with $M_{\rm BH} = 10^5$ - $10^6\,{\rm M}_\odot$.

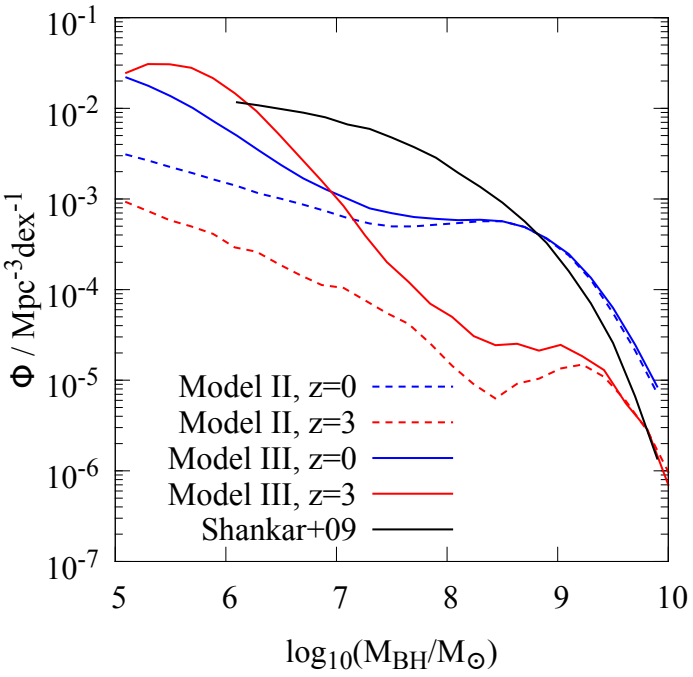

Figure 5.3. SMBH mass function predicted by Model II (colored dashed lines) and Model III (colored solid lines) at $z = 0$ (blue) and $z = 3$ (red). The black line is the mass function of Shankar et al. (2009).

Figure 5.3 shows the SMBH mass functions predicted by Model II (dashed curves) and Model III (solid curves) at $z \approx 0$ (blue) and $z \approx 3$ (red). For $z = 3$ we have adopted a $M_{\rm BH}/M_{\rm cb}$ ratio that is three times as large as that for $z = 0$ (Kormendy and Ho, 2013). For $M_{\rm BH} > 10^8\,{\rm M}_\odot$ and $z = 0$, the predictions of the two models are almost identical. For $M_{\rm BH} < 10^7\,{\rm M}_\odot$, however, the mass function predicted by Model III is much steeper than that predicted by Model II, particularly at high z. According to Model III, most SMBH in this mass range, which are hosted by galaxies with sub-Galactic masses, have already in place at $z \approx 3$. They formed together with their host galaxies in dwarf halos with $M_{\rm h} \lesssim 10^{11}\,{\rm M}_\odot$ at $z > 2$, where star formation is efficient and gas-rich mergers of galaxies are still frequent (§ 4.1). The consequence of such SMBH formation for the evolution of their host galaxies is discussed in the next section.

For comparison, we also plot the SMBH mass function estimated by Shankar et al. (2009), which is widely adopted in the literature. Without distinguishing between pseudo-bulges and classic bulges, their estimates assumed that all late-type galaxies have a bulge-to-total light ratio of ≈ 0.27 (Fukugita et al., 1998). This ratio is about 7 times larger than our prediction for classical bulges, and is also the reason why their mass function is much higher than ours at $M_{\rm BH} \sim 10^{7.5}\,{\rm M}_\odot$.

5.3 AGN feedbck from low-mass galaxies

As a SMBH grows, the power of the energy output can be written as

$$\frac{d\,E}{d\,t} = \epsilon \dot{M}_{\rm BH} c^2\,, \tag{5.2}$$

where ϵ is an efficiency factor. The total energy output is

$$E = \bar{\epsilon} M_{\rm BH} c^2\,, \tag{5.3}$$

where $\bar{\epsilon}$ is the mean efficiency. Since in our model $M_{\rm BH}$ is roughly proportional to $M_{\rm cb}$, and the energy feedback from star formation is also $\propto M_{\rm cb}$, the above equation (and our

modeling below) also applies even if the feedback effect of classical bulge stars is taken into account.

Suppose a fraction f_ϵ of this energy output is eventually transferred to and retained in a total amount of gas of mass $M_{\rm gas}$ and write this mass in terms of the mass of the host halo: $M_{\rm gas} = \lambda f_{\rm B} M_{\rm h}$, with λ a constant loading factor and $f_{\rm B} \approx 0.17$ the universal baryon fraction. The effective temperature of the gas due to the energy transfer can then be written as

$$\frac{kT\lambda f_{\rm B} M_{\rm h}}{\mu} = f_\epsilon \bar{\epsilon} M_{\rm BH} c^2 \,, \tag{5.4}$$

where μ is the mean molecular weight of the gas. This can be written in a more useful form:

$$\frac{kT}{\mu} = \frac{f_\epsilon \bar{\epsilon} c^2}{\lambda f_{\rm B}} \left(\frac{M_{\rm BH}}{M_\star}\right) \left(\frac{M_\star}{M_{\rm h}}\right) . \tag{5.5}$$

For dwarf galaxies with stellar masses between 10^8 and $10^9\,{\rm M}_\odot$ at $z \sim 3$, which were the progenitors of present day sub-Galactic galaxies ($10^9 - 10^{10}\,{\rm M}_\odot$), $M_\star/M_{\rm h} \approx 10^{-2.5}$ (§ 3). These progenitors are typically bulge dominated because of galaxy-galaxy major mergers. If we adopt a bulge to total ratio of 1/2 and scale Eq. (5.1) up by a factor of 3 for these high-redshift progenitors (Kormendy and Ho, 2013), we get $M_{\rm BH}/M_\star \approx 0.3\%$. Assuming μ to be 0.6 times the proton mass, and $\bar{\epsilon} = 0.1$ as is usually assumed for an AGN in the quasar mode, we have

$$T_6 \approx 1.0 \times \left(\frac{f_\epsilon}{0.03}\right) \left(\frac{\lambda}{10}\right)^{-1} , \tag{5.6}$$

where $T_6 \equiv T/10^6{\rm K}$.

For the current cosmology, the age of the universe at $z > 1$ is roughly

$$t \approx 3.0 \times \left(\frac{1+z}{3}\right)^{-3/2} {\rm Gyr} \,, \tag{5.7}$$

and the cooling time of the gas medium with an over-density δ and temperature T at z is

$$t_{\rm cool} \approx 3.0 \times \left(\frac{1+z}{3}\right)^{-3} \Lambda_{-23}^{-1} T_6 \left(\frac{1+\delta}{60}\right)^{-1} {\rm Gyr} \,, \tag{5.8}$$

where Λ_{-23} is the cooling function in units of $10^{-23}\mathrm{erg\,s^{-1}cm^3}$. Setting $t_{\rm cool} = t$ and using the fact that $\Lambda_{-23} \approx T_6^{-1}$ for a low-metallicity gas with temperature in the range 10^5 - 10^6K, we get

$$T_6 \approx 1.0 \times \left(\frac{1+z}{3}\right)^{3/4} \left(\frac{1+\delta}{60}\right)^{1/2} . \tag{5.9}$$

This temperature corresponds to a specific entropy

$$\begin{aligned} S &\equiv \frac{kT}{n_{\rm e}^{2/3}} \approx 15.0 \times T_6 \left(\frac{1+\delta}{60}\right)^{-2/3} \left(\frac{1+z}{3}\right)^{-2} \mathrm{keV\,cm^2} \\ &\approx 15.0 \times \left(\frac{1+\delta}{60}\right)^{-1/6} \left(\frac{1+z}{3}\right)^{-5/4} \mathrm{keV\,cm^2} . \end{aligned} \tag{5.10}$$

Note that S depends only weakly on δ at a given z.

If the medium is heated above the temperature given by Eq. (5.9), gas cooling will be suppressed. Thus, for the AGN feedback to have a significant impact on subsequent galaxy formation and evolution, the temperature given by Eq. (5.6) should be higher than that given by Eq. (5.9). This gives a constraint on λ,

$$\frac{\lambda}{10} \lesssim \left(\frac{f_\epsilon}{0.03}\right) \left(\frac{1+z}{3}\right)^{-3/4} \left(\frac{1+\delta}{60}\right)^{1/2} . \tag{5.11}$$

Numerical simulations show that to reproduce the observed $M_{\rm BH}$ - galaxy velocity dispersion relation requires 5% of the total energy output be coupled with the surrounding gas (e.g. Di Matteo et al., 2008). The energy will propagate through the host galaxies and the host halos as blast waves or super-bubbles in which radiative cooling is expected to be slow because of the low-density of the halo gas. So f_ϵ is not expected to be much lower than a few percent. The over-density $\delta \approx 60$ is about the value appropriate for the exteriors of dark matter halos.[1] Finally, as shown in the previous section, most of the SMBH in the dwarf galaxies formed at $z > 2$. Setting $f_\epsilon \approx 0.03$, $z \approx 3$ and $\delta \approx 60$ gives

[1]For a singular isothermal halo defined so that the mean density within its virial radius is 200 times the mean density of the universe, the over-density at the virial radius is $\delta \approx 66$.

$\lambda \lesssim 8$. As shown in Zhao et al. (2009), a low-mass halo has typically increased its mass by a factor of 5 since $z \sim 3$. Thus, the AGN feedback is, in principle, capable of affecting all the gas to be accreted by a halo in its subsequent growth.

How might such a feedback proceed? In general, the impact of the feedback from an AGN is expected to first affect the gas close to the galaxy. Since the IGM is expected to be clumpy and filamentary, the dense part of the medium may still be able to cool and circumvent total disruption by the wind. This part of the gas can then move against the outflow toward the halo center to feed the galaxy and SMBH. The growth of a SMBH and star formation may thus go through a number of bursts, as the dense clouds are accreted episodically. As the local medium is heated and expands, an negative gradient of entropy may develop, producing gas convection and mixing gas of different specific entropies. Eventually a large envelope of roughly constant specific entropy may develop around each low-mass halo. This process continues until most of the gas in the IGM around the galaxy has an specific entropy such that its radiative cooling time is comparable to the age of the universe. As demonstrated above, the likely epoch for this to accomplish is at $z \sim 2$ to 3, and the specific entropy is of the order of 10 - 15 $\mathrm{keV\,cm^2}$. The subsequent formation and growth of dark matter halos are then in a preheated medium where gas accretion and cooling in low-mass halos can be reduced significantly. This scenario of galaxy formation in a preheated medium was first proposed in Mo and Mao (2002). As shown in Mo et al. (2005), Lu and Mo (2007) and Lu et al. (2015a), the level of specific entropy predicted here is roughly what is needed to suppress gas accretion and star formation in low-mass halos to match the observed stellar mass and HI-mass functions at the low-mass end.

5.4 Discussion

In this chapter we have shown that the observed SMBH masses are consistent with the assumption that they are directly proportional to the mass of the classical bulges of their hosts, even in late-type galaxies. In particular, this assumption combined with the

classical bulge masses reproduces well the rapid decrease of $M_{\rm BH}$ with decreasing $M_\star$ at $M_\star \lesssim 10^{11}\,{\rm M}_\odot$. Our Model III, constrained by various observations of galaxies, predicts a new low-mass sequence in the $M_{\rm BH}$ - $M_\star$ relation, where the SMBH mass is roughly proportional to galaxy mass, but with the amplitude ~ 50 times lower than that of the high-mass sequence.

The assumptions in and consequences of our proposed scenario can be tested with more detailed modeling. At the moment, it is still unclear whether SMBH in low-mass and late-type galaxies form in a similar way as in their more massive counterparts. For example, it is unclear whether they grow from seeds similar to that for massive SMBH, and whether their growth is also dominated by quasar modes, although they are predicted to be produced by mergers of gas-rich galaxies. To answer these questions, observations of low-luminosity AGNs at high z and accurate determinations of the SMBH masses in low-mass galaxies are essential. If our scenario is correct, we expect to observe a large number of low-luminosity AGNs at $z > 2$ and a large number of SMBH with masses in the range from 10^5 to $10^6\,{\rm M}_\odot$ at $z \sim 0$. The classical bulges within which these SMBH have formed are expected to be compact, not only because their host halos at $z > 2$ are small, but also because major galaxy-galaxy mergers can reduce the angular momentum of the cold disk gas, making the merger remnants even smaller. How such formation is related to the observed ultra compact dwarfs (UCD, e.g. Norris et al., 2015, and references therein) is clearly an interesting and open question.

In our scenario an implicit assumption is that secular evolution of galaxy disks does not play a major role in the growth of SMBH. This assumption can be tested by observing barred disks, in which secular evolution is on-going, to see whether their AGN activities are elevated or not. So far observational evidence is negative for such elevation (e.g. Kormendy et al., 2011), but current data are still sparse.

We have also shown that the formation of SMBH in low-mass galaxies provides a new scenario of preventative feedback, in which gas is prevented from being accreted into

a dark matter halo due to preheating. Such a scenario is plausible based on our simple arguments, but the details need to be worked out. For example, how does the AGN-driven outflow propagates into the IGM and affect its properties? What is the structure of the circum-galactic medium (CGM) produced in this way? Will such formation produce a multiphase medium (Mo and Miralda-Escude, 1996) and how will it be observed in QSO absorption line studies? Can the interaction between the gas and dark matter reduce the phase space density of dark matter halos, as advocated in Mo and Mao (2004)? We will come back to some of these questions in the future.

CHAPTER 6

GALAXY ECOSYSTEMS: GAS CONTENTS, INFLOWS AND OUTFLOWS

6.1 Introduction

The processes that can affect star formation are generally divided into three categories: gas inflow, outflow, and star formation laws. The low star formation efficiency can either be caused by a reduced gas inflow, a strong gas outflow driven by some feedback processes, or the distribution, thermal and chemical states of the cold gas disk, but how the processes work in detail is still unclear. For high-mass galaxies, the quenching of star formation is believed to stem from the suppression of gas inflow into the galaxies by processes, such as AGN heating (e.g. Croton et al., 2006), that can heat the gas supply.

For less massive galaxies, one popular scenario is strong gas outflow driven by supernova (SN) explosions and radiation pressure from massive stars (e.g. Dekel and Silk, 1986; Oppenheimer and Davé, 2008). Because of the relatively shallow potential wells associated with low-mass galaxies, outflows may drive gas out of their host halos, reducing the gas supply for star formation. Preventative scenarios have also been proposed for low-mass galaxies. For example, using a simple analytic model and observational constraints, Bouché et al. (2010) suggests that gas accretion must be suppressed if the halo mass is $< 10^{11}\,\mathrm{M}_{\odot}$. The physical mechanism is uncertain. Heating by the UV background (Ikeuchi, 1986; Rees, 1986). is found to be effective only in halos with masses below $\sim 10^{10}\,h^{-1}\mathrm{M}_{\odot}$ (e.g. Gnedin, 2000). For halos with mass $\sim 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ other heating sources have been suggested, such as gravitational pancaking (Mo et al., 2005), blazar heating (Chang et al., 2012), and galactic winds (e.g. Mo and Mao, 2002; Mo and Mao, 2004; van de Voort et al., 2011).

It is also possible that the low star formation efficiency is caused by a low efficiency of converting cold gas into stars. Indeed, as shown by Krumholz and Dekel (2012) using the metallicity-regualted star formation model developed in Krumholz et al. (2008) and Krumholz et al. (2009), the star formation in very low-mass galaxies can be completely shut off owing to their low metallicities.

To distinguish between the different scenarios, one ultimately needs direct observational constraints on inflow, outflow and the gas distribution in high redshift galaxies. In the absence of such direct observational data at the moment, observational measurements such as the metallicity of the interstellar medium (Tremonti et al., 2004; Erb et al., 2006; Kewley and Ellison, 2008), have been used to constrain gas flows and star formation in galaxies (Dalcanton, 2007; Erb, 2008; Peeples and Shankar, 2011; Lilly et al., 2013; Zahid et al., 2014). For instance, using a simple analytic chemical evolution model, Erb (2008) constrained the outflow and argued that the mass loading factor of the outflow should be about unity in order to match the mass-metallicity relation. Using a more sophisticated chemical evolution model that takes into account inflow/outflow of the gas and star formation in the ISM, Lilly et al. (2013) were able to infer, from a set of simple but plausible assumptions, how the mass loading factor of the outflow depends on the mass of host galaxies

In this chapter, we construct a model for the galactic ecosystem, which includes the gas content (both atomic and molecular), inflow and outflow of gas and the enrichment of metals. Using current observational data on the gas mass fraction of local galaxies (Peeples and Shankar, 2011; Papastergis et al., 2012) and the evolution of gas phase metallicity-stellar mass relation (Maiolino et al., 2008), together with the empirically constrained SFR-halo mass relations (§2), we infer how gas inflow and outflow regulate star formation. This chapter is organized as follows. The basic equations that govern the evolution of different components of a galaxy are described in §6.2. The observational constraints adopted are presented in §6.3. In §6.4, we use two optional star formation

laws (models) to infer the cold gas mass for galaxies with different masses and at different redshifts. In §6.5, we constrain the mass and metal exchange between galaxies and their environments to shed light on gas inflow and outflow, using information about the major components of galaxies, such as dark halos, stars, cold gas and metals. Finally we summarize our conclusions and discuss their implications for physically motivated galaxy evolution models in §6.6.

Unless stated otherwise, we adopt the stellar population synthesis model of Bruzual and Charlot (2003) and a Chabrier IMF Chabrier (2003). Solar metallicity is defined as $Z_{\odot} = 0.0142$ in terms of the total metal mass fraction, and as $Z_{\odot,\mathrm{O}} = 0.0056$ in terms of the oxygen mass fraction (Asplund et al., 2009).

6.2 Galaxy Ecology

6.2.1 The basic equations

The ecosystem of a galaxy consists of stellar mass ($M_{\star}$), cold gas mass (M_{g}), and metal mass (M_{Z}). The evolution of these components in a galaxy can be described by the following set of equations:

$$\frac{\mathrm{d}M_{\star}}{\mathrm{d}t} = (1-R)\Psi(M_{\mathrm{g}},\, R_{\mathrm{g}},\, Z)\,; \tag{6.1}$$

$$\frac{\mathrm{d}M_{\mathrm{g}}}{\mathrm{d}t} = \epsilon_{\mathrm{acc}} f_{\mathrm{b}} \dot{M}_{\mathrm{h}} - \dot{M}_{\mathrm{w}} + \dot{M}_{\mathrm{r}} - (1-R)\Psi\,; \tag{6.2}$$

$$\begin{aligned}\frac{\mathrm{d}M_{\mathrm{Z}}}{\mathrm{d}t} = {} & \epsilon_{\mathrm{acc}} f_{\mathrm{b}} \dot{M}_{\mathrm{h}} Z_{\mathrm{IGM}} \\ & - \dot{M}_{\mathrm{w}} Z_{\mathrm{w}} + \dot{M}_{\mathrm{r}} Z_{\mathrm{r}} \\ & - (1-R)\Psi Z + y\Psi\,.\end{aligned} \tag{6.3}$$

In this set, Eq. (6.1) specifies the change in stellar mass, with Ψ being the SFR and R being the return mass fraction of evolved stars. Here we make instanteneous recycling approximation, which is reasonable in the redshift range we consider ($0 \leq z \lesssim 2$) because the lifetime of the stars that contribute most of the recycling is short compared to the

Hubble time. The SFR depends on the mass ($M_{\rm g}$), distribution (characterized by the size of the cold gas disk, $R_{\rm g}$), metallicity (Z), and perhaps other properties of the interstellar medium (ISM), as specified by a star formation law.

Eq. (6.2) describes the evolution in cold gas mass. The first term on the right hand side is the inflow rate of pristine gas, written in terms of mass accretion rate of the host dark halo, $\dot{M}_{\rm h}$ (see §6.2.2) multiplied by the universal baryon mass fraction, and a gas accretion efficiency, $\epsilon_{\rm acc}$. In normal circumstances, the efficiency factor $\epsilon_{\rm acc} \leq 1$, and its value may depend on halo mass and redshift. This efficiency may be affected by a variety of physical processes. For instance, if a halo is embedded in a preheated gas, the accretion into the halo may be reduced, making $\epsilon_{\rm acc} < 1$ (Lu and Mo, 2007). It is also possible that the halo can accrete gas at a rate of $f_{\rm b}\dot{M}_{\rm h}$, but that certain heating sources such as "radio-mode" AGN feedback in massive halos (Croton et al., 2006) or photoionization heating by local sources (Cantalupo, 2010; Kannan et al., 2014), can prevent the coronal gas from cooling, making $\epsilon_{\rm acc} < 1$. $\dot{M}_{\rm w}$ on the right hand of Eq. (6.2) is gas outflow, and $\dot{M}_{\rm r}$ is the re-accretion rate of the gas mass that has been ejected at earlier times. Finally, the last term on the right hand side is the cold gas consumption rate of star formation.

Eq. (6.3) describes the chemical evolution. y is the intrinsic metal yield from stars. The metal yield is assumed to be instanteneous, which is a good approximation, because we only consider oxygen produced by short-lived massive stars (§6.2.3). $Z_{\rm IGM}$, Z, $Z_{\rm w}$, and $Z_{\rm r}$ are the metallicities of the intergalactic medium (IGM), ISM, wind, and the re-accreted material, respectively. Note that we distinguish between the accretion of the pristine gas from the IGM and the re-accretion of the recycled wind material from the galaxy. In general, the metallicity of the wind and recycled material is not necessarily equal to that of the ISM; for example supernova ejecta and stellar wind may directly carry away metals (Mac Low and Ferrara, 1999), giving $Z_{\rm w} \geq Z$.

Note that this set of equations is only valid for galaxies with no satellites of comparable masses. Otherwise, the central galaxies may obtain a significant amount of metals by

Table 6.1. The oxygen yield as a function of initial stellar metallicity. Results obtained from two different stellar evolution models are presented: P98 is for Portinari et al. (1998) and K06 for Kobayashi et al. (2006). Chabrier IMF (Chabrier, 2003) is used in both models.

Model	$Z_{\rm i} = 0.0004$	$Z_{\rm i} = 0.004$	$Z_{\rm i} = 0.02$
P98	0.0168	0.0180	0.0163
K06	0.0134	0.0110	0.0103

accreting the enriched hot gas of the satellites after halo merger and the ISM once galaxy mergers occur. In this paper we focus only on galaxies with their halo mass in the range $10^{11}\,{\rm M}_{\odot}$ to $2 \times 10^{12}\,{\rm M}_{\odot}$. In this range the mass of the satellites are typical much smaller than the centrals and major galaxy-galaxy mergers are negligible (§ 4.1). For more massive galaxies, Eqs 6.2 and 6.3 are not sufficient unless the gas and metals brought in by mergers are properly taken into account. Another reason for not to extending to higher halo mass is the limitation of the observational constraints we use. The star formation history (§6.3.1) of more massive galaxies, most of which are quenched, is an average over star formation and quenched galaxies, while the gas phase metallicity (§6.3.2) are limited to star forming galaxies.

6.2.2 The Halo Assembly History

Our empirical model follows galaxy evolution in the context of realistic halo assembly histories. The assembly of individual dark matter halos is modeled using the halo merger tree generator proposed by Parkinson et al. (2008). Given a halo mass and a redshift, we only follow the average assembly history instead of individual merger trees. The average assembly history is obtained by averaging over the main-branch progenitors of different trees. The intrinsic scatter in both the models and the observational constraints, such as the gas phase metallicity-stellar mass relation and the gas mass fraction-stellar mass relations, is ignored for simplicity.

6.2.3 The Intrinsic Metal Yield

The intrinsic yield of a simple stellar population can be estimated from the stellar IMF and the adopted stellar evolution model:

$$y = \int_{m_u}^{m_l} mp(m)\phi(m)\mathrm{d}m\,, \tag{6.4}$$

where $\phi(m)$ is the IMF and $p(m)$ is the mass fraction of certain metals produced by stars of an initial mass m. Here we adopt two models, one is from Portinari et al. (1998) and the other is from Kobayashi et al. (2006). Table 6.1 lists the yield of oxygen for different initial metallicities of the stellar population. For both models the oxygen yield depends mildly on the initial stellar metallicity. However, the variance between different models is considerable. The yield predicted by the Kobayashi et al. (2006) model is about 2/3 of that by the Portinari et al. (1998) model. In this paper, we choose the Portinari et al. (1998) model as our fiducial model, since it is consistent with a broad range of stellar evolutions models in the literature (Peeples et al., 2014). The consequence of using a smaller yield will be discussed whenever needed.

6.3 Observational Constraints

6.3.1 Star Formation History

The SFR - halo mass relation of Model III is adopted as the constraints. The SFR as a function of halo mass and redshift and the consequent stellar mass to halo mass ratio are shown in Figure 6.1 for reference, with the bands representing the inferential uncertainty. In the mass range we are interested in here (between the two vertical grey lines), this uncertainty is quite small. We therefore ignore the scatter and use the best fit parameters to characterize the star formation histories.

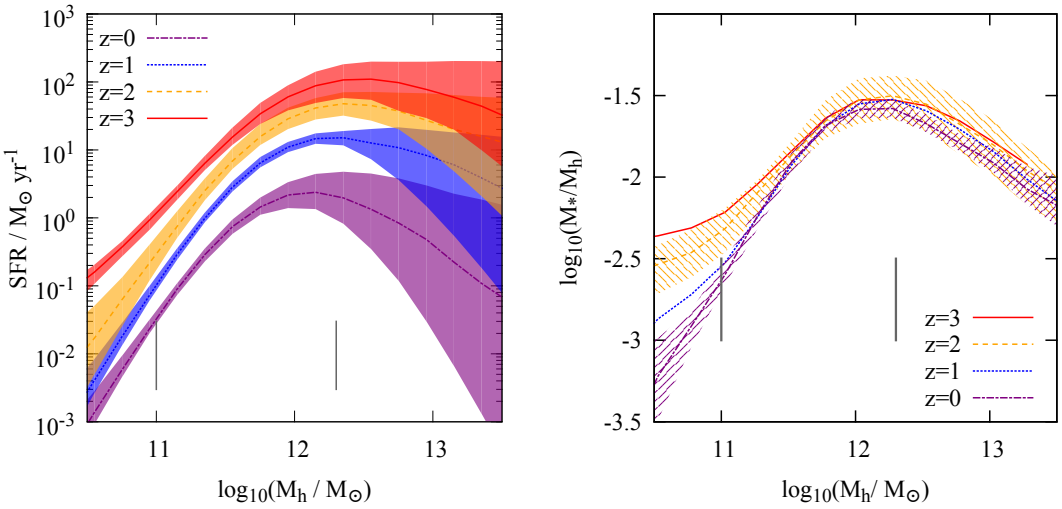

Figure 6.1. The dependence of SFR (left) and stellar mass to halo mass ratio (right) on halo mass and redshift. The bands represent the 95% credible intervals. The grey vertical lines bracket the halo mass range we focus on in this work.

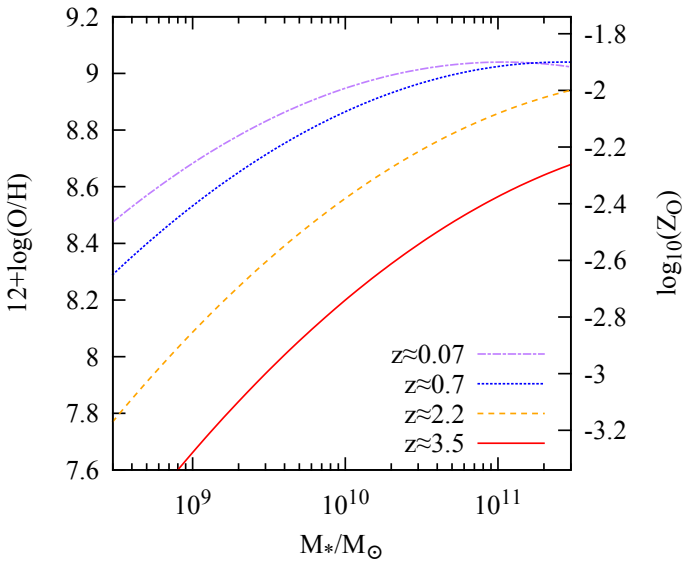

Figure 6.2. Gas phase metallicity compiled by Maiolino et al. (2008). The y-axis on the right is the corresponding oxygen mass fraction.

6.3.2 Gas Phase Metallicity

Another observational constraint is the gas phase metallicity, which is usually measured from the emission lines of the HII regions of star forming galaxies. We adopt the metallicity measurements compiled by Maiolino et al. (2008). Figure 6.2 shows the metallicity-stellar mass relations obtained from their fitting formula. It is important to realize that the metallicity measurements have significant systematic error. For local galaxies the random error in the measurements is only about 0.03 dex (Tremonti et al., 2004), but the systematic uncertainty due to different ways to convert the emission lines into abundances is as large as 0.7 dex (Kewley and Ellison, 2008). There are two ways to estimate the metal abundance from such observations: the electron temperature ($T_{\rm e}$) method and the theoretical method. In the T_e method, the ratio between the [OIII]λ4363 auroral line and [OIII]λ5007 is used to estimate the mean electron temperature, which is in turn used to estimate the oxygen abundance (Peimbert and Costero, 1969). In the theoretical method, a sophisticated photoionization model is fit to the strong line ratios, such as $R_{23} = ([\mathrm{OII}]\lambda 3737 + [\mathrm{OIII}]\lambda 4959, 5007)/\mathrm{H}\beta$. Empirical calibrations based on the two methods often show a discrepancy as large as 0.7 dex. Stasińska (2005) pointed out that due to the temperature fluctuation or gradient in high metallicity [$12 + \log_{10}(\mathrm{O/H}) > 8.6$] HII regions, the T_e method can underestimate the metallicity by as much as 0.4 dex. Meanwhile the systematics in the photoionization modeling can be as large as 0.2 dex (Kewley and Ellison, 2008). In Maiolino et al. (2008), both of the two methods described above are used to derive the relations between the strong line ratios and metallicity. Specifically, the T_e method is only applied to metal poor galaxies ($12 + \log_{10}(\mathrm{O/H}) < 8.6$) to avoid bias. The empirical calibrations derived in this way cover a large metallicity range and therefore can be applied to galaxies over a large redshift range.

6.3.3 Gas Fraction in Local Galaxies

In addition, we also include the observations of gas contents in local galaxies compiled by Peeples and Shankar (2011) as a constraint. The data points in Figure 6.3, which show

Table 6.2. Parameters in the Kennicutt-Schmidt model, $A_{\rm K}$ and $\Sigma_{\rm c}$, and in the Krumholz model, $\tau_{\rm sf}$ and c, tuned together with the disk-size parameter, $\mathcal{L}$, to match the gas mass to stellar mass ratio of local galaxies (data points in Figure 6.3). The second column lists the fitting results and the third column lists the equations which define the corresponding parameters.

parameter	value	
$A_{\rm K}$ $/{\rm M}_\odot{\rm yr}^{-1}{\rm pc}^{-2}$	$(2.5 \pm 1.2) \times 10^{-4}$	Eqs. (6.5) - (6.6)
$\Sigma_{\rm c}$ $/{\rm M}_\odot{\rm pc}^{-2}$	9.8 ± 2.6	
$\mathcal{L}$	2.1 ± 0.4	Eq (6.12)
$\tau_{\rm sf}/{\rm Gyr}$	2.5 ± 0.3	Eqs. (6.7) - (6.9)
c	2.6 ± 1.5	
$\mathcal{L}$	3.3 ± 1.0	Eq. (6.12)

the total gas mass to stellar mass ratios, are taken from Peeples and Shankar (2011). The binned data points are compiled from several different sources, taken into account HI, helium and molecular hydrogen. Here both the mean relation and the uncertainties, taken as random errors, are used in the data constraint.

6.4 Evolution of cold gas content of galaxies

Given the observational constraints for the star formation histories in §6.3.1 and for the gas phase metallicity in §6.3.2, we can solve Eq. (6.1) to obtain the gas mass $M_{\rm g}$ by adopting specific models for the SFR and for the structure of the cold gas distribution. In this section, we first introduce the star formation (§6.4.1) and disk structure (§6.4.2) models we adopt, we then show the predictions for the cold gas mass in high redshift galaxies (§6.4.3).

6.4.1 The Star Formation Models

We consider two different star formation models widely adopted in the literature. The first is the Kennicutt-Schmidt Law (Kennicutt, 1998), an empirical relation between the SFR surface density, $\Sigma_{\rm SFR}$, and the cold gas surface density, $\Sigma_{\rm g}$,

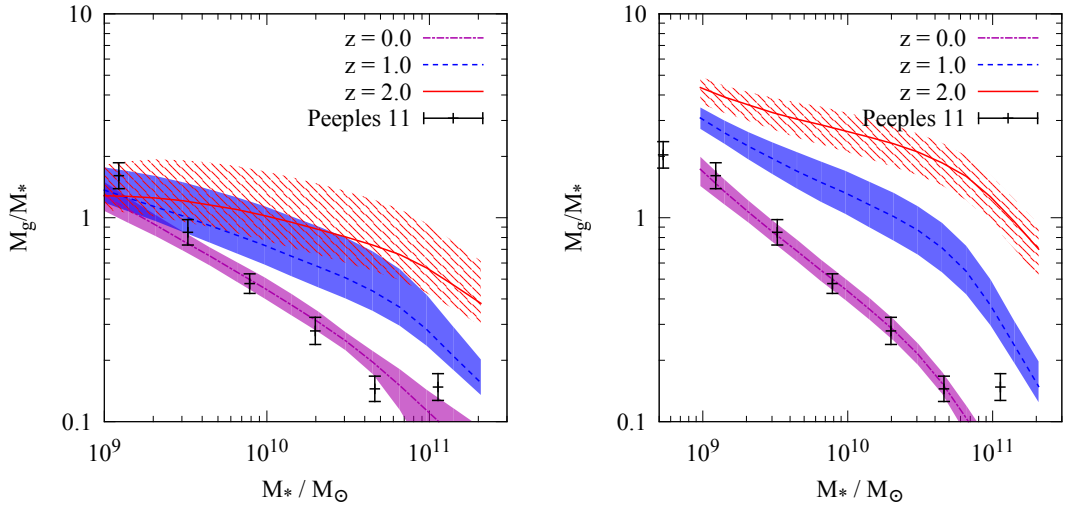

Figure 6.3. Cold gas to stellar mass ratio as a function of stellar mass at different redshifts calculated using the Kennicutt-Schmidt Law (left) and the Krumholz model (right). The lines are the predictions of the best fitting model in Table 6.2 and the bands are obtained by marginalizing the uncertainties in the parameters. The data points are compilation of Peeples and Shankar (2011) from different observations of local galaxies.

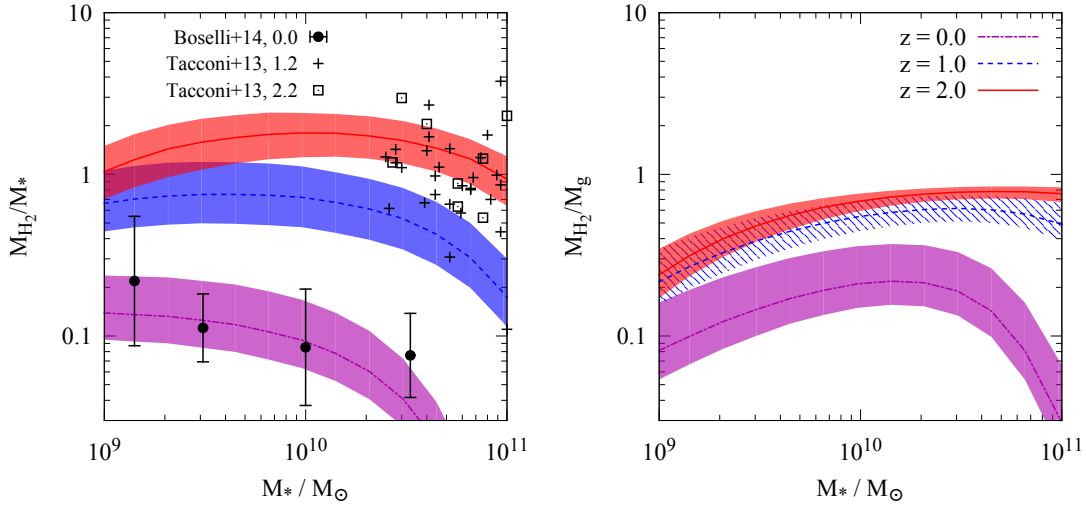

Figure 6.4. Molecular gas to stellar mass ratio (left), molecular gas to total gas mass ratio (right) as a function of stellar mass, all predicted by the Krumholz model. The lines are the predictions of the best fitting model in Table 6.2 and the bands are obtained by marginalizing the uncertainties in the parameters. The data points from Tacconi et al. (2013) are individual galaxies the binned data points for local galaxies are from Boselli et al. (2014).

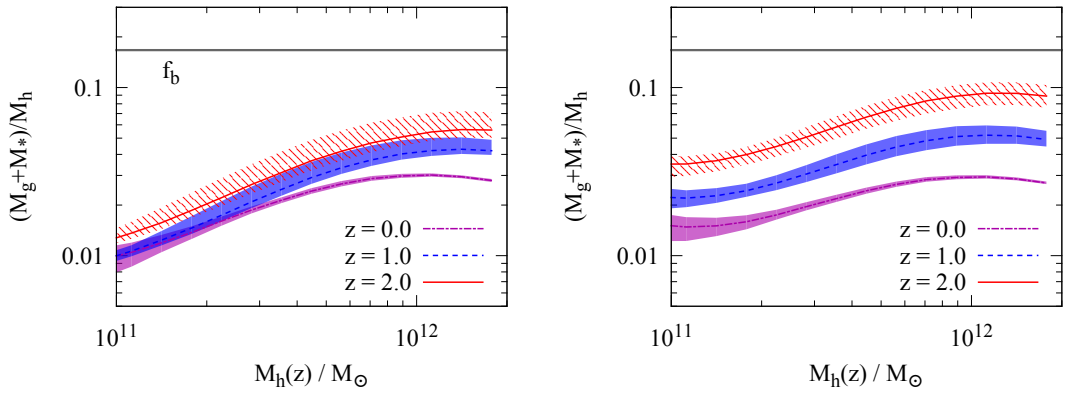

Figure 6.5. Baryon mass (in both stars and cold gas) to halo mass ratio as a function of halo mass calculated using the Kennicutt-Schmidt Law (left) and the Krumholz Law (right). The horizontal lines indicate the universal baryon fraction. The lines are the predictions of the best fitting model in Table 6.2 and the bands are obtained by marginalizing the uncertainties in the parameters.

$$\Sigma_{\rm SFR} = A_{\rm K} \left(\frac{\Sigma_{\rm g}}{\rm M_\odot pc^{-2}} \right)^{N_{\rm K}} , \tag{6.5}$$

where the power index $N_{\rm K} \approx 1.4$, and $A_{\rm K}$ is a constant amplitude. In this model, star formation is assumed to occur only in cold gas disks where the surface density exceeds a threshold $\Sigma_{\rm c}$. Assuming the cold gas disk follows an exponential profile, the total SFR can be obtained as,

$$\Psi = \begin{cases} \frac{2\pi A_{\rm K} \Sigma^{N_{\rm K}} R_{\rm g}^2}{N_{\rm K}^2} \left[1 - \left(1 + N_{\rm K} \frac{r_{\rm c}}{R_{\rm g}} \right) \exp\left(-N_{\rm K} \frac{r_{\rm c}}{R_{\rm g}} \right) \right} \\ \qquad \text{if } \Sigma_0 \geq \Sigma_{\rm c} \\ 0 \\ \qquad \text{if } \Sigma_0 < \Sigma_{\rm c} \, , \end{cases} \tag{6.6}$$

where $R_{\rm g}$ is the scale radius of the disk, $\Sigma_0 \equiv \frac{M_{\rm g}}{2\pi R_{\rm g}^2}$ is the surface density at the disk center, and $r_{\rm c} = \ln(\Sigma_0/\Sigma_{\rm c}) R_{\rm g}$ is the critical radius, within which star formation can happen. Both $A_{\rm K}$ and $\Sigma_{\rm c}$ are treated as free parameters to be determined by observational constraints.

The other star formation model adopted here is the one proposed by Krumholz et al. (2008) and Krumholz et al. (2009), in which the SFR is assumed to be directly related to the properties of the molecular cloud:

$$\Sigma_{\rm SFR} = \frac{\epsilon_{\rm ff}}{t_{\rm ff}} \Sigma_{\rm H_2} \,, \tag{6.7}$$

where $\Sigma_{\rm H_2}$ is the surface density of molecular hydrogen, and $t_{\rm ff}$ is the local free fall time scale. The ratio $\epsilon_{\rm ff}/t_{\rm ff}$ depends on the total gas surface density:

$$\frac{\epsilon_{\rm ff}}{t_{\rm ff}} = \frac{1}{\tau_{\rm sf}} \begin{cases} \left(\Sigma_{\rm g}/85 {\rm M}_\odot {\rm pc}^{-2}\right)^{-0.33} & \text{if } \Sigma_{\rm g} < 85 {\rm M}_\odot {\rm pc}^{-2} \\ \left(\Sigma_{\rm g}/85 {\rm M}_\odot {\rm pc}^{-2}\right)^{0.33} & \text{if } \Sigma_{\rm g} \geq 85 {\rm M}_\odot {\rm pc}^{-2} \,, \end{cases} \tag{6.8}$$

where $\tau_{\rm sf}$ is a constant, treated as a free parameter. The fraction of molecular gas, $f_{\rm H_2} = \Sigma_{\rm H_2}/\Sigma_{\rm g}$, depends primarily on the surface density and metallicity of the cold gas, and is modeled as

$$f_{\rm H_2} = \begin{cases} 1 - \frac{3}{4}\left(\frac{s}{1+0.25s}\right) & \text{if } s \leq 2 \\ 0 & \text{if } s > 2 \end{cases} \tag{6.9}$$

$$s = \frac{\ln(1 + 0.6\chi + 0.01\chi^2)}{0.6\tau_{\rm c}}$$

$$\chi = 3.1 \frac{1 + Z_o^{0.365}}{4.1}$$

$$\tau_{\rm c} = 320 c Z_o \frac{\Sigma_{\rm g}}{{\rm g/cm^2}} \,.$$

Here Z_o is the metallicity normalized to the solar value, c is a constant treated as a free parameter, and $s = 2$ defines a threshold surface density for the formation of molecular hydrogen. Note that s is roughly inversely proportional to the gas phase metallicity, so that a high metallicity corresponds to a lower surface density threshold.

6.4.2 Disk Size

To determine the distribution of cold gas, we assume that the cold gas disk follows an exponential radial profile with a disk size proportional to the stellar disk. We estimate the stellar disk size using the empirical size-stellar mass relation obtained by Dutton et al. (2011) for nearby galaxies ($z \approx 0.1$),

$$R_{50} = R_0 \left(\frac{M_\star}{M_0}\right)^\alpha \left[\frac{1}{2} + \frac{1}{2}\left(\frac{M_\star}{M_0}\right)^\gamma\right]^{(\beta-\alpha)/\gamma}, \tag{6.10}$$

where R_{50} is the half light radius of the stellar disk, $\log_{10}(M_0/M_\odot) = 10.44$, $\log_{10}(R_0/\mathrm{kpc}) = 0.72$, $\alpha = 0.18$, $\beta = 0.52$ and $\gamma = 1.8$. With the assumption that the shape of the relation holds at all redshifts, the time evolution of the disk size is given by the offset

$$\Delta \log_{10}(R_{50}) = 0.018 - 0.44 \log_{10}(1+z). \tag{6.11}$$

This redshift dependence is slightly shallower than the more recent observational calibration by van der Wel et al. (2014), which is $\propto (1+z)^{-0.75}$, but our results are not sensitive to it. As shown by Dutton et al. (2011) the star formation activity typically has a more extended distribution than the stellar disk, with a size about two times the stellar disk, and the relation does not evolve strongly with time. The gas disk is traced by the star formation to some extent. Using a sample of local galaxies that covers a broad range of stellar mass and morphological types, Kravtsov (2013) showed that the sizes of the cold gas disks are typically larger than the stellar disks by a factor of ≈ 2.6. Investigating a semi-analytic model that implements detailed treatments of gas distribution and star formation, Lu et al. (2015a) found that the size ratio between the cold gas disk and the stellar disk ranges from 2 to 3 for galaxies with mass in the range considered here. In our model we therefore assume that

$$R_\mathrm{g} = \mathcal{L} R_\star \,, \tag{6.12}$$

with $\mathcal{L}$ treated as a free parameter to be tuned along with some other parameters in the star formation models (Table 6.2) to match the observed gas fraction of local galaxies.

6.4.3 The Cold Gas Contents

To make use of the models described above, we first calibrate the parameters in the star formation laws and the gas disk size parameter $\mathcal{L}$ using the observed gas mass/stellar mass ratio of local galaxies (Peeples and Shankar, 2011). The best fits and the 1 σ uncertainties of the tuned parameters are listed in Table 6.2. The predicted cold gas contents as functions of stellar mass are shown in Figures 6.3 and 6.4.

Both of the star formation laws can successfully reproduce the cold gas fraction of local galaxies by tuning the corresponding model parameters. This is in contrast with the finding of Peeples and Shankar (2011) that the Schmidt-Kennicut law fails to match the high gas mass fraction in dwarf galaxies. We find that the critical surface density $\Sigma_{\rm c}$, which was not taken into account in Peeples and Shankar (2011), is crucial in reproducing the steep gas mass fraction-stellar mass relation. Similarly, the Krumholz star formation model also has a critical surface density for molecule formation, which is roughly inversely proportional to the gas phase metallicity [Eq. (6.9)]. The key difference between the two star formation models is that the critical surface density in the Schmidt-Kennicutt law is a constant, while that in the Krumholz model changes with time and the mass of the host galaxies. According to the observed gas phase metallicity (Figure 6.2), the critical surface density in Krumholz model increases with redshift. To sustain the same amount of star formation, the gas fraction derived from this model is thus higher than that derived from the Schmidt-Kennicutt law, especially for dwarf galaxies with stellar masses $< 10^9\,{\rm M}_\odot$. Using a molecule-regulated star formation model, Dutton et al. (2010) inferred that the gas to stellar mass ratio changes only weakly with time, in contrast to our results shown in Figure 6.3. The major reason for the difference is that in their model the formation of molecular hydrogen is determined by the total gas surface density, while in the Krumholz model the evolution of metallicity plays a crucial role. Clearly, the gas mass in high

redshift galaxies is sensitive to the assumed star formation model, and more models need to be explored and checked with future observations (Popping et al., 2012; Popping et al., 2014).

The Krumholz model also allows us to infer the gas fraction in molecular phase. The left panel in Figure 6.4 shows the molecular gas to stellar mass ratio. At $z = 0$ the ratio is about 0.1, and it increases by an order of magnitude at $z = 2$. The predictions are consistent with the recent measurements from Boselli et al. (2014) and Tacconi et al. (2013). The right panel shows the molecular gas to total gas mass ratio as a function of stellar mass. At $z > 1$, most of the gas is in the molecular phase.

Regardless which star formation model is adopted, the ratio between the total baryon mass settled in the galaxies and the host halo mass is always much less than the universal baryon mass fraction (see Figure 6.5). This deficit of baryon mass strongly indicates that star formation models alone cannot account for the low star formation efficiency in low-mass halos. Processes that control the gas exchange between the surrounding medium and galactic medium in forms of gas inflow and outflow must have played a major role. In the following section, we infer limits on the inflow and outflow rates in low-mass galaxies from the constrained star formation histories, cold gas fractions, and metallicity measurements.

6.5 Inflow and Outflow

As described above, the main components of galaxies, such as halo mass $M_{\rm h}$, stellar mass $M_\star$, gas mass $M_{\rm g}$ and mass in gas phase metals $M_{\rm Z} \equiv M_{\rm g}Z$, and their time derivatives can either be obtained directly from observational constraints (§6.3) or from modeling (§6.4). In this section, we go a step forward by constraining the terms pertaining to inflow and outflow in Eqs. (6.2) and (6.3). As we will see below, these terms cannot be completely determined, but stringent limits can be obtained for them.

To proceed we rewrite Eqs. (6.2) and (6.3) in more transparent forms. Since the metallicity of the IGM is expected to be much lower than that of the ISM, we set $Z_{\rm IGM} = 0$

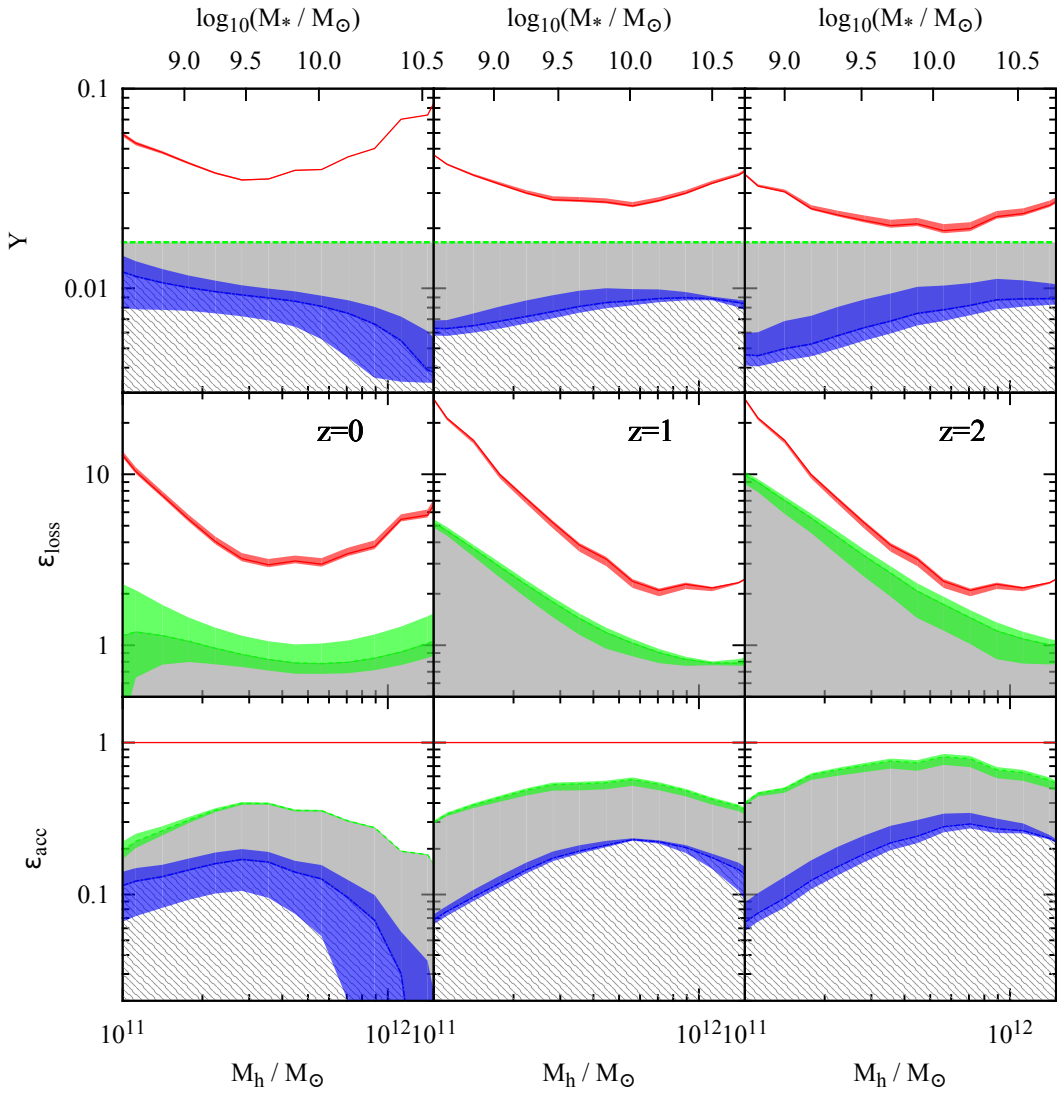

Figure 6.6. The net yield $\mathcal{Y}$, loading factor of gas mass loss rate $\epsilon_{\rm loss}$ and gas accretion efficiency $\epsilon_{\rm acc}$. The red lines (bands) are obtained by assuming $\epsilon_{\rm acc} = 1$, i.e. galaxies accrete at the maximum rate. The green lines (bands) are obtained by assuming $\mathcal{Y} = y$, which means full mixing of newly produced metals in the ISM, and no recycling or instanteneous recycling of the ejected material. The blue lines (bands) are obtained by setting $\epsilon_{\rm loss} = 0$. The areas that are not shaded are forbidden by the observational constraints. Here the Kennicutt-Schmidt law is assumed. The lines are the predictions of the best fitting model in Table 6.2 and the bands are obtained by marginalizing the uncertainties in the parameters. The hatched regions correspond to constraints when $\epsilon_{\rm loss}$ is allowed to be negative (see text).

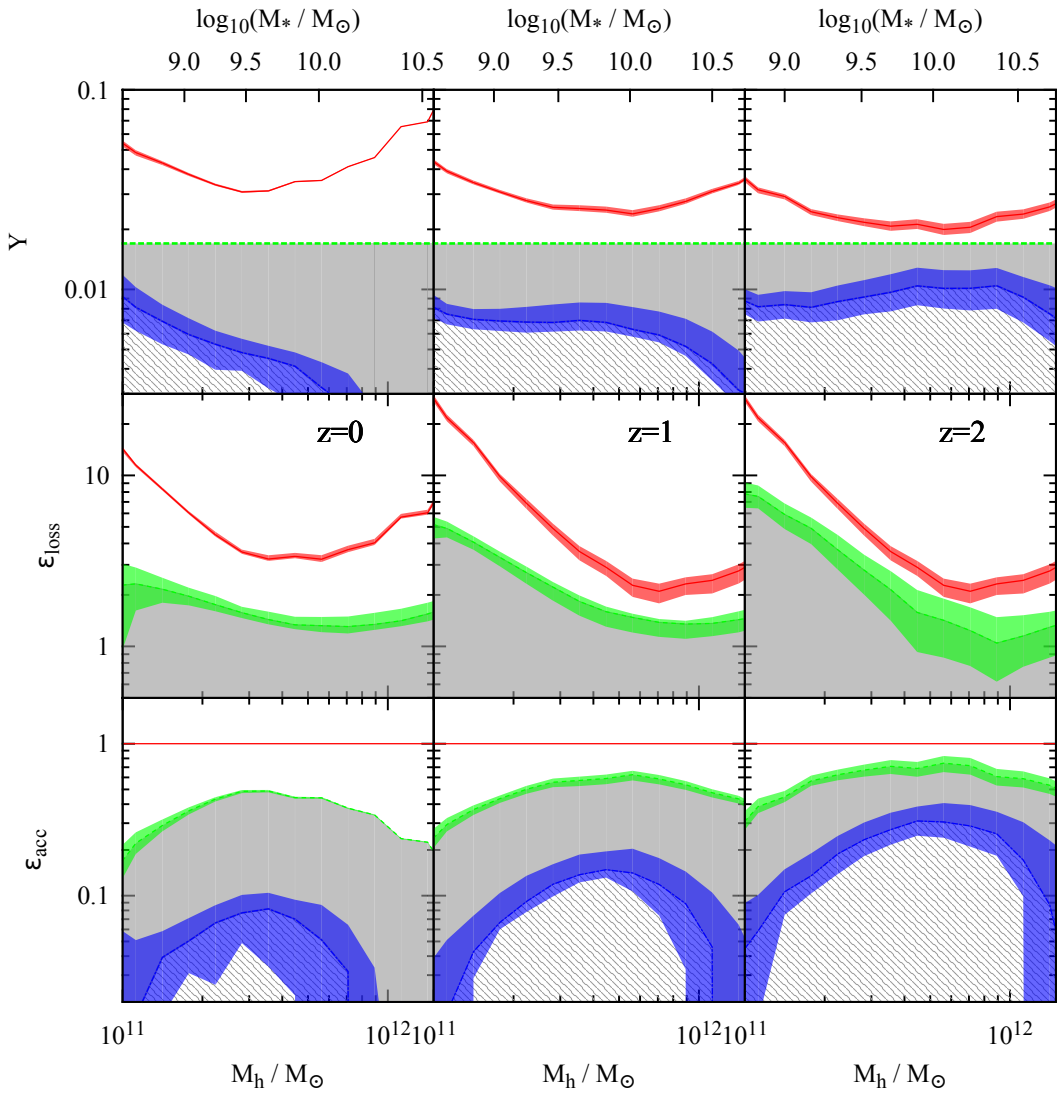

Figure 6.7. The same as Figure 6.6 but here the Krumholz star formation model is assumed.

for simplicity. The gas and chemical evolution equations are then reduced to

$$\frac{\mathrm{d}M_{\rm g}}{\mathrm{d}t} = \epsilon_{\rm acc} f_{\rm b} \dot{M}_{\rm h} - \epsilon_{\rm loss} \Psi - (1-R)\Psi \,; \tag{6.13}$$

$$\frac{\mathrm{d}M_{\rm Z}}{\mathrm{d}t} = -\epsilon_{\rm loss,Z} \Psi - (1-R)\Psi Z + y\Psi \,, \tag{6.14}$$

where

$$\begin{aligned} \epsilon_{\rm loss} &\equiv \epsilon_{\rm w} - \epsilon_{\rm r} \\ &\equiv \frac{\dot{M}_{\rm w}}{\Psi} - \frac{\dot{M}_{\rm r}}{\Psi} \end{aligned} \tag{6.15}$$

is the loading factor of net mass loss, and

$$\epsilon_{{\rm loss},Z} \equiv \epsilon_{\rm w} Z_{\rm w} - \epsilon_{\rm r} Z_{\rm r} \tag{6.16}$$

is the loading factor of net metal loss. With some combinations and re-arrangements, Eqs. (6.13) and (6.14) can be written as

$$1 - R + \epsilon_{\rm loss} = \epsilon_{\rm acc} \mathcal{E}_{\rm SF}^{-1} \left(1 - \epsilon_{\rm acc}^{-1} \frac{\dot{M}_{\rm g}}{f_{\rm b} \dot{M}_{\rm h}} \right) \,; \tag{6.17}$$

$$\frac{\mathcal{Y}}{Z} = \epsilon_{\rm acc} \mathcal{E}_{\rm SF}^{-1} \left(1 + \epsilon_{\rm acc}^{-1} \frac{M_{\rm g}}{f_{\rm b} \dot{M}_{\rm h}} \frac{\dot{Z}}{Z} \right) \,, \tag{6.18}$$

where $\mathcal{E}_{\rm SF} \equiv \Psi / (f_{\rm b} \dot{M}_{\rm h})$ is the star formation efficiency (Eq 3.1), and

$$\mathcal{Y} \equiv y - \epsilon_{\rm w} \left(Z_{\rm w} - Z \right) + \epsilon_{\rm r} \left(Z_{\rm r} - Z \right) \,. \tag{6.19}$$

This quantity can be interpreted as the "net yield". For instance, the second term $\epsilon_{\rm w} \left(Z_{\rm w} - Z \right)$ represents the metals taken away by the galactic wind without being mixed with the ISM.

The above equations are general and are used to make model predictions to be described below. Before presenting the results, let us look at these equations under certain approximations, which will help us to understand the results obtained from the full model and to make connections to results obtained earlier under similar approximations. Since $M_{\rm g}$ is typically much smaller than $f_{\rm b}M_{\rm h}$, as shown in Figure 6.5, $\dot{M}_{\rm g}/(f_{\rm b}\dot{M}_{\rm h})$ in Eq. (6.17) and $[M_{\rm g}/(f_{\rm b}\dot{M}_{\rm h})](\dot{Z}/Z)$ in Eq. (6.18) are expected to be much less than unity. Thus, if $\epsilon_{\rm acc}$ is of the order of unity or $\epsilon_{\rm acc} \gg [M_{\rm g}/(f_{\rm b}\dot{M}_{\rm h})](\dot{Z}/Z)$ and $\epsilon_{\rm acc} \gg \dot{M}_{\rm g}/(f_{\rm b}\dot{M}_{\rm h})$, the above equations can be simplified to

$$1 - R + \epsilon_{\rm loss} \approx \epsilon_{\rm acc}\mathcal{E}_{\rm SF}^{-1}\,; \tag{6.20}$$

$$\frac{\mathcal{Y}}{Z} \approx \epsilon_{\rm acc}\mathcal{E}_{\rm SF}^{-1}\,. \tag{6.21}$$

In this case, the star formation efficiency ($\mathcal{E}_{\rm SF}$) and the chemical evolution is completely determined by the gas exchange between the galaxies and their environment, independent of the gas content of the galaxy. The choice of star formation law is also not important unless it gives a gas mass that is comparable to $f_{\rm b}M_{\rm h}$. This set of equations is basically equivalent to equations (16) and (18) in Davé et al. (2012), which are derived directly from the assumption that gas inflow, outflow and consumption by star formation are in equilibrium. This approximate model was adopted by Henry et al. (2013) to evaluate the plausibility of different wind models (and models with no wind). We caution, however, that this simplified model is not general, and is only valid under the assumptions described above.

6.5.1 Models with strong gas outflow

A commonly adopted assumption in galaxy formation models is that halos accrete baryons at the maximum rate, $f_{\rm b}\dot{M}_{\rm h}$. For halos with mass below $10^{12}\,{\rm M}_{\odot}$, where the radiative cooling timescale is always shorter than halo dynamical time, the gas accre-

tion onto the central galaxy is also expected to follow the halo accretion. We test the consequence of this basic assumption using our constrained model.

Setting $\epsilon_{\rm acc} = 1$, i.e. assuming galaxies are accreting at the maximum rate, we can calculate the net yield $\mathcal{Y}$ and the mass loading factor $\epsilon_{\rm loss}$ using Eqs. (6.17) and (6.18). The results are shown as the red curves in Figure 6.6 for the Kennicutt-Schmidt star formation model and in Figure 6.7 for the Krumholz model, respectively. Although the two star formation models lead to sizable differences in the gas mass, the predicted mass loading factors and net yields are very similar, suggesting that this uncertainty does not strongly affect the estimates of the yield and mass loading factor. The reason for this is that the conditions leading to the approximate model given by Eqs. (6.20) and (6.21) are valid, so that $\mathcal{E}_{\rm SF}$ and $\mathcal{Y}$ are independent of $M_{\rm g}$. In this case the gas exchange between the galaxy and the environments is rapid. For example, the required loading factor for $10^{11}\,{\rm M}_{\odot}$ halos can be as high as 10 to 20.

With the use of the fiducial M-Z relations as constraints, the net yield $\mathcal{Y}$ predicted exceeds the intrinsic yield, y (shown as green horizontal lines in the upper panels of Figures 6.6 and 6.7) at least since $z \approx 2$. The value of $\mathcal{Y}$ defined above is related to a number of factors: (i) the intrinsic yield y; (ii) the value of $Z_{\rm w}$ which is determined by how well the metals produced by stars are mixed with the ISM; and (iii) the value of $Z_{\rm r}$ which is determined by the history of the galaxies. In general, the value of $\mathcal{Y}$ cannot exceed that of y because metals in both inflow and outflow must have been diluted. In large scale cosmological simulations (e.g. Davé et al., 2012) and semi-analytic model of galaxy evolution (e.g. Lu et al., 2014b), metals produced by stars are assumed to be fully mixed with the ISM, so that $Z_{\rm w} = Z$ is expected. Also, the observed metallicity of the ISM generally increases monotonically with time, so that $Z_{\rm r} < Z$. Putting all these together implies $\mathcal{Y} = y + \epsilon_{\rm r}\,(Z_{\rm r} - Z) < y$. On the other hand, in the case of no wind recycling, as assumed in Lilly et al. (2013), $\mathcal{Y} = y - \epsilon_{\rm r}\,(Z_{\rm w} - Z) \leq y$. Generally, as long as the recycled material is less enriched than the wind, $\mathcal{Y}$ should always be no larger than the intrinsic

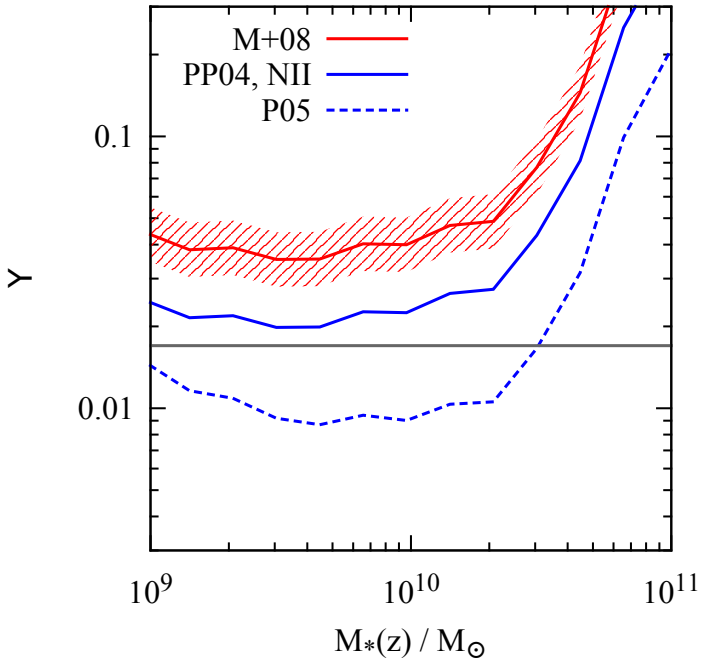

Figure 6.8. The net yield at $z = 0$ calculated using Eq. (6.21) assuming $\epsilon_{\rm acc} = 1$. The red solid line is based on the fiducial M-Z relation from Maiolino et al. (2008) and the band indicates the systematic uncertainty from the photoionization modeling. The blue solid line and the blue dashed line are based on the metallicity measurements using the calibrations from Pettini and Pagel (2004) and Pilyugin and Thuan (2005) respectively, both are based on $T_{\rm e}$ method. The horizontal line shows the intrinsic yield.

yield y. Thus, under the assumption that gas accretion follows the accretion of the host dark halos, the gas outflow would be required to be under enriched in metals than what is to be expected, suggesting that the assumption $\epsilon_{\rm acc} = 1$ is invalid.

As shown in Eq. (6.21), the net yield $\mathcal{Y}$ is roughly proportional to the gas phase metallicity measured Z. This provides a simple way to understand the systematic effects in the measured gas phase metallicity. These effects are carefully analyzed in Kewley and Ellison (2008). The variance between different measurements using the photonionization modeling (the second method briefly described in §6.3) is about 0.2 dex, and the resultant uncertainty in the net yields is shown as the red band shown in Figure 6.8. It is clear that the net yield $\mathcal{Y}$ required is always larger than the intrinsic value y, in conflict with the expectation that $\mathcal{Y} < y$.

We have also used the M-Z relations derived from the $T_{\rm e}$ method (or calibrations based on this method) to estimate the net yield, and the results are shown as the blue lines in Figure 6.8. For HII regions with $12 + \log_{10}({\rm O/H}) > 8.6$ the metallicity derived from this method is systematically lower. In particular, the M-Z relation from Pilyugin and Thuan (2005) is lower by 0.7 dex. The values of $\mathcal{Y}$ so derived are consistent with the intrinsic yield from stellar evolutions models, except at the massive end. Unfortunately, this agreement cannot be taken seriously, because theoretical investigations have demonstrated that the $T_{\rm e}$-based methods tend to underestimate the metallicity in metal rich HII regions (e.g. Stasińska, 2005). What is clear, though, is that accurate measurements of the gas phase metallicity can provide stringent constraints on galactic inflow and outflow.

6.5.2 Constraining gas inflow and outflow

As discussed in the previous subsection, the natural assumption that the net yield ought to be lower than the intrinsic yield requires a reduced rate for gas exchange between galaxy and its ambient medium. If the baryon mass exchange is too rapid via inflow of pristine gas or outflow of metal enriched ISM, the predicted gas phase metallicity would be too low when a reasonable value is assumed for the net yield. What this means is that we can constrain the upper limit for the inflow and outflow efficiencies, $\epsilon_{\rm acc}$ and $\epsilon_{\rm loss}$, by setting $\mathcal{Y}$ to its upper limit, namely setting $\mathcal{Y} = y$. It can be shown that, as long as $Z_{\rm w} \geq Z$ and $Z_{\rm w} \geq Z_{\rm r}$, the relation $\mathcal{Y} = y$ requires both $Z_{\rm w} = Z$ and $Z_{\rm w} = Z_{\rm r}$. As mentioned above, $Z_{\rm w} = Z$ implies that the metals produced from star formation is fully mixed with the ISM. In this case, gas outflow is the least efficient in carrying metals out of galaxies. The second condition, $Z_{\rm w} = Z_{\rm r}$, implies that some of the ejected gas is recycled instantaneously while the rest is permanently lost. The upper limits to $\epsilon_{\rm acc}$ so obtained are shown as the green lines in the lower panels of Figures 6.6 and 6.7, while the upper limits to $\epsilon_{\rm loss}$ are shown as the green lines in the middle panels of the same figures.

The two different star formation models produce similar results. The variation in the gas content does not cause much variation in the estimate of $\epsilon_{\rm acc}$. The low star formation

efficiency in low-mass galaxies is due to strong outflow at $z = 2$, and to inefficient accretion at $z = 0$. At $z = 2$, the mass loading is roughly proportional to $M_{\rm h}^{-1}$ and is about 10 for $10^{11}\,{\rm M}_{\odot}$ halos. At $z = 0$ the mass loading depends only weakly on halo mass, with values close to 1. Both the accretion efficiency, $\epsilon_{\rm acc}$, and the effective wind loading factor, $\epsilon_{\rm loss}$, drop by a factor of 2 from $z = 2$ to $z = 0$. These drops are direct results of the evolution in the observed M-Z relation, as our model assumes full mixing. At high redshift, a large fraction of metals are required to be lost with the ejected ISM in order to reproduce the relatively low metallicity, while at low redshift most of the metals are retained so as to reproduce the increased metallicity.

We also consider another special case in which there is no wind recycling, i.e. $\epsilon_{\rm loss} \to 0$. [1] In this case, the accretion efficiency $\epsilon_{\rm acc}$ is required to be much lower than unity in order to maintain the total amount of cold gas in the disk. In this limit, the approximations given by Eqs. (6.20) and (6.21) are not valid anymore, and the derived $\epsilon_{\rm acc}$ depends on the star formation models adopted. For instance, since the gas fraction at high z predicted by the Krumholz model is systematically higher than the prediction of the Kennicutt-Schmidt model, the required gas accretion at low redshift is much lower, because the star formation at low redshift can be fueled by the gas accumulated earlier in the galaxy. If reincorporation of ejected gas is taken into account, it is possible that $\epsilon_{\rm loss} < 0$. In this case, the corresponding $\epsilon_{\rm acc}$ and $\mathcal{Y}$ will occupy the grey hatched areas shown Figures 6.6 and 6.7.

The boundaries we draw are based on the fiducial M-Z relations and the fiducial intrinsic oxygen yield. As mentioned above, the systematic uncertainty in the metallicity estimate using detailed photoionization modeling is ± 0.1 dex around the mean. Since in the full mixing model, which gives the upper limits of $\epsilon_{\rm acc}$ and $\epsilon_{\rm loss}$, the simple proportionality in Eq. (6.21) holds, a change by ± 0.1 dex in metallicity simply leads to a change of

[1] As shown in the following subsection, outflow of metals is always required to ensure $\mathcal{Y} \leq y$, and so strictly speaking $\epsilon_{\rm loss}$ cannot be exactly zero.

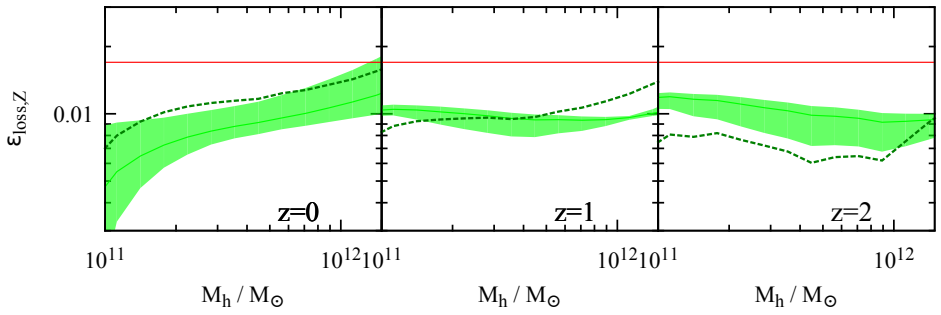

Figure 6.9. The loading factor of metal loss rate. Light green represents the Kennicutt-Schmidt star formation law and dark green represents the Krumholz star formation model. The curves are the predictions of the best fitting model in Table 6.2 and the bands are obtained by marginalizing the uncertainties in the parameters. The red horizontal lines are the intrinsic yield of oxygen.

∓ 0.1 dex in $\epsilon_{\rm acc}$ and to a change of ± 0.1 dex in $\epsilon_{\rm loss}$. However, if the $T_{\rm e}$-based metallicity is used as model constraint, the assumption that gas accretion into the galaxy follows the accretion of dark matter, i.e. $\epsilon_{\rm acc} \approx 1$, is still permitted by the metallicity measurements, as shown in Figure 6.8.

The oxygen yield from Kobayashi et al. (2006) is about 0.01. This number is quite close to the blue lines in the upper panels of Figures 6.6 and 6.7, which are obtained by setting $\epsilon_{\rm loss} = 0$. This suggests that the combination of the Kobayashi et al. (2006) chemical evolution model with the metallicity measurements using detailed photoionization modeling strongly prefers a weak outflow scenario, even at $z \approx 2$.

6.5.3 Metal loss

The loading factor of metal loss rate $\epsilon_{\rm loss,Z}$ can be directly estimated from Eq. (6.14), and the estimate is independent of the rates of gas inflow and outflow. Figure 6.9 shows $\epsilon_{\rm loss,Z}$ as a function of halo mass at three different redshifts. As one can see, net metal outflow is always required, i.e. $\epsilon_{\rm loss,Z} > 0$, for different halos at different redshifts, regardless of the gas outflow. The loading factor predicted with the Kennicutt-Schmidt law is about 0.01, which is about 60% of the yield (indicated by the horizontal lines), and

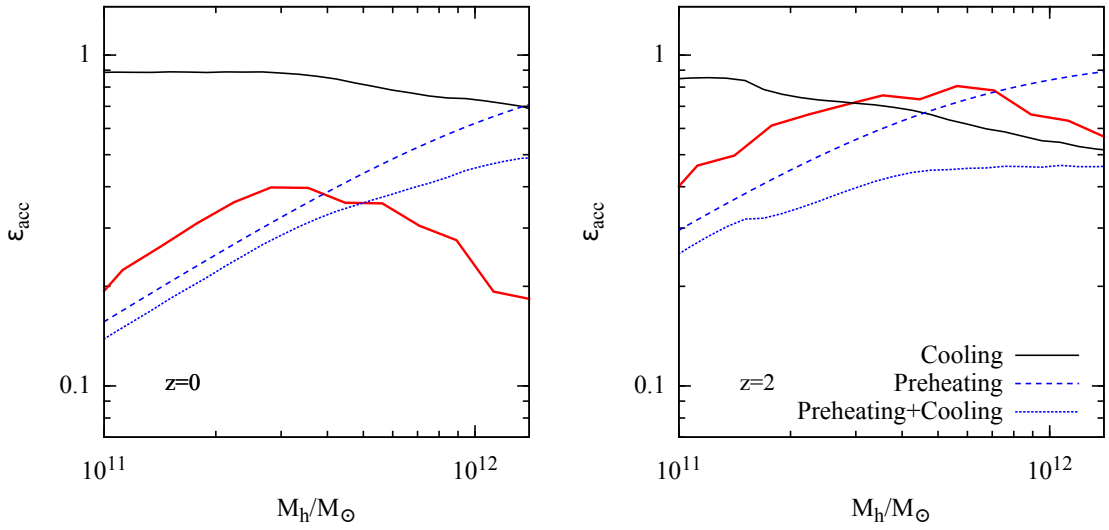

Figure 6.10. Comparison between the upper limit of the accretion efficiency derived from our empirical model (red lines) with a number of physical models. The black solid lines are cooling efficiency of halo gas (Croton et al., 2006). The blue dashed lines are the efficiency of accretion into dark matter halos in the preheating model of Lu et al. (2015a) and the blue dotted lines show the accretion to the galaxies in the same model.

depends only weakly on redshift and the mass of the host halos. This is consistent with the finding of Peeples et al. (2014), that is about 75% of the metals ever produced do not stay in the host galaxies. The prediction using the Krumholz model is similar except that at $z = 2$ the mass loading factor is lower. The reason for this difference is that the Krumholz model predicts higher cold gas fraction at $z = 2$, and so a larger fraction of newly produced metals can be stored in the ISM instead of going out with the wind.

6.6 Conclusions and Implications

In the present paper, we have combined up-to-date observational constraints, including the star formation - halo mass relations, the gas phase metallicity - stellar mass relations (Maiolino et al., 2008), and the gas mass fraction of local galaxies (Peeples and Shankar, 2011), and used a generic model to investigate how the contents, inflow and outflow of gas and metals evolve in the ecosystem of a low-mass galaxy. The goal is to understand the

underlying physics responsible for the low star formation efficiency in halos with masses between $10^{11}\,\mathrm{M}_{\odot}$ and $10^{12}\,\mathrm{M}_{\odot}$. Our conclusions are summarized in the following.

We adopt both the Kennicutt-Schmidt and the Krumholz models of star formation estimate the gas contents in galaxies up to $z = 2$. We find that (i) The gas mass to stellar mass ratio in general increases with redshift because of the increase of SFR; (ii) The Krumholz model predicts a higher gas mass fraction at high redshift than the Kennicutt-Schmidt model, especially in dwarf galaxies, because of its dependence on metallicity and because of the metallicity evolution of the ISM; (iii) The Krumholz model predicts that the ISM of galaxies is dominated by the molecular gas at $z > 1$, with the molecular gas to stellar mass ratio increasing from ~ 0.1 at $z = 0$ to ~ 1 at $z = 2$; (iv) The baryon mass ratio, $(M_{\rm g} + M_{\star})/M_{\rm h}$, is, since $z = 2$, always much less than the universal baryon mass fraction.

Using the gas mass estimated from the star formation laws together with other observational data, we derive constraints on the gas inflow and outflow rates. Independent of the gas outflow rate, metal outflow is always required at different redshift. The metal mass loading factor is about 0.01, or about 60% of the metal yield, and this factor depends only weakly on halo mass and redshift.

In spite of the degeneracy between gas inflow and outflow, and the uncertainties in modeling how metals are mixed with the medium, we can still put constraints on gas inflow and outflow. As the galactic wind material is expected to be more metal enriched than both the ISM and the material ejected at an earlier epoch, we can derive stringent upper limits on the accretion rate of primordial gas and on the net gas mass loss rate in the outflow. We find that (i) At $z \sim 0$, the low star formation efficiency is mainly caused by the low accretion rate. The maximum loading factor of the mass loss is about one while the maximum accretion efficiency factor $[\dot{M}_{\rm acc}/(f_{\rm b}\dot{M}_{\rm h})]$ is between 0.3 and 0.4; (ii) At $z \sim 2$, strong gas mass loss is allowed. The maximum loading factor allowed by the observational constraints is about 10 for $10^{11}\,\mathrm{M}_{\odot}$ halos, and is inversely proportional

to halo mass. These upper limits do not depend significantly on the star formation laws adopted, because the exact amount of gas in the galaxies is irrelevant in estimating the rate of gas exchange, as long as $M_{\rm g} \ll f_{\rm b} M_{\rm h}$, which is roughly the case based on our model inferences.

In a typical semi-analytic model of galaxy formation, the mass accretion into a *halo* is usually assumed to be $f_{\rm b}\dot{M}_{\rm h}$, and the mass accretion into the *central galaxy* is determined by the cooling rate of the gaseous halo. Following the cooling model of Croton et al. (2006) (also in Lu et al. 2011b) and assuming a metallicity of $0.1Z_{\odot}$ in the coronal gas, we calculate the efficiency of mass accretion in such a process, and the value of $\epsilon_{\rm cool} \equiv \dot{M}_{\rm cool}/(f_{\rm b}\dot{M}_{\rm h})$ is shown as the black solid lines in Figure 6.10. We see that $\epsilon_{\rm cool} \sim 1$ for a $10^{11}\,{\rm M}_{\odot}$ halo, and is roughly proportional to $M_{\rm h}^{-0.2}$ in the halo mass range shown in the figure. At $z = 2$ $\epsilon_{\rm cool}$ is close to the upper limit of $\epsilon_{\rm acc}$ we have derived, but at $z = 0$ it is significantly larger. The discrepancy between this prediction and our empirically derived constraint suggests that either accretion of the IGM into dark halos must be reduced or the cooling of the halo gas must be slowed down.

The scenario of galaxy formation in a preheated medium was first proposed in Mo and Mao (2002) in order to explain the observed SMFs and HI mass function. Lu et al. (2015a) suggested that the extended gas disks provide independent supports to such a scenario. They considered an "isentropic" accretion model, in which the IGM is assumed to be preheated to a certain level at $z < 2$ so that the gas accretion rate into low-mass halos is reduced. The hot gaseous halos formed in this way are less concentrated and cooling can happen even in the outer part of a halo, where the specific angular momentum is higher, producing a disk size - stellar mass relation that matches observation. Using the entropy model explored in Lu et al. (2015a) we have calculated the accretion efficiency of the pre-heated IGM into *dark matter halos*, which is shown as the blue dashed lines in Figure 6.10. The corresponding accretion rate of the central galaxies due to radiative cooling of the gaseous halos is shown as the blue dotted lines. At $z = 2$, the accretion

efficiency lies below the upper limit, and so the model is compatible with our results. At $z = 0$, the predicted accretion efficiency is consistent with the upper limit we obtained for halos with masses below $4 \times 10^{11}\,\mathrm{M}_{\odot}$ but is higher by a factor of ~ 2 for Milky Way mass halos.

It is still unclear how the IGM is preheated. In addition to the possibilities listed in §6.1, the intermediate mass central black holes can serve as a promising source (§ 5.3). Such black holes form from the major merger between dwarf galaxies at $z > 2$ and is able to heat the surrounding IGM to a entropy tested in Lu et al. (2015a).

For Milky Way mass halos, preventing the IGM from collapsing with the dark matter requires an entropy level that is much larger than what Lu et al. (2015a) suggests. Such a high level of preheating may over-quench star formation in smaller galaxies. It is more likely that some other *preventive* (rather than ejective) mechanisms may reduce *gas cooling* in such galaxies at low redshift, instead of preventing gas accretion into the *host halo*. For example, a central black hole may keep halo gas hot via the "radio mode" feedback, preventing it from further cooling (Croton et al., 2006). Clearly, it is important to examine if such "radio mode" feedback is also operating in Milky Way size galaxies, or other processes have to be invoked. Using hydrodynamic simulation of Milky Way mass galaxies, Kannan et al. (2014) found that the ionizing photons from local young and aging stars can effectively reduce the cooling of the halo gas, which may provide another promising preventive mechanism to reduce star formation efficiency in such galaxies at low redshift, as predicted by our empirical model.

CHAPTER 7

STAR FORMATION AND QUENCHING OF SATELLITE GALAXIES

7.1 Introduction

The impact of the intra-halo environments on the star formation in satellite galaxies has been noticed for decades (see §15.5 in Mo et al., 2010), and has been investigated in detail using groups/clusters of galaxies. For example, using the group catalogue constructed from SDSS DR2, Weinmann et al. (2006a) examined how the fractions of late and early type satellite galaxies change with galaxy luminosity and the mass of the halo that host the satellites, and found that the early type fraction increases strongly with halo mass. More recently, Wetzel et al. (2012) carried out a similar but more thorough analysis. They found that the sSFR of satellite galaxies in halos with different masses follows a bimodal distribution, but that the sSFR in the star forming sequence shows no strong dependence on halo mass. Wetzel et al. (2012) also investigated the excess in the quenched fraction of satellites relative to that of centrals, a quantity that reflects the strength of the impact of the intra-halo environments in quenching star formation, and found that this quantity increases with halo mass but does not depend strongly on the stellar mass of the satellites.

On the theory side, galaxy formation models are still struggling to reproduce the basic observational results regarding the quenching of star formation by environmental effects. By comparing results obtained from galaxy groups with the prediction of the semi-analytical model of Croton et al. (2006), Weinmann et al. (2006b) suggested that the simple "strangulation" model adopted by Croton et al. (2006) and many other versions of semi-analytic model, in which the diffuse gaseous halo is assumed to be stripped

immediately after a galaxy is accreted into a larger halo, significantly over-quenches the satellite population. Font et al. (2008) found that this "over-quenching" problem cannot be solved by a more physically motivated "strangulation" model, but may be solved by assuming that the ejected gas from a satellite by supernova feedback is recycled back to the satellite to sustain further star formation.

In a traditional semi-analytic model, a large number of prescriptions and model parameters are adopted to describe a variety of physical processes relevant to the formation and evolution of the galaxy population. This comprehensive nature of the model makes it hard to put tight constraints on specific physical processes, because of the degeneracy among different parts of the model (e.g. Lu et al., 2011b). In particular, even the latest semi-analytic models still have trouble in reproducing the bimodal distributions of the central galaxies in sSFR and color and their evolution (e.g. Lu et al., 2014c). Since satellite galaxies are believed to be evolved from centrals at higher redshift, the failure of a model in correctly reproducing the properties of central galaxies can also hinder the interpretation of the model prediction for the quenching of satellites. Several investigations have attempted to isolate the problem of satellite quenching from the total evolutionary process of galaxy evolution. For example, Balogh et al. (2000) studied the halo-centric distribution of quenched satellites by assigning the SFRs of present-day centrals to satellites at the time of accretion and modeling the subsequent star formation of satellites with the Kennicutt law (Kennicutt, 1998) and the cold gas reservoirs inherited from the progenitor centrals. Wetzel et al. (2013) used a set of observational constraints to initialize the SFRs of satellites, and modeled the subsequent evolution of star formation in the satellites by assuming that a rapid quenching (an exponential decline in SFR) of a satellite occurs at a certain time after it is accreted by a larger halo. These phenomenological models can be used to translate observational constraints into certain physically meaningful quantities, such as the quenching time scale, but they themselves are not linked directly to any particular physical processes.

In this chapter, we study the quenching of satellites by intra-halo environments. Our approach is different from those of previous investigations in two aspects. First, the stellar masses, SFRs and gas contents of the progenitors of satellite galaxies at the time of accretion are set by using the results shown in Chapter 6 constrained empirically by a broad range of observations. Second, the subsequent evolution of satellite galaxies is modeled with physically motivated processes. We test our model predictions by comparisons with results obtained from groups of galaxies. This chapter is arranged as follows. The properties of the progenitors of satellite galaxies are described in §7.2. Our models for the evolution of satellite galaxies in dark matter halos are described in §7.3. Model predictions are presented in §7.4, and a summary of our main results is given in §7.5.

7.2 Progenitors of Satellite Galaxies

We start to trace the evolution of a satellite at the time when it was first accreted into a bigger halo. The distribution of sub-halos in accretion redshift $z_{\rm acc}$ and in initial mass $M_{\rm s,i}$ is obtained from halo merger trees generated with the algorithm of Parkinson et al. (2008). More details can be found in § 2.2.

Each sub-halo at the time of accretion is assigned a stellar mass and a SFR according to its mass and accretion redshift using the empirical model (§ 6.3). A cold gas mass is derived from its SFR using a star formation law, as described in § 6.4. In the present model, stars and cold gas are assumed to be distributed in thin disks. For simplicity, we assume that the surface densities of both the stellar and gas disks follow truncated exponential profiles. For the gas, it is

$$\Sigma_{\rm g}\left(r\right)=\begin{cases}\Sigma_{\rm g,0}\exp\left(-\frac{r}{R_{\rm g}}\right) & \text{if } r<R_{\rm tr,g}\\ 0 & \text{otherwise}\end{cases} \tag{7.1}$$

where $\Sigma_{\rm g,0}$ is the central surface density, $R_{\rm g}$ the scale radius, and $R_{\rm tr,g}$ the truncation radius. At the time of accretion of a satellite, the truncation radius is set to ∞. The sub-

sequent evolution of the truncation radius is determined by different stripping processes, as described in the following sections. The total mass within the truncation radius is

$$M_{\rm g} = 2\pi R_{\rm g}^2 \Sigma_{\rm g,0} \left[1 - \left(1 + \frac{R_{\rm tr,g}}{R_{\rm g}}\right) \exp\left(-\frac{R_{\rm tr,g}}{R_{\rm g}}\right)\right] . \tag{7.2}$$

The stellar disks are assumed to follow a profile similar to the gas but with a different scale radius $R_\star$ and a different truncation radius $R_{\rm tr,\star}$. Both $R_\star$ and $R_{\rm g}$ are set using the model described in § 6.3, in which the ratio $\mathcal{L} \equiv R_{\rm g}/R_\star$ is assumed to be a constant larger than one.

The initial gas mass $M_{\rm g,i}$ is obtained from the initial SFR (SFR) through an adopted star formation law. Specifically, we write

$$\mathrm{SFR} = \int_0^\infty \Sigma_{\rm SF}\left(\Sigma_{\rm g}\right) 2\pi R \mathrm{d}R \,, \tag{7.3}$$

where $\Sigma_{\rm SF}$ is the SFR surface density, assumed to be determined by the cold gas mass surface density, $\Sigma_{\rm g}$. In this paper we implement the Kennicutt-Schmidt law (Kennicutt, 1998),[1] which gives

$$\Sigma_{\rm SFR} = \begin{cases} A_{\rm K} \left(\frac{\Sigma_{\rm g}}{\rm M_\odot pc^{-2}}\right)^{N_{\rm K}} & \text{if } \Sigma_g \geq \Sigma_{\rm c}, \\ 0 \quad \text{if } \Sigma_{\rm g} < \Sigma_{\rm c} \,. \end{cases} \tag{7.4}$$

The power index $N_{\rm K} \approx 1.4$, and the amplitude $A_{\rm K} \approx 2.5 \times 10^{-4}\,{\rm M_\odot yr^{-1} pc^{-2}}$. These parameters are constrained using the gas mass - stellar mass relation of local galaxies (Lu et al., 2015b, and references therein).

[1]We have made a test using the star formation model of Krumholz et al. (2009). The results are qualitatively the same.

7.3 Evolution of Satellite Galaxies

7.3.1 Orbits of Satellites

The orbits of satellite galaxies in host dark matter halos are modeled according to the orbits of sub-halos within which the satellites are located. The host halos are assumed to follow the NFW profile (Navarro et al., 1996):

$$\rho_{\rm D}(r) = \frac{M_{\rm h}}{4\pi R_{\rm vir}^3} \frac{c^2}{\ln(1+c) - c/(1+c)} \frac{1}{x(1+cx)^2} \tag{7.5}$$

where $M_{\rm h}$ is the halo mass, $R_{\rm vir}$ the radius of the halo, and $x \equiv r/R_{\rm vir}$. The concentration parameter, c, is modeled with the c-$M_{\rm h}$ relation given by Bullock et al. (2001). Within the mass and redshift ranges we are interested in, the model is a reasonably good approximation to the more recent models of the halo concentration - mass relation (e.g. Zhao et al., 2009; Dutton and Macciò, 2014). The corresponding gravitational potential of this density profile is

$$\Phi_{\rm D}(r) = -V_{\rm vir}^2 \frac{1}{\ln(1+c) - c/(1+c)} \frac{\ln(1+cx)}{x}, \tag{7.6}$$

where $V_{\rm vir}^2 \equiv GM_{\rm h}/R_{\rm vir}$ is the virial speed of the halo.

At the time of accretion, that is when a sub-halo crosses the virial radius ($R_{\rm vir}$) of a larger halo, the infall velocity of the sub-halo is set to be exactly the virial velocity $V_{\rm vir}$. This is a natural prediction of the spherical collapse model, and is used here as an approximation. The initial orbital angular momentum of the sub-halo is described by its orbital circularity, η, which is defined to be the ratio between the true orbital angular momentum of the sub-halo and the orbital angular momentum of a circular orbit with the same orbital energy as the sub-halo. The distribution of this quantity has been determined using high resolution simulations. Here we use the result of Zentner et al. (2005), which gives

$$P(\eta) \propto \eta^{1.2}(1-\eta)^{1.2}\,. \tag{7.7}$$

The subsequent evolution of the sub-halo is traced by following the model of Taylor and Babul (2001), which includes dynamic friction and tidal stripping. As the sub-halo moves through the background of dark matter particles, a drag force (dynamic friction) acts to reduce its angular momentum and energy. This drag force is modeled by Chandrasekhar's formula,

$$\begin{aligned} \mathbf{F}_{df} &= -4\pi G^2 M_{\rm s}^2 \ln\Lambda \rho_{\rm D}(\mathbf{r}) \\ &\left[\mathrm{erf}\left(\frac{v_{\rm s}}{V_{\rm c}}\right) - \frac{2}{\sqrt{\pi}}\frac{v_{\rm s}}{V_{\rm c}}\mathrm{e}^{-(v_{\rm s}/V_{\rm c})^2}\right]\frac{\mathbf{v}_{\rm s}}{v_{\rm s}^3}\,, \end{aligned} \tag{7.8}$$

where M_s is the instantaneous mass of the sub-halo, $\mathbf{v}_{\rm s}$ is its velocity, and $V_{\rm c}$ is the local circular velocity of the host halo. The Coulomb logarithm is estimated as $\ln\Lambda = \ln(M_{\rm h}/M_{\rm s})$, where $M_{\rm h}$ is mass of the host halo.

The instantaneous tidal radius of the sub-halo is estimated using

$$R_{\rm tidal} \approx \left(\frac{GM_{\rm s}\,(<R_{\rm tidal})}{\omega^2 - \frac{\partial^2\Phi(r)}{\partial r^2}}\right)^{1/3}, \tag{7.9}$$

where ω is the instantaneous orbital angular velocity. This tidal radius is calculated at each time step when the orbit of the satellite is integrated. Following Taylor and Babul (2001), the mass outside the tidal radius is assumed to be stripped at the rate given by

$$\dot{M}_{\rm s} = \frac{\omega}{2\pi} M_{\rm s}\,(\geq R_{\rm tidal})\ . \tag{7.10}$$

The sub halo mass so obtained is then used in the dynamic friction calculation at each time step. Note that the disk contained in a sub-halo is also assumed to be tidally truncated, so that the disk truncation radius, $R_{\rm tr,g}$, introduced in §7.2, is at most $R_{\rm tidal}$.

The orbit of each sub-halo (and of the galaxies it contains) is integrated starting from the time of accretion with the initial conditions specified above, taking into account both gravity and dynamical friction. We stop tracing the orbit of a sub-halo when it has lost

all its orbital angular momentum. At this point, the satellite galaxy associated with the sub-halo is assumed to have merged with the central galaxy of the host halo.

7.3.2 High-order Substructures

When a group of galaxies merges with a more massive system, some of the galaxies and substructures may still be associated with their parent halos, becoming high order substructures. Others may be tidally stripped, moving on independent orbits in the new, bigger halo. Whether a substructure will become a high-order substructure after merging depends on how they are gravitationally bound to each other as well as on the orbit of their parent halo. An accurate treatment of the high-order substructures, therefore, requires a recursive scheme: we need to trace not only the trajectories of the sub-halos but also the trajectories of the sub-sub-halos within them. A detailed treatment of this makes the model very complicated and is beyond the goal of this paper. Instead, we adopt the pruning scheme developed by Taylor and Babul (2004), in which the substructures that are loosely bound to the parent halo are assumed to be directly stripped when the merger happens. An empirical criterion suggested by the authors is that if a substructure has spent less than a period of $n_0 P_{\rm rad}$ in the parent halo, it is considered loosely bound, where $n_0 = 2$ is a free parameter calibrated using more accurate models, and $P_{\rm rad}$ is the mean orbital period. Each of the stripped sub-halos is assigned a new orbital parameter and is followed the same way as the other sub-halos in the new halo. About 70% of the substructures are stripped from the parent halos according to this model. The remaining, tightly bound part is assumed to stay with their parent until the parent itself is destroyed.

7.3.3 Star Formation and Outflow

After a satellite is accreted by a larger halo, its cold gas can be depleted by different processes. The first is star formation and galactic outflows driven by the star formation. The gas mass depletion rate due to this process is

$$\dot{M}_{\rm g} = (1 - \mathcal{R} + \epsilon_{\rm w}) \int_0^{R_{\rm tr,g}} \Sigma_{\rm SF}\left(\Sigma_{\rm g}\right) 2\pi R {\rm d}R\,, \tag{7.11}$$

where $\Sigma_{\rm SF}\left(\Sigma_{\rm g}\right)$ is the star formation law, $\mathcal{R}$ is the recycled fraction from stellar evolution and $\epsilon_{\rm w}$ is the loading factor of the mass outflow driven by star formation. For simplicity, we assume that this process changes neither the scale radius nor the truncation radius of the disk, but only reduces the central surface density $\Sigma_{\rm g,0}$, so that

$$\frac{\dot{\Sigma}_{\rm g,0}}{\Sigma_{\rm g,0}} = \frac{(1 - R + \epsilon_{\rm w})\, SFR}{M_{\rm g}}\,. \tag{7.12}$$

The star formation in the gas disk is assumed to follow the same model as described in §7.2.

In addition to star formation and outflow, other processes, such as ram-pressure stripping and tidal interactions, may also deplete the disk gas outside a certain radius at which the external forces are balanced by the gravitational potential of the satellite itself. We assume that the stripping does not change the inner structure of the disk, so that $\Sigma_{\rm g,0}$ and $R_{\rm g}$ are not affected. The detailed modeling of these processes are described in the following.

7.3.4 Ram Pressure Stripping

The host halos are assumed to be filled with diffuse hot gas. The total mass of the diffuse gas within the virial radius is modeled on the basis of the hot mode accretion presented in Kereš et al. (2005) and Lu et al. (2011b), and the hot gas fraction is given by

$$f_{\rm hot} = \frac{1}{2}\left(1 + {\rm erf}\left[\frac{\log_{10}\left(M_{\rm h}/M_{\rm tran}\right)}{\sigma_{\rm tran}}\right]\right) \tag{7.13}$$

where $\log_{10}\left(M_{\rm tran}\right) = 11.4$ is the transition halo mass, and $\sigma_{\rm tran} = 0.4$.

The diffuse gas is assumed to be isothermal and has the virial temperature $\frac{1}{2}(\mu m_{\rm p}/k_{\rm b})V_{\rm vir}^2$, where μ is the mean molecular weight of the gas. Assuming hydrostatic equilibrium, the

density profile of the diffuse gas is determined by the gravitational potential of the host halo,

$$\rho_{\rm X}(r) = \rho_{\rm X,0} \exp\left[-\frac{\mu m_{\rm p}}{k_{\rm B} T_{\rm vir}} \Phi_{\rm D}\left(r\right)\right] \tag{7.14}$$

The integral constant $\rho_{\rm X,0}$ is given by the total mass of the diffuse gas through

$$f_{\rm hot} \frac{\Omega_{\rm b,0}}{\Omega_{\rm m,0}} M_{\rm h} = \int_0^{R_{\rm vir}} \rho_{\rm X}(r) 4\pi r^2 {\rm d}r \,. \tag{7.15}$$

The gravitational restoring force per unit area of a infinitesimally thin disk is

$$P_{\rm gr} = 2\pi G \Sigma_\star \Sigma_{\rm g}, \tag{7.16}$$

where $\Sigma_\star$ and $\Sigma_{\rm g}$ are the surface densities of stars and gas at the radius in question, respectively. For a face-on disk, the truncation radius owing to ram pressure stripping can be calculated by setting the gravitational restoring force to be equal to the ram pressure $\rho_{\rm X}\left(r\right) v_{\rm s}^2$, which gives

$$R_{\rm rp} = \frac{R_\star R_{\rm g}}{R_\star + R_{\rm g}} \ln\left(\frac{2\pi G \Sigma_{\star,o} \Sigma_{\rm g,o}}{\rho_{\rm X}\left(r\right) v_{\rm s}^2}\right) . \tag{7.17}$$

In general, if the disk is inclined relative the moving direction, the ram pressure exerted on the disk depends on the inclination angle. Analytically deriving this inclination dependence is difficult, as the stripping becomes asymmetric when the disk in inclined. Using a set of hydrodynamic simulations, Roediger and Brüggen (2006) measured the mean truncation radius as a function of both the ram pressure and the inclination angle. It is found that the prediction of Eq. (7.17) only deviates from the simulations when the inclination angle $\theta \geq 60°$. To take this into account, we add a correction term to Eq. (7.17) so that it is consistent with the simulation result:

$$R_{\rm rp} = \frac{R_\star R_{\rm g}}{R_\star + R_{\rm g}} \left[\ln\left(\frac{2\pi G \Sigma_{\star,o} \Sigma_{\rm g,o}}{\rho_{\rm X}\left(r\right) v_{\rm s}^2}\right) - 0.8 \ln\left(\cos\theta\right)\right] , \tag{7.18}$$

where $\cos\theta = |\mathbf{v}_s \cdot \mathbf{s}|/v_s$. The spin axis $\mathbf{s}$ of the disk is assigned at the time of accretion and is assumed to be independent of any other quantities.

In addition to ram pressure stripping, gas disk can also lose mass because of Kelvin-Helmholtz instability. However, according to Roediger and Brüggen (2007), this mass loss rate is far less efficient than ram pressure stripping, and so it will be ignored in our model.

7.3.5 Tidally Triggered Starburst

Interaction between galaxies is another possible process that can lead to the quenching of star formation. As a satellite galaxy passes by the central galaxy, the tidal field from the central region can strip the outskirts of the disk of the satellite. This direct tidal stripping is modeled with the tidal radius described in §7.3.1. In addition, the tidal interaction can also exert a torque on the disk, causing the disk gas to lose angular momentum and funneling it into the nuclear region to trigger a star burst. This process may also quench the satellite by consuming the star forming gas quickly.

The enhancement of the star formation in flyby satellites has been found both observationally (Li et al., 2008) and theoretically (Di Matteo et al., 2007). In particular, Li et al. (2008) found that, for star-forming galaxies, (i) the enhancement of star formation in the flyby satellites occurs when the projected distance is less than $5r_{90}$, where r_{90} is the radius that encloses 90% of the light of the central galaxies; (ii) the enhancement is a factor of about 3; (iii) the results depend only weakly on the mass of the companion.

Motivated by the findings of Li et al. (2008), we consider a simple phenomenological model to emulate the star burst triggered by tidal interaction. We assume that the tidal interaction becomes important only when a satellite is within some radius from the halo center. This radius is assumed to scale with the size of the central galaxy as $R_{\rm sb} = x_{\rm sb}R_\star$, where $R_\star$ is the scale radius of the stellar disk and is estimated following the redshift-dependent galaxy size - stellar mass relation as described in §6.3. When a satellite reaches $R_{\rm sb}$, the satellite is switched to a "star burst" mode, in which the SFR is assumed to be

enhanced by a factor of $\mathcal{E}_{\rm enhance}$ relative to its SFR immediately before it reaches $R_{\rm sb}$. This enhanced SFR is assumed to last until the gas is completely exhausted or the satellite gets out of the sphere, whichever comes first. If the gas is not exhausted before the satellite moves out of the sphere, the galaxy will resume its pre-burst mode of star formation, but with SFR calculated from the reduced amount of cold gas. We also assume that the star burst generates an outflow with a loading factor $\epsilon_{\rm w,sb}$, which may be different from $\epsilon_{\rm w}$.

7.3.6 Environmental Quenching versus Self-Quenching

Quenched galaxies exist in a variety of environments, although satellites tend to have a higher fraction of quenched population than centrals of similar stellar masses. It is believed that the physical processes responsible for the quenching of centrals are not related to environments. A widely adopted scenario is quasar-mode feedback, which can eject the cold gas in a galaxy and quench its star formation in a very short time scale. Such mechanisms are usually referred to as "self-quenching", to distinguish them from environmental processes. Self-quenching is expected to have also operated in the observed population of satellites before or after the their accretions by larger halos. If one assumes that self-quenching and environmental quenching are independent of each other, the excess of quenched fraction in satellites relative to that in centrals can be used as an indicator of the efficiency of environmental quenching. This excess can be described by the ratio

$$f^{\rm sat}_{\rm Q,excess} = \frac{f^{\rm sat}_{\rm Q,now} - f^{\rm cen}_{\rm Q,now}}{f^{\rm cen}_{\rm A,now}}\,, \tag{7.19}$$

where $f_{\rm Q}$ denotes the quenched fraction, $f_{\rm A}$ the total fraction, and the superscripts, 'cen' and 'sat', indicate centrals and satellites, respectively. This quantity has been estimated by Wetzel et al. (2012) and Wetzel et al. (2013) using galaxy groups selected from the SDSS. In our investigation, we will focus on environmental quenching without modeling the mechanisms that are responsible for the quenching of centrals. In what follows, we will

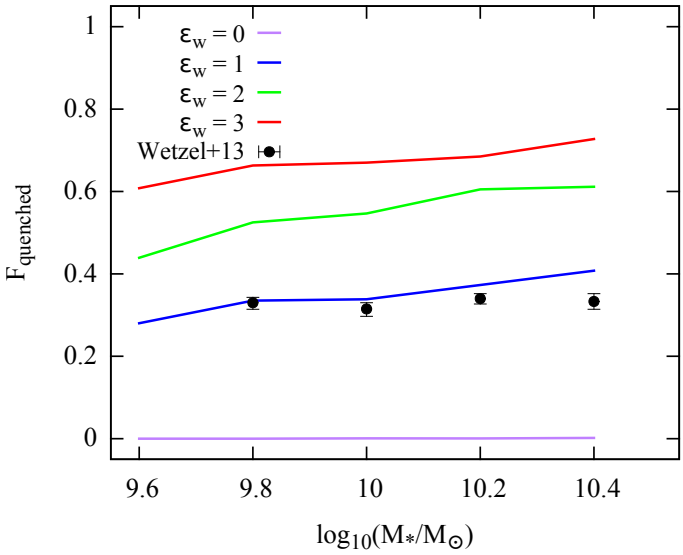

Figure 7.1. Fractions of quenched galaxies predicted by the model in which disk gas is consumed only by star formation and associated outflow. Different colors represent different outflow mass loading factors, as indicated in the panel. The data points are taken from Wetzel et al. (2013).

refer $f_{\rm Q,excess}^{\rm sat}$ as the quenched fraction without including the population that is quenched before the galaxies become satellites.

7.4 Model Predictions

7.4.1 Quenching by Strangulation

In the literature, strangulation refers to a hypothesis that a galaxy in a sub-halo will stop accreting new gas as soon as the sub-halo is accreted into a big halo. Because of strangulation, star formation in satellite galaxies will gradually slow down and eventually stop as the gas reservoir is consumed by star formation, even in the absence of any other additional processes. Galactic wind driven by star formation speeds up the depletion of the cold gas. This is generally how red satellites are produced in many semi-analytical models in the literature (e.g. Lu et al., 2014c). To understand how this process works, we consider a model in which all other environmental processes (ram pressure stripping

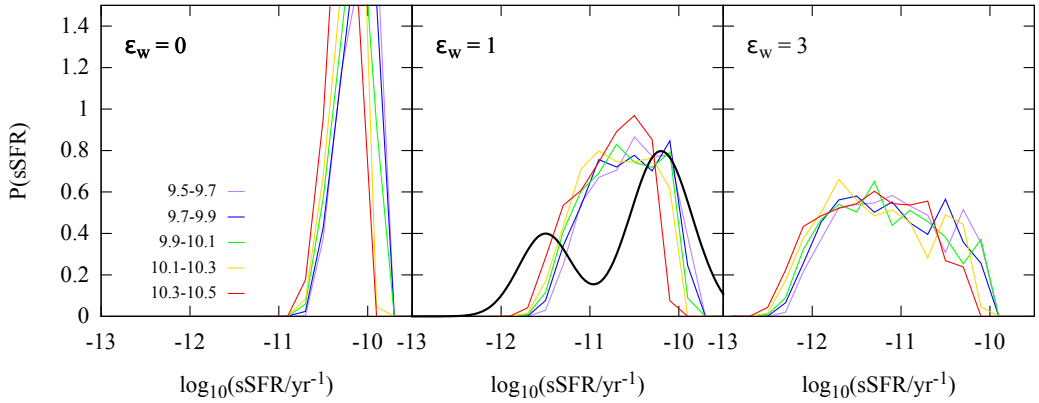

Figure 7.2. The distribution of the sSFR of satellite galaxies predicted by the model in which disk gas is consumed only by star formation and associated outflow. Different panels correspond to different outflow mass loading factors. The stellar mass range of satellites is coded in different colors, as indicated in the left panel. The black solid line in the middle panel is the observed distribution inferred from Wetzel et al. (2012) and Wetzel et al. (2013).

and tidally triggered starburst) are switched off, so that we can focus on quenching by gradual gas consumption owing to star formation and the associated outflow.

Figure 7.1 shows the impact of the outflow loading factor, $\epsilon_{\rm w}$, on the predicted quenched fraction of satellites. All the satellites stay as star forming galaxies if $\epsilon_{\rm w}$ is set to be 0. This means that the cold gas reservoir left over from the time of accretion is sufficient to keep satellite galaxies active in star formation all the way to the present day if there is no other gas depletions than star formation. The observational constraints seem to favor $\epsilon_{\rm w} = 1$ and adopting a loading factor $\epsilon_{\rm w} > 1$ tends to lead to over-quenching of the satellite population. Unfortunately, the loading factor of outflow is still poorly understood. In hydrodynamic simulations and semi-analytic models, a large value seems to be required to lower the efficiency of star formation to reconcile the tension between the observed galaxy luminosity/SMF (dominated by central galaxies) and the halo mass function at the low-mass end e.g. Yang et al., 2003. For instance, in the Somerville model studied in Lu et al. (2014c), the loading factor of a Milky-Way like halo is about 3 and is more than 10 for a $10^{11}\,{\rm M}_{\odot}$ halo. In many semi-analytic models, the outflow loading factor

assumed for satellite galaxies is the same (high value) as for centrals, and so the satellite population is expected to be over-quenched according to the results shown in Figure 7.1.

Indeed, satellite galaxies are predicted to be too red, with too little current star formation, in many semi-analytic models (e.g. Lu et al., 2014c), and strangulation is usually blamed for this over-quenching problem (Weinmann et al., 2006b). However, detailed modeling of the stripping of the diffuse gas halo suggests that the problem cannot be solved by continuous accretion of fresh gas from the parent halo (Weinmann et al., 2010; Font et al., 2008). In order to solve this problem, Weinmann et al. (2010) and Kang and van den Bosch (2008) had to assume that ram pressure does not make any significant stripping of the diffuse halo gas. Indeed, Font et al. (2008) assumed that most of the gas reheated by star formation - driven wind in a satellite will not be stripped, but will instead be re-accreted by the satellite, so that the gas reservoir originally associated with a satellite is always available for star formation, even though large amounts of gas is assumed to be in the wind.

The assumption of strong galactic winds for satellite galaxies needs to be revisited, however. In fact, to match theory with observation does not require a large loading factor for central galaxies. For instance, using empirically constrained star formation-halo mass relation and the gas phase metallicity, Lu et al. (2015b) found that the mass loading factor of galaxy wind required to suppress star formation in present-day central galaxies is no larger than unity if the accretion of pristine gas into a halo can be suppressed by preheating. If satellite galaxies inherit the same star formation and feedback from their progenitors, then there will be no over-quenching problem, as shown above.

As shown in Figure 7.2, the sSFR distribution predicted by this gradual gas depletion model is always unimodal, and increasing the loading factor only broadens the distribution without changing the unimodal nature. This is in contrast with the observational results that the sSFR of satellite galaxies follows bimodal distribution (shown by the smooth thick solid curve in the middle panel) and such a distribution persists in different environments

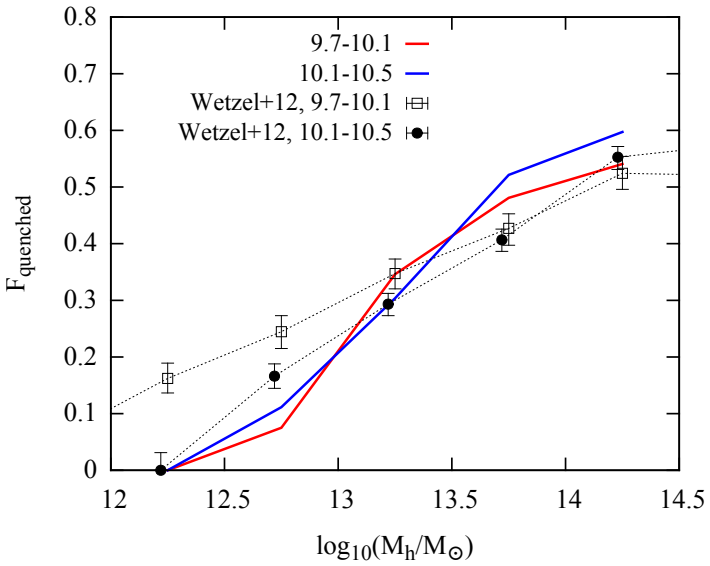

Figure 7.3. Fraction of quenched satellites as a function host halo mass, predicted by the model in which disk gas is consumed only by star formation and associated outflow. The solid colored lines are predictions by the model assuming $\epsilon = 1$. The data points with error bars are taken from Wetzel et al. (2012).

(e.g. Wetzel et al., 2013). This demonstrates that strangulation alone cannot explain the observed quenching of satellite galaxies.

Figure 7.3 shows the quenched fraction as a function of host halo mass. Here results are only shown for the $\epsilon_{\rm w} = 1$ model, as it correctly reproduces the overall quenched fraction. The quenched fraction increases with halo mass, because satellites of the same mass were accreted earlier in more massive halos and so a larger fraction of them have exhausted their gas by present day. The model under-predicts the quenched fraction in halos with masses $< 10^{13}\,{\rm M}_{\odot}$ in comparison with observational data, demonstrating again that strangulation alone cannot explain the observed quenching of satellite galaxies in different halos, even if the outflow loading factor is tuned so that the model correctly predicts the overall quenched fraction of the satellite population.

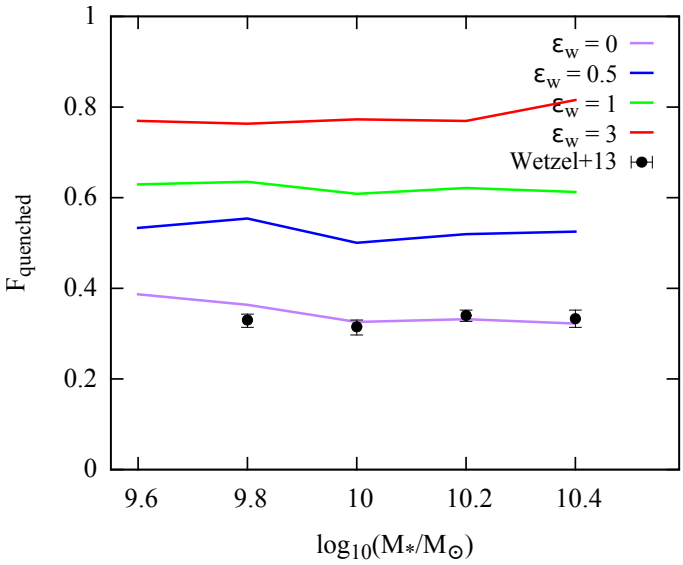

Figure 7.4. The predicted fractions of quenched galaxies by the model with ram-pressure stripping. Different colors represent different outflow mass loading factors, as indicated in the panel. The data points are taken from Wetzel et al. (2013).

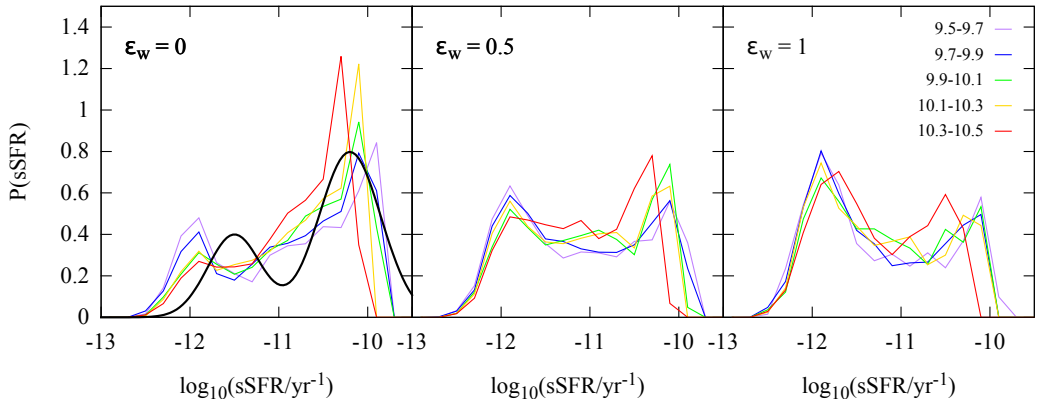

Figure 7.5. The distribution of sSFR of satellite galaxies predicted by the model of ram-pressure stripping. Different panels correspond to different assumed outflow mass loading factors. The stellar mass range of satellites is coded in different colors, as indicated in the right panel. The black solid line in the first panel is the observational distribution inferred from Wetzel et al. (2012) and Wetzel et al. (2013).

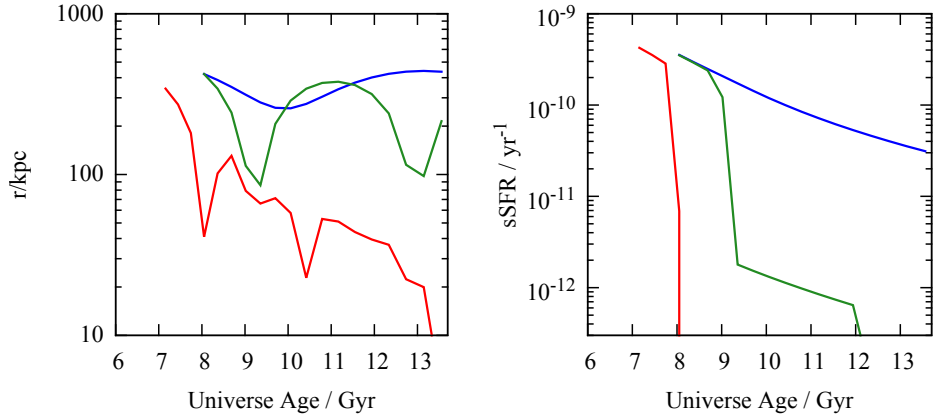

Figure 7.6. The evolution history of three satellites on three different initial orbits (red: nearly radial; blue nearly circular; green: intermediate). The left panel shows the distance to the host center as a function of time, and the right panel shows the sSFR as a function of time. The host halos are of $10^{13}\,h^{-1}\mathrm{M}_{\odot}$ at present day.

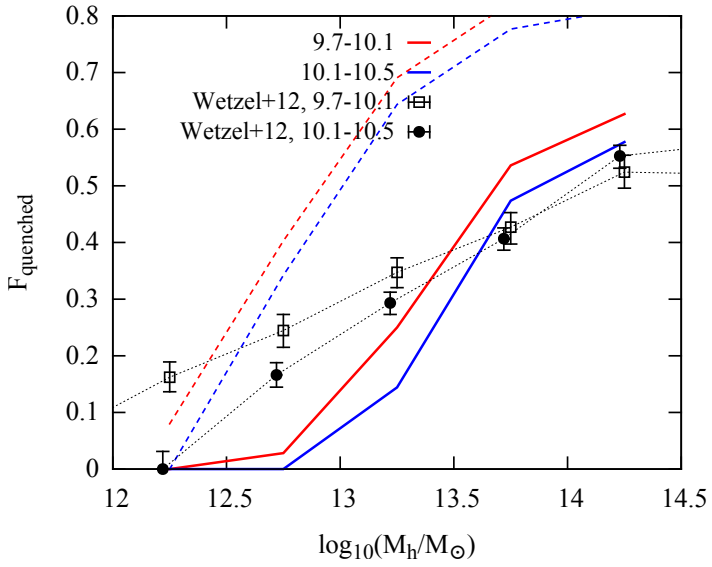

Figure 7.7. The fraction of quenched satellites as a function host halo mass. The colored solid lines are predictions by the model with ram pressure stripping and with $\epsilon = 0$, and the colored dashed lines are predictions of the model with $\epsilon = 1$. Data points with error bars are taken from Wetzel et al. (2012).

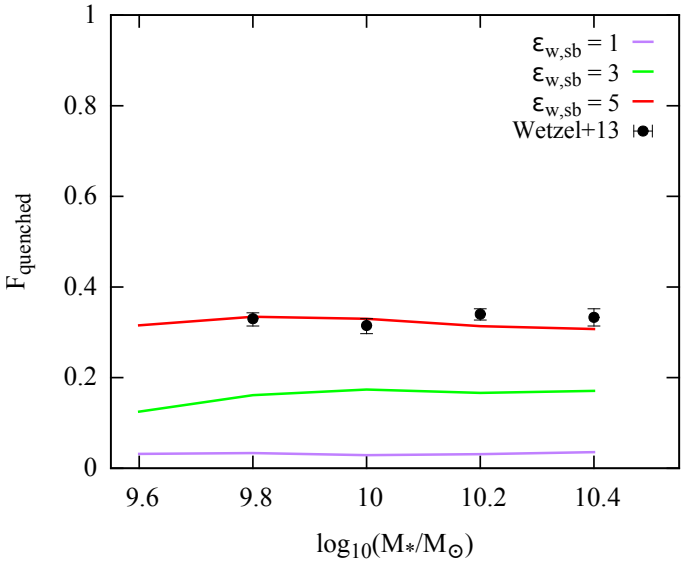

Figure 7.8. The predicted fractions of quenched galaxies by the model of tidally-triggered starburst. Different colors represent different outflow mass loading factors assumed for the starburst phase, as indicated in the panel. The data points are taken from Wetzel et al. (2013).

7.4.2 Quenching by ram pressure stripping

Next we include ram-pressure stripping in a way described in §7.3.4. Figure 7.4 shows the overall quenched fraction predicted by this model and how it changes with the assumed outflow loading factor. As one can see, when ram-pressure stripping is included, any loading factor $\epsilon_{\rm w} > 0$ tends to over-quench the satellite population, and the model prefers $\epsilon_{\rm w} = 0$, i.e. no outflow. Setting $\epsilon_{\rm w}$ to 1, the model already overestimates the quenched fraction by 80%. The two depletion mechanisms, outflow and ram-pressure stripping, do not act independently. Galactic wind depletes the gas at the central region of the disk, where most stars form. This reduces the gas surface density, thereby reducing the gravitational restoring force and making the ram pressure stripping more effective [see Eq. (7.18)].

Unlike gradual gas consumption, ram pressure stripping produces a bimodal distribution in the sSFR, as shown in Figure 7.5. With $\epsilon_{\rm w} = 0$, the model match the observed distribution (the left panel) qualitatively. To understand how the bimodal distribution is

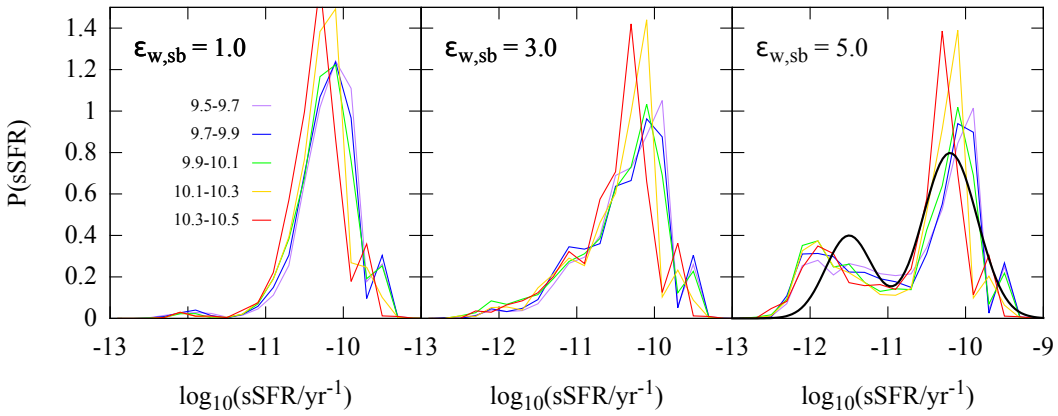

Figure 7.9. The distribution of the sSFR of satellite galaxies predicted by the model of tidally-triggered starburst. Different panels correspond to different outflow mass loading factors assumed for the starburst. The stellar mass range of satellites is coded in different colors, as indicated in the left panel. The black solid line in the last panel is the observational distribution inferred from Wetzel et al. (2012) and Wetzel et al. (2013).

produced, let us have a close look at how quenching proceeds in this model. Based on the model described in §7.3.4, a complete stripping of the gas disk occurs when the density of the surrounding medium reaches a critical value, which can be obtained from Eq. (7.17) by setting $R_{\rm rp} = 0$,

$$\rho_{\rm X,c} \sim \frac{2\pi G \Sigma_{\star,o} \Sigma_{\rm g,o}}{V_{\rm vir}^2}, \tag{7.20}$$

where we have replaced the speed of the satellite, $v_{\rm s}$, by the virial speed $V_{\rm vir}$ of the host halo. Only galaxies that have reached such a high density region can be completely quenched. Figure 7.6 shows the evolution of satellites on three typical orbits in a host halo with mass of $10^{13}\,{\rm M}_\odot$: a nearly circular orbit (blue), a nearly radial orbit (red) and an elliptical orbit in the middle (green). Although it has been accreted for about 6 Gyrs, the satellite on the circular orbit is still actively forming stars at present day. In contrast, the satellite on the radial orbit has almost completely lost its gas during the first approach to the peri-center, which happened less than 1 Gyr after accretion.

Compared with observational results, the predicted quenched fraction of satellites by ram-pressure stripping alone has much stronger dependence on host halo mass, quite

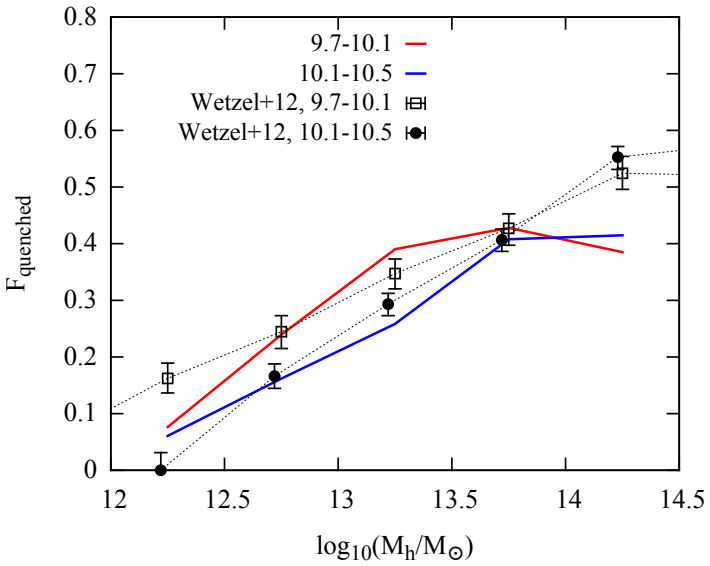

Figure 7.10. The fraction of quenched satellites as a function host halo mass by the model of tidally-triggered starburst with $\epsilon_{\rm sb,w} = 5$. Data points with error bars are taken from Wetzel et al. (2012).

independent of the assumed loading factor, as shown in Figure 7.7. There is a transition at halo mass about $2 \times 10^{13}\,{\rm M}_\odot$ (for $\epsilon = 0$). In more massive halos satellites are slightly over-quenched (by about 10%), while in halos with masses less than $5 \times 10^{12}\,{\rm M}_\odot$ ram pressure by the halo gas seems unable to quench any satellites. This strong dependence on host halo mass is also responsible for the bimodal distribution seen in Figure 7.5: most of the star-forming satellites are from the low-mass halos ($< 10^{13}\,{\rm M}_\odot$) while most the quenched satellites are from more massive halos ($> 5^{13}\,{\rm M}_\odot$).

A simple analysis can be used to understand this strong dependence on halo mass. The critical density for complete stripping given by Eq. (7.20) is proportional to $\rho_{\rm X,c} \propto V_{\rm vir}^{-2}$. Assuming isothermal sphere gas distribution for the host halo, this critical density corresponds to a critical radius,

$$x_c \equiv \frac{r_{\rm c}}{R_{\rm vir}} \propto V_{\rm vir}\,. \tag{7.21}$$

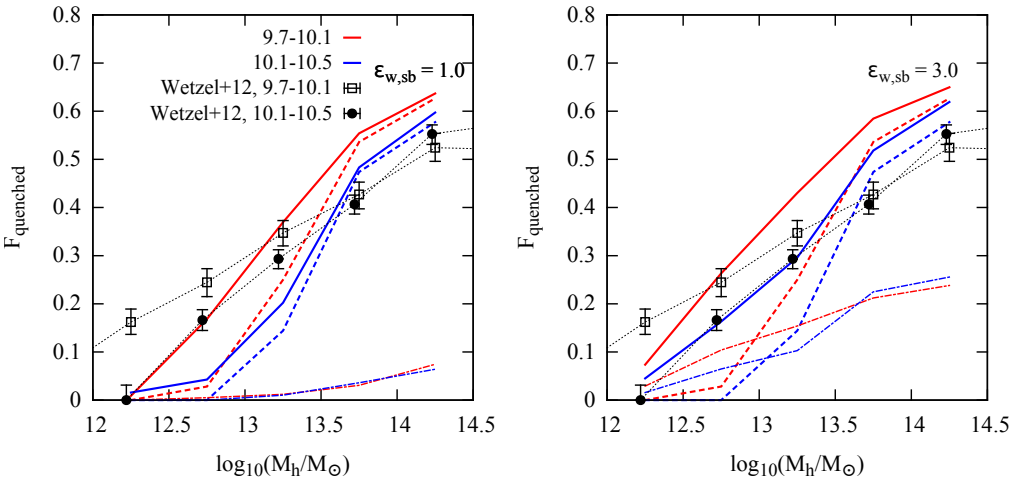

Figure 7.11. The fraction of quenched satellites as a function host halo mass. The dashed, dotted and solid lines correspond to models with only ram pressure, with only tidally triggered starburst, and with both processes, respectively. Left and right panels assume $\epsilon_{\rm sb,w} = 1$ and 3, respectively.

It is this dependence on the virial velocity that produces the strong dependence of the quenched fraction on the host halo mass shown in Figure 7.7.

7.4.3 Quenching by Tidally Triggered Starburst

Next, we switch off ram pressure stripping and investigate how the tidally-triggered star burst shapes the population of the present day satellites. We can obtain a rough estimate of $x_{\rm sb}$ from Li et al. (2008). For an exponential disk $r_{90} \approx 4R_\star$. As shown by Li et al. (2008), significant enhancement occurs at a radius smaller than $\sim 5r_{90}$, which leads to $x_{\rm sb} \approx 20$. For simplicity we set $\epsilon_{\rm w}$ to 0 and $\mathcal{E}_{\rm enhance}$ to 3 and only tune the value of $\epsilon_{\rm w,sb}$ to investigate the behavior of this model.

The overall quenched fractions as functions of stellar mass are shown in Figure 7.8, and the distributions of sSFR are shown in Figure 7.9. A modest outflow during the burst phase ($\epsilon_{\rm w,sb} \leq 3$) produces a long tail on the quenched side, but only a limited number of galaxies are quenched after the star burst. A significant bimodal distribution in the sSFR and a reasonable overall quenched fraction can be produced if $\epsilon_{\rm w,sb} \approx 5$ is

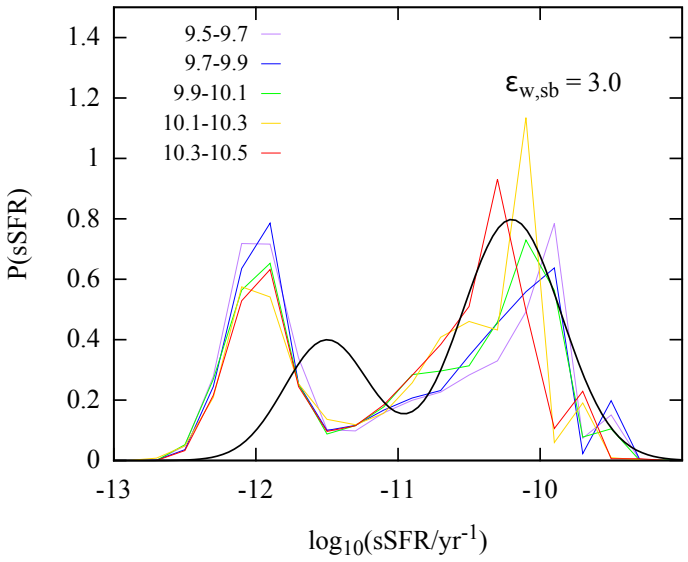

Figure 7.12. The distribution of sSFR predicted by the model in which both ram-pressure stripping and tidally-triggered starburst (with $\epsilon_{sb,w} = 3$) are included.

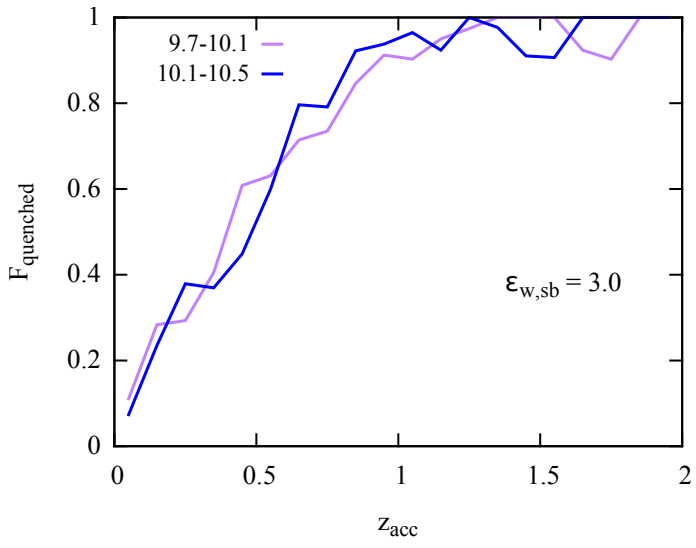

Figure 7.13. The fraction of quenched satellites as a function of the redshift of accretion, predicted by the model in which both ram-pressure stripping and tidally-triggered starburst (with $\epsilon_{sb,w} = 3$) are included.

assumed. This can be understood as follows. The time for a satellite to cross the region within $R_{\rm sb}$ is about a few 10^8yr, while the gas depletion time scale by star formation, $\tau_{\rm dep} \equiv M_{\rm g}/{\rm SFR}$, is an order of magnitude longer. Thus, star formation alone can hardly consume a significant amount of gas during the star burst. (See Di Matteo et al., 2007, who demonstrated that the extra gas consumed by star burst is not significant). In order to get a complete quenching, $\mathcal{E}_{\rm enhance}\,(1-\mathcal{R}+\epsilon_{\rm w,sb}) > 10$.

In contrast to the ram pressure stripping model, the quenched fraction predicted by the tidally-induced starburst model with $\epsilon_{\rm w,sb} = 5$ has a much weaker dependence on the host halo mass (Figure 7.10). The reason is that the size of a central galaxy scales roughly with the virial radius of the host halos (Mo et al., 1998). If the starburst (including the outflow driven by the burst) is strong enough to quench a flyby satellite in a short time scale, the dimensionless critical radius for satellite quenching $x_{\rm c} \approx \frac{R_{\rm sb}}{R_{\rm vir}}$ is nearly a constant.

7.4.4 Conspiracy between ram pressure stripping and tidally triggered starburst

The existence of both the ram pressure stripping and tidally-triggered starburst have observational supports. Thus, both processes should be taken into account in a realistic model. The two processes do not act independently. Here we demonstrate how they conspire to quench satellites in different environments. Figure 7.11 shows the predictions of two models, one with $\epsilon_{\rm w,sb} = 1$ and the other with $\epsilon_{\rm w,sb} = 3$, both including the joint effects of ram pressure stripping and tidally-triggered starburst. For comparison, the predictions of models including only ram pressure or only starburst are shown as dashed lines and dot-dashed lines, respectively. The sSFR distribution is shown in Figure 7.12. As one can see, in halos with masses below $10^{13}\,{\rm M}_\odot$, starburst with a modest outflow ($\epsilon_{\rm w,sb} = 3$) alone can hardly quench any satellites. Ram pressure stripping acting alone is also too weak to strip the gas discs in such halos, as already discussed in §7.4.2. However, when both processes are turned on, the starburst can reduce the gas surface density significantly so as to make the ram pressure stripping more effective. In contrast, in

more massive halos, ram pressure starts to work long before the starburst is triggered, so that the inclusion of starburst has only a modest effect. Overall the prediction of the combined model matches roughly the observed quenched fraction as a function of halo mass. The largest discrepancy occurs for low-mass galaxies in massive halos, where the model over-predict the quenched fraction by about 15%. For low-mass galaxies, the predicted halo mass dependence of the quenched fraction also appears to be stronger than in the data. Therefore, some halo mass-dependent prescriptions seem to be needed for the model to match the data better. For example, if the strength of tidally triggered starburst decreases with halo mass, or if the efficiency of ram pressure stripping in massive halos is lower than that given by the model, the match between model and data can be improved. The first possibility is actually likely. In halos with masses above $10^{13.5}\,\mathrm{M}_{\odot}$, ram pressure stripping starts to work long before the starburst is triggered. By the time when a satellite galaxy reaches $R_{\rm sb}$, the cold gas mass in its disk has already been reduced by a large amount, and the interaction of the gas-poorer disk may not be able to trigger a strong starburst. In the extreme case where no starburst is triggered for satellites in these massive halos, the quenched fractions would be reduced to those represented by the dashed lines in Figure 7.11, which match the data better. This possibility is not in conflict with the result of Li et al. (2008), because the result is for star-forming population which is dominated by galaxies in relatively low-mass halos.

The wind loading factor, $\epsilon_{\rm w,sb} = 3$, required for low-mass halos appears to be reasonable for starburst galaxies. Observations of galactic winds show that the mass in the outflow gas is typically a few times the SFR for starburst galaxies (e.g. Martin, 1999). Once the gas is ejected from a satellite and becomes diffuse, ram-pressure and tidal force may strip it from the sub-halo of the satellite, even if the wind is not energetic enough to escape the sub-halo by itself.

The quenched fraction is a strong function of the accretion time of the satellites (Figure 7.13). It increases rapidly from 0.5 at $z \sim 0.5$ (corresponding to a look-back time of

$\sim$ 5 Gyr) to 1 at $z \sim 1$ (corresponding to a look-back time of 7.5 Gyr). This suggests that a satellite accreted at $z > 0.5$ is more likely to have already been quenched, while a satellite accreted at a later time is more likely to be observed as a star forming-galaxy. This is consistent with the finding of Wetzel et al. (2013) with a empirical model, that on average a galaxy can remain as star forming for $\approx 4\,\mathrm{Gyr}$ after accretion before it is quenched abruptly.

7.5 Summary

We investigated the star formation quenching of satellite galaxies using a set of simplified but physically motivated models. The models are built upon the dark halo merger histories of Parkinson et al. (2008) and the sub-halo model of Taylor and Babul (2001) and Taylor and Babul (2004) that traces the the orbits, tidal stripping and disruption of sub-halos. Instead of following the whole histories of present-day satellites, we only keep track of the evolution of a satellite after it is accreted into a more massive system. The progenitors of satellites at the time of accretion, which are themselves central galaxies at that time, are initialized according to the dependences of stellar mass, star formation and cold gas content of star forming galaxies on halo mass given in previous chapters. Such dependences are constrained using current observations of the galaxy populations at different redshifts. This approach avoids some of the serious uncertainties in the assumptions made in previous models of satellite galaxies, and allows us to focus on the quenching of star formation by intra-halo environments and processes. Our model does not assume a diffuse gas reservoir in sub-halos, and as such it assumes instantaneous "strangulation".

We tested a series of models against the observational constraints derived from galaxy groups by Wetzel et al. (2012) and Wetzel et al. (2013), including the sSFR distribution, its dependence on stellar mass of the satellite and on the mass of the halo in which the satellite resides. We found that these observational constraints provide important information regarding the process of environmental quenching.

We first considered a model in which the satellites are quenched by gradual gas consumption due to star formation and associated outflow. Without any outflow, all the survived satellites would lie in the star forming sequence. This means that the cold gas reservoir brought in at the time of accretion is large enough to sustain the star formation in satellites all the way to present-day. A mass loading factor larger than unity produces an over-quenched satellite population. This critical loading factor is modest compared with what is traditionally assumed in semi-analytic models of galaxy formation. The over-quenching of satellites in these semi-analytic models has usually been attributed to the instantaneous "strangulation", and solutions to it have been focused on slowing down the stripping of the diffuse gas from sub-halos (Kang and van den Bosch, 2008; Weinmann et al., 2010). For instance, Weinmann et al. (2010) suggested that no ram pressure of diffuse gas should happen in order to reproduce the correct quench fraction, in conflict with simple physical expectations. An alternative solution to this over-quenching problem is to assume a modest outflow, as is demonstrated by our calculation and supported by other studies of central galaxies (Lu et al., 2014c; Lu et al., 2015b). Thus, we believe that the so called "over-quenching" problem is primarily due to the assumption of a high loading factor in these models. However, gradual gas consumption by star formation and outflow alone cannot produce a bimodal distribution for sSFR, suggesting that other quenching mechanisms are needed.

The second model we tested is the ram pressure stripping of the cold gas by hot halo gas. With realistic assumptions about the hot gas profiles in dark matter halos, quenching of star formation is found to occurs in an abrupt manner when satellites reach deep enough into the halo center, where the ram pressure is sufficiently large to completely strip the cold gas. To obtain an overall quenched fraction that matches observation, this model does not need any net outflow powered by star formation. Ram pressure stripping is also able to reproduce a bimodal distribution in the sSFR. However, this bimodality is largely due to the strong dependence of the ram pressure stripping on halo mass: in

massive halos ($> 10^{13.5}\,\mathrm{M}_\odot$), the satellite population is slightly over-quenched compared with observation, while in low mass groups ($< 10^{13}\,\mathrm{M}_\odot$), ram pressure is too weak to produce any quenched satellites. This is in conflict with the observational data in which a significant fraction of satellites in low-mass halos are also quenched.

In a third model, we proposed a phenomenological prescription to describe starbursts that are triggered by tidal interactions of satellites as they pass by the central galaxy in the halo center. This model is motivated by and calibrated with the observational results of Li et al. (2008). The SFR is assumed to be enhanced by a constant factor $\mathcal{E}_{\rm enhance} = 3$ once a satellite reaches a critical radius $r_{\rm sb} \sim 20R_\star$. The enhanced star formation alone cannot quench the flyby satellites, as the crossing time scale of the sphere $r_{\rm sb}$ is too short. Without ram pressure, a strong outflow with a loading factor $\epsilon_{\rm w,sb} > 5$ during the starburst phase is required. The dependence of the quenched fraction on halo mass in this model is found to be much weaker than the prediction of the ram-pressure stripping model. In the halo mass range $< 5 \times 10^{13}\,\mathrm{M}_\odot$, the prediction matches observation, while in more massive halos the critical radius is too small to produce a large enough population of quenched satellites to match observation.

Finally, we considered a model where both ram-pressure stripping and tidally-triggered starburst operate together. With ram pressure, the wind loading factor in the starburst required to match observation is reduced. In low mass halos, ram pressure and a starburst with $\epsilon_{\rm w,sb} \approx 3$ working together can produce a reasonable match to the observed quenched fraction. In high mass halos, quenching is always dominated by ram pressure.

In conclusion, our investigation demonstrates that, once the progenitors of satellites are properly modeled, the star formation properties of the satellite population can be understood in terms of physical processes that are expected to operate in galaxy systems. Further investigations are clearly needed. The assumptions regarding ram pressure stripping and tidally induced starburst need to be tested with detailed numerical simulations. Detailed model predictions can be made by controlled simulations which follow

the evolutions of satellites with properly-set initial conditions in different halos. Further observational consequences, such as spatial distributions and velocity dispersions of quenched and non-quenched satellites in their host dark matter halos, should be explored to test the model in more detail. We will come back to some of these problems in the future.

CHAPTER 8

CONCLUSIONS AND OUTLOOK

8.1 Conclusions

8.1.1 The star formation history across cosmic time

We develop an empirical approach to infer the SFR in dark matter halos from the galaxy SMF at different redshifts and the local CGLF, which has a steeper faint end relative to the SMF of local galaxies. Since satellites are typically old galaxies which have been accreted earlier, this feature can cast important constraint on the formation of low-mass galaxies at high-redshift. The evolution of the SMFs suggests the star formation in high mass halos ($> 10^{12}\,h^{-1}\mathrm{M}_{\odot}$) has to be boosted at high redshift beyond what is expected from a simple scaling of the dynamical time. The faint end of the CGLF implies a characteristic redshift $z_c \approx 2$ above which the SFR in low mass halos with masses $< 10^{11}\,h^{-1}\mathrm{M}_{\odot}$ must be enhanced relative to that at lower z. This is not directly expected from the standard stellar feedback models.

The model predicts that the slope of galaxy SMFs at $z > 2$ should be quite steep at the low mass end, beyond the current detection limit. At $z > 4$, the star formation rate density is mainly contributed by dwarf galaxies with mass $\sim 10^8\,\mathrm{M}_{\odot}$. It also implies that the low-mass end upturn is ubiquitous in the satellite population of halos of different masses. These are strong predictions that can be directly tested against future observations.

We characterize the galaxy star formation and assembly or merger histories in details. Most of the stars in cluster centrals, corresponding to BCGs in observations, formed earlier than $z \approx 2$ but have been assembled much later. Typically, they have experienced ≈ 5 major mergers since their star formation was quenched. Milky Way mass galaxies

have had on-going star formation without significant mergers since $z \approx 2$, and are thus free of significant (classic) bulges produced by major mergers. Dwarf galaxies in halos with $M_{\rm h} < 10^{11} h^{-1} M_{\odot}$ or $M_{\star} < 10^{9} M_{\odot}$ have experienced a star formation burst at $z > 2$, followed by a nearly constant SFR after $z = 1$, and the stellar age decreases with stellar mass, contrary to the 'downsizing' trend for more massive galaxies. Major mergers are not uncommon during the early burst phase and may result in the formation of old spheroids in dwarf galaxies. We also characterise the stellar population of halo stars in different halos.

8.1.2 Reverse Engineering

The relation between SFR and the host halo mass and the its evolution with time serve as basis to infer the physics that regulates galaxy evolution.

The assumption that the SMBH mass ($M_{\rm BH}$) is directly proportionalto the classical bulge mass, with the proportionality given by that for massive galaxies, predicts a $M_{\rm BH}$ - $M_{\star}$ relation that matches well the observed relation for different types of galaxies. In particular, the model reproduces the strong transition at $M_{\star} = 10^{10.5}$ - $10^{11}\,{\rm M}_{\odot}$, below which $M_{\rm BH}$ drops rapidly with decreasing $M_{\star}$. Our model predicts a new sequence at $M_{\star} < 10^{10.5}\,{\rm M}_{\odot}$, where $M_{\rm BH} \propto M_{\star}$ but the amplitude is a factor of ~ 50 lower than the amplitude of the sequence at $M_{\star} > 10^{11}\,{\rm M}_{\odot}$. If all SMBH grow through similar quasar modes with a feedback efficiency of a few percent, the energy produced in low-mass galaxies at redshift $z \gtrsim 2$ can heat the circum-galactic medium up to a specific entropy level that is required to prevent excessive star formation in low-mass dark matter halos.

A set of observational data for galaxy cold gas mass fraction and gas phase metallicity is used to constrain the content, inflow and outflow of gas in central galaxies hosted by halos with masses between $10^{11} M_{\odot}$ to $10^{12} M_{\odot}$. The gas contents in high redshift galaxies are obtained from the empirical star formation histories and the star formation laws . The total baryon mass, that is the stellar mass and cold gas mass combined, in low-mass galaxies is always much less than expected from the universal baryon mass fraction since

$z = 2$. The data for the evolution of the gas phase metallicity require net metal outflow at $z \lesssim 2$, and the metal loading factor is constrained to be about 0.01, or about 60% of the metal yield. Based on the assumption that galactic outflow is more enriched in metal than both the interstellar medium and the material ejected at earlier epochs, we are able to put stringent constraints on the upper limits for both the net accretion rate and the net mass outflow rate. It is shown that low-mass galaxies at $z \approx 0$ is not compatible with strong outflow. We speculate that the low star formation efficiency of low-mass galaxies is owing to some preventative processes that prevent gas from accreting into galaxies in the first place.

The empirical relation between stellar mass, SFR and halo mass is used initialize the progenitors of present-day satellites. With this initial condition, we study the quenching of satellite galaxies by gradual depletion of gas due to star formation, by ram-pressure striping and by tidally triggered starburst. Eevn if there is no further cold gas supply from the cooling of the halo gas after a galaxy is accreted by its host, the cold gas remaining in the satellites is sufficient to sustain the star formation until present day. Gradual depletion alone predicts a unimodal distribution in specific star formation, in contrast to the bimodal distribution observed, and under-predicts the quenched fraction in low mass halos. Ram-pressure stripping nicely reproduces the bimodal distribution but under-predicts the quenched fraction in low-mass halos. Starbursts in gas-rich satellites triggered by tidal interactions with central galaxies can nicely reproduce the quenched satellite population in low-mass halos, but become unimportant for low-mass satellites in massive halos. The combined processes, together with the constrained progenitors, can reproduce the observed star formation properties of satellites in halos of different masses.

8.2 Prospective Science Projects

8.2.1 A Generic Model of Galaxy-Halo Connection

The bimodal distribution suggests that the star formation must be quenched in an abrupt manner and its ubiquitousness indicates that such quenching processes occur in halos of different masses and in different environments. Given that the hierarchical formation of dark matter halos is the backbone of galaxy formation, understanding the physics of the quenching processes requires a solid understanding of the how the they are related to the properties of the dark matter halos. The bimodality of the galaxy population has been overlooked in the empirical model presented in this paper, and many other studies of galaxy-halo connections. Such models can only be considered as first order approximation since i) they ignore the bimodal distribution in the SFR; ii) it is assumed that it is the halo mass that determines a galaxy's property (SFR and $M_\star$).

To understand the "dark side" of quenching of star formation, a more generic model of galaxy-halo connection will be developed based on previous studies. Such a generic model can be written as a conditional distribution

$$\mathcal{P}\left(\mathrm{SFR}, M_{\mathrm{star}}, z|\mathcal{H}\right) \tag{8.1}$$

where $\mathcal{H}$ represents a vector of halo properties, including halo mass, maximum circular velocity and assmebly history, etc. The model translates the observational constraints into a set of physically meaningful quantities, for example, the life time of star forming galaxies and how it depends on the properties of the host halos, such as the environment, mass and formation time.

To get the exact form of Eq 8.1 is a difficult task. The Bayesian approach adopted in this paper can be used construct empirical models like Eq 8.1. The approach is based on the improved nested sampler, which is used to search the high dimensional parameter spaces, and Bayesian evidence for model selection. However, as the observational constraints accumulate and become more accurate, developing such empirical models with

efficiency and statistical rigorousness is even more challenging. The model complexity increases dramatically with amount and accuracy of the data and can be hard to write down as a combination of elementary functions. In this case, more efficient sampling tools as well as non-parametric approaches need to be developed.

8.2.2 Modelling the Evolution of Dwarf Galaxies

It is generally believed that monolithic collapse, major mergers and BH formation and feedback are only associated with the progenitors of the present-day massive elliptical galaxies and are the major reason for the existence of this quiescent and homegeneous population. However, our findings suggest that these processes are also relevant to the dwarf galaxies in the early Universe. The formation of dwarf galaxies ($10^8\,\mathrm{M}_\odot$ in stars) starts with a strong star burst before $z \approx 2$, which makes up a significant fraction of the total stellar mass at present day. Also, major mergers occurred in this period. If central black holes (BH) could form and grow through gas rich mergers, the dwarf galaxies are supposed to host BHs of $10^6\,\mathrm{M}_\odot$ by $z \approx 2$.

Based on such findings, analytical and numerical models are to be built to investigate how such early violent formation phase would shape the population of present day dwarf galaxies, which remains a mystery in many aspects.

One of the major problem concerning dwarf galaxies is that their stellar mass to dark halo mass ratio are orders of magnitude lower than unity, The early star burst and BH formation provide a promising solution to this problem. If output energy is dumped into the surrouding IGM, the subsequent accretion will be slowed down and the star formation will be delayed. However, to understand how the feedback will affect the subsequent star formation until present day, many questions need to be answered: How does the energy couple with the surrouding gas? If the AGN feedback is in quasar mode, is dust absorption an efficent channel in dwarf galaxies that are poor in metals? Can the energy propogate into the IGM? Is it subject to efficent radiative loss on the way since the dwarf galaxies lie in the fast cooling regime?

Another outstanding problem with dwarf galaxies is the cusp/core problem: halo density profiles predicted by the CDM models are more cuspy than what has been found. Different theories have been proposed, which suggest that energy feedback from stars are responsible for the core formation (Mo and Mao, 2004; Pontzen and Governato, 2012). The theories can be used to investigate whether the early violent formation phase would leave an imprint in the structure of the host dark halos. Also dwarf galaxies are considerably diverse in morphology compared to massive spirals and ellipticals. Besides the common dwarf ellipticals and dwarf spheroids, a population of compact dwarfs (Norris et al., 2015), with surface density orders of magnitude larger, have been detected. The early major mergers can probably shed some light on the origin of such compact systems, as the mergers are gas rich and could remove large amount of energy and angular momentum from the systems.

8.2.2.1 Star Formation History in Hydrodynamic Simulations

Galaxy evolution is a multi-scale problem. In cosmological hydrodynamic simulations small scale processes, such as the interaction between stars and ISM, cannot be resolved and is usually controlled by sub-grid prescriptions. But processes on large scales come out naturally as a result of relevant dynamics of the gas and dark matter. Therefore, the connection between the statistics of the simulated galaxy population and the input prescriptions is not straightforward. For instance, the mass loading factor in Oppenheimer et al. (2010) scales with the host halo mass as

$$\eta = 1.7 \left(\frac{M_{\rm h}}{10^{12}\,{\rm M}_{\odot}} \right)^{-0.32} \tag{8.2}$$

The dependence on halo mass is much weaker than the mass loading inferred from semi-analytic models (Lu et al., 2014a) or the stellar mass - halo mass relation (§3). As shown in Oppenheimer et al. (2010), this is because recycling of ejected gas depends on the potential well of the host halos. To understand how the sub-grid physics is connected

to the statistics of the modeled galaxies, it is necessary to investigate the star formation histories of modeled galaxies and compare with the prediction of the empirical models.

APPENDIX

MULTINEST SAMPLING OF THE POSTERIOR DISTRIBUTION

The main goal of nested sampling is to evaluate the Bayesian evidence, and to provide samples of the posterior distribution. Briefly, the algorithm works as follows. At the beginning of the process, one randomly draws N points in parameter space from the adopted prior distribution. These points are called the active set. Each of the points has a likelihood value L_i $(i = 0, ..., N-1)$, and associated with it is an isolikelihood surface defined by the value of L_i. The volume (modulated by the prior distribution) within the surface is X_i. The point with the lowest likelihood value is denoted by $L^{(0)}$ and the corresponding prior volume, $X^{(0)}$, can be approximated by the total volume of the prior space. This point is removed from the active set and is added to another list called the inactive set. A new point with likelihood bigger than $L^{(0)}$ is then drawn from the prior distribution and is added to the active set. Before going to the next iteration, it is important to know the volume (again modulated by the prior distribution) occupied by the new active set. Note that the ratios, $t_i \equiv X_i/X^{(0)}$ $(i = 0, ..., N-1)$, can be considered as N random numbers drawn from the uniform distribution within $[0, 1)$.

Define $t^{(1)} \equiv \max(t_i)$ and the corresponding likelihood is $L^{(1)}$. Then the volume occupied by the new active set is simply $X^{(1)} = t^{(1)}X^{(0)}$. The exact value of $t^{(1)}$ is unknown but it must satisfy the following distribution:

$$p(t) = Nt^{N-1}, \tag{A.1}$$

with the expectation of $\ln(t)$ equal to $-1/N$ and the standard deviation in $\ln(t)$ equal to $1/N$. Thus $t^{(1)}$ may be approximated by the expectation value, $\exp(-1/N)$. The first

step, therefore, ends with a new active set that occupies a total volume that is smaller by a factor of $t^{(1)} \approx \exp(-1/N)$ than the old set, and with a new member, $\left(L^{(0)}, X^{(0)}\right)$, in the inactive set.

By repeating the above process for the new active set produced at each subsequent step, a list of points are drawn from the posterior distribution:

$$\{(L^{(k)}, X^{(k)})\} \qquad \text{with} \quad X^{(k+1)} = t^{(k+1)} X^{(k)}, \tag{A.2}$$

which defines a series of nested shells in the parameter space and can be used to sample the posterior distribution. As described above, the exact value of $t^{(k+1)}$ is unknown but can be approximated by $\exp(-1/N)$ and the uncertainty can also be quantified with the use of Equation (A.1). The Bayesian evidence is simply

$$Z = \sum_k L^{(k)} \frac{[X^{(k+1)} - X^{(k-1)}]}{2} \,. \tag{A.3}$$

One iterates until the value of Z reaches a chosen accuracy, and we choose it to be 0.5 in natural logarithmic scale.

The efficiency of this algorithm depends on how efficiently the active set can be replenished at each iteration. Drawing new samples blindly from the prior leads to a lower and lower acceptance rate as the iso-likelihood surface shrinks. However, if the surface can be approximated by some regular shapes, then the active set can be efficiently replenished. The MULTINEST package developed by Feroz et al. (2009) provides such a method. At each iteration, multiple ellipsoids are used to approximate the iso-likelihood surface of the new point drawn. An ellipsoid can either overlap with others or be isolated. Too few big ellipsoids would result in a bad approximation, while too many small ellipsoids would result in too much overlap. In both cases the acceptance rate would be low. Optimal ellipsoidal decomposition is found by minimising

$$F = \frac{\sum_j V(E_j)}{V(S)} \geq 1 \,, \tag{A.4}$$

where $V(S)$ is the volume within the iso-likelihood surface and $\sum_j V(E_j)$ is the total volume of all the ellipsoids. The acceptance rate is simply the inverse of F.

The parameter that controls the process of posterior exploration is the size of the active set. A large active set can slow down the speed of going uphill because after each iteration the size of the volume enclosed by the iso-likelihood surface shrinks by $\exp(-1/N)$. Conversely, to detect all the modes that are statistically significant in a high dimensional parameter space the size of the active set cannot be too small. Unfortunately, there is no good way to find the optimal active set size. For this work, we use $2,000$ active points for Model I, Model II, and Model IIb, $5,000$ for Model III and $10,000$ for Model IV. We arrived at these values empirically by increasing the active set size until the value of the marginal likelihood did not change appreciably.

BIBLIOGRAPHY

[1] M. Asplund, N. Grevesse, A. J. Sauval, et al. "The Chemical Composition of the Sun". In: ARA&A 47 (Sept. 2009), pp. 481–522.

[2] E. Bañados, L.-W. Hung, R. De Propris, et al. "The Faint End of the Galaxy Luminosity Function in A1689: A Steep Red Faint End Upturn at z = 0.18". In: ApJL 721 (Sept. 2010), pp. L14–L18.

[3] I. K. Baldry, S. P. Driver, J. Loveday, et al. "Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at z $<$ 0.06". In: MNRAS 421 (Mar. 2012), pp. 621–634.

[4] M. L. Balogh, J. F. Navarro, and S. L. Morris. "The Origin of Star Formation Gradients in Rich Galaxy Clusters". In: ApJ 540 (Sept. 2000), pp. 113–121.

[5] W. A. Barkhouse, H. K. C. Yee, and O. López-Cruz. "The Luminosity Function of Low-Redshift Abell Galaxy Clusters". In: ApJ 671 (Dec. 2007), pp. 1471–1496.

[6] J. Barnes and G. Efstathiou. "Angular momentum from tidal torques". In: ApJ 319 (Aug. 1987), pp. 575–600.

[7] A. J. Barth, L. C. Ho, R. E. Rutledge, et al. "POX 52: A Dwarf Seyfert 1 Galaxy with an Intermediate-Mass Black Hole". In: ApJ 607 (May 2004), pp. 90–102.

[8] P. S. Behroozi, C. Conroy, and R. H. Wechsler. "A Comprehensive Analysis of Uncertainties Affecting the Stellar Mass-Halo Mass Relation for 0 $<$ z $<$ 4". In: ApJ 717 (July 2010), pp. 379–403.

[9] P. S. Behroozi, R. H. Wechsler, and C. Conroy. "On the Lack of Evolution in Galaxy Star Formation Efficiency". In: ApJL 762, L31 (Jan. 2013), p. L31.

[10] P. S. Behroozi, R. H. Wechsler, and C. Conroy. "The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8". In: ApJ 770, 57 (June 2013), p. 57.

[11] A. A. Berlind and D. H. Weinberg. "The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass". In: ApJ 575 (Aug. 2002), pp. 587–616.

[12] E. J. Bernard, A. M. N. Ferguson, J. C. Richardson, et al. “The nature and origin of substructure in the outskirts of M31 - II. Detailed star formation histories”. In: MNRAS 446 (Jan. 2015), pp. 2789–2801.

[13] M. Bernardi, A. Meert, R. K. Sheth, et al. “The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile”. In: MNRAS 436 (Nov. 2013), pp. 697–704.

[14] M. Béthermin, L. Wang, O. Doré, et al. “The redshift evolution of the distribution of star formation among dark matter halos as seen in the infrared”. In: A&A 557, A66 (Sept. 2013), A66.

[15] R. Bezanson, P. G. van Dokkum, T. Tal, et al. “The Relation Between Compact, Quiescent High-redshift Galaxies and Massive Nearby Elliptical Galaxies: Evidence for Hierarchical, Inside-Out Growth”. In: ApJ 697 (June 2009), pp. 1290–1298.

[16] F. Bigiel, A. Leroy, F. Walter, et al. “The Star Formation Law in Nearby Galaxies on Sub-Kpc Scales”. In: AJ 136 (Dec. 2008), pp. 2846–2871.

[17] Y. Birnboim and A. Dekel. “Virial shocks in galactic haloes?” In: MNRAS 345 (Oct. 2003), pp. 349–364.

[18] M. R. Blanton, R. H. Lupton, D. J. Schlegel, et al. “The Properties and Luminosity Function of Extremely Low Luminosity Galaxies”. In: ApJ 631 (Sept. 2005), pp. 208–230.

[19] J. S. Bolton and M. G. Haehnelt. “The observed ionization rate of the intergalactic medium and the ionizing emissivity at z $>=$ 5: evidence for a photon-starved and extended epoch of reionization”. In: MNRAS 382 (Nov. 2007), pp. 325–341.

[20] J. R. Bond, S. Cole, G. Efstathiou, et al. “Excursion set mass functions for hierarchical Gaussian fluctuations”. In: ApJ 379 (Oct. 1991), pp. 440–460.

[21] A. Boselli, L. Cortese, M. Boquien, et al. “Cold gas properties of the Herschel Reference Survey. II. Molecular and total gas scaling relations”. In: A&A 564, A66 (Apr. 2014), A66.

[22] N. Bouché, A. Dekel, R. Genzel, et al. “The Impact of Cold Gas Accretion Above a Mass Floor on Galaxy Scaling Relations”. In: ApJ 718 (Aug. 2010), pp. 1001–1018.

[23] F. Bournaud, A. Dekel, R. Teyssier, et al. “Black Hole Growth and Active Galactic Nuclei Obscuration by Instability-driven Inflows in High-redshift Disk Galaxies Fed by Cold Streams”. In: ApJL 741, L33 (Nov. 2011), p. L33.

[24] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, et al. “UV-continuum Slopes at z ˜ 4-7 from the HUDF09+ERS+CANDELS Observations: Discovery of a Well-defined UV Color-Magnitude Relationship for z > = 4 Star-forming Galaxies”. In: ApJ 754, 83 (Aug. 2012), p. 83.

[25] R. G. Bower, A. J. Benson, R. Malbon, et al. “Breaking the hierarchy of galaxy formation”. In: MNRAS 370 (Aug. 2006), pp. 645–655.

[26] M. Boylan-Kolchin, C.-P. Ma, and E. Quataert. “Dynamical friction and galaxy merging time-scales”. In: MNRAS 383 (Jan. 2008), pp. 93–101.

[27] L. D. Bradley, M. Trenti, P. A. Oesch, et al. “The Brightest of Reionizing Galaxies Survey: Constraints on the Bright End of the z ˜ 8 Luminosity Function”. In: ApJ 760, 108 (Dec. 2012), p. 108.

[28] G. Bruzual and S. Charlot. “Stellar population synthesis at the resolution of 2003”. In: MNRAS 344 (Oct. 2003), pp. 1000–1028.

[29] J. S. Bullock, T. S. Kolatt, Y. Sigad, et al. “Profiles of dark haloes: evolution, scatter and environment”. In: MNRAS 321 (Mar. 2001), pp. 559–575.

[30] M. Cacciato, F. C. van den Bosch, S. More, et al. “Cosmological constraints from a combination of galaxy clustering and lensing - III. Application to SDSS data”. In: MNRAS 430 (Apr. 2013), pp. 767–786.

[31] S. Cantalupo. “Stars quenching stars: how photoionization by local sources regulates gas cooling and galaxy formation”. In: MNRAS 403 (Mar. 2010), pp. L16–L20.

[32] G. Chabrier. “Galactic Stellar and Substellar Initial Mass Function”. In: PASP 115 (July 2003), pp. 763–795.

[33] P. Chang, A. E. Broderick, and C. Pfrommer. “The Cosmological Impact of Luminous TeV Blazars. II. Rewriting the Thermal History of the Intergalactic Medium”. In: ApJ 752, 23 (June 2012), p. 23.

[34] S. Cole. “Modeling galaxy formation in evolving dark matter halos”. In: ApJ 367 (Jan. 1991), pp. 45–53.

[35] S. Cole, J. Helly, C. S. Frenk, et al. “The statistical properties of Λ cold dark matter halo formation”. In: MNRAS 383 (Jan. 2008), pp. 546–556.

[36] C. Conroy and R. H. Wechsler. “Connecting Galaxies, Halos, and Star Formation Rates Across Cosmic Time”. In: ApJ 696 (May 2009), pp. 620–635.

[37] C. Conroy, R. H. Wechsler, and A. V. Kravtsov. "Modeling Luminosity-dependent Galaxy Clustering through Cosmic Time". In: ApJ 647 (Aug. 2006), pp. 201–214.

[38] D. J. Croton, V. Springel, S. D. M. White, et al. "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies". In: MNRAS 365 (Jan. 2006), pp. 11–28.

[39] E. Daddi, M. Dickinson, G. Morrison, et al. "Multiwavelength Study of Massive Galaxies at z˜2. I. Star Formation and Galaxy Growth". In: ApJ 670 (Nov. 2007), pp. 156–172.

[40] J. J. Dalcanton. "The Metallicity of Galaxy Disks: Infall versus Outflow". In: ApJ 658 (Apr. 2007), pp. 941–959.

[41] R. Davé, K. Finlator, and B. D. Oppenheimer. "An analytic model for the evolution of the stellar, gas and metal content of galaxies". In: MNRAS 421 (Mar. 2012), pp. 98–107.

[42] A. Dekel and J. Silk. "The origin of dwarf galaxies, cold dark matter, and biased galaxy formation". In: ApJ 303 (Apr. 1986), pp. 39–55.

[43] P. Di Matteo, F. Combes, A.-L. Melchior, et al. "Star formation efficiency in galaxy interactions and mergers: a statistical study". In: A&A 468 (June 2007), pp. 61–81.

[44] T. Di Matteo, N. Khandai, C. DeGraf, et al. "Cold Flows and the First Quasars". In: ApJL 745, L29 (Feb. 2012), p. L29.

[45] T. Di Matteo, J. Colberg, V. Springel, et al. "Direct Cosmological Simulations of the Growth of Black Holes and Galaxies". In: ApJ 676 (Mar. 2008), pp. 33–53.

[46] N. Drory, K. Bundy, A. Leauthaud, et al. "The Bimodal Galaxy Stellar Mass Function in the COSMOS Survey to z ˜ 1: A Steep Faint End and a New Galaxy Dichotomy". In: ApJ 707 (Dec. 2009), pp. 1595–1609.

[47] Y. Dubois, M. Volonteri, J. Silk, et al. "Black hole evolution - II. Spinning black holes in a supernova-driven turbulent interstellar medium". In: MNRAS 440 (May 2014), pp. 2333–2346.

[48] Y. Dubois, C. Pichon, M. Haehnelt, et al. "Feeding compact bulges and supermassive black holes with low angular momentum cosmic gas at high redshift". In: MNRAS 423 (July 2012), pp. 3616–3630.

[49] A. A. Dutton and A. V. Macciò. “Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles”. In: MNRAS 441 (July 2014), pp. 3359–3374.

[50] A. A. Dutton, F. C. van den Bosch, S. M. Faber, et al. “On the evolution of the velocity-mass-size relations of disc-dominated galaxies over the past 10 billion years”. In: MNRAS 410 (Jan. 2011), pp. 1660–1676.

[51] A. A. Dutton, F. C. van den Bosch, and A. Dekel. “On the origin of the galaxy star-formation-rate sequence: evolution and scatter”. In: MNRAS 405 (July 2010), pp. 1690–1710.

[52] D. K. Erb. “A Model for Star Formation, Gas Flows, and Chemical Evolution in Galaxies at High Redshifts”. In: ApJ 674 (Feb. 2008), pp. 151–156.

[53] D. K. Erb, A. E. Shapley, M. Pettini, et al. “The Mass-Metallicity Relation at z > ˜2”. In: ApJ 644 (June 2006), pp. 813–828.

[54] F. Feroz, M. P. Hobson, and M. Bridges. “MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics”. In: MNRAS 398 (Oct. 2009), pp. 1601–1614.

[55] A. V. Filippenko and L. C. Ho. “A Low-Mass Central Black Hole in the Bulgeless Seyfert 1 Galaxy NGC 4395”. In: ApJL 588 (May 2003), pp. L13–L16.

[56] A. S. Font, R. G. Bower, I. G. McCarthy, et al. “The colours of satellite galaxies in groups and clusters”. In: MNRAS 389 (Oct. 2008), pp. 1619–1629.

[57] F. Fontanot, G. De Lucia, P. Monaco, et al. “The many manifestations of downsizing: hierarchical galaxy formation models confront observations”. In: MNRAS 397 (Aug. 2009), pp. 1776–1790.

[58] M. Fukugita, C. J. Hogan, and P. J. E. Peebles. “The Cosmic Baryon Budget”. In: ApJ 503 (Aug. 1998), pp. 518–530.

[59] A. Gallazzi, S. Charlot, J. Brinchmann, et al. “The ages and metallicities of galaxies in the local universe”. In: MNRAS 362 (Sept. 2005), pp. 41–58.

[60] M. Geha, P. Guhathakurta, and R. P. van der Marel. “Internal Dynamics, Structure, and Formation of Dwarf Elliptical Galaxies. II. Rotating versus Nonrotating Dwarfs”. In: AJ 126 (Oct. 2003), pp. 1794–1810.

[61] N. Y. Gnedin. “Effect of Reionization on Structure Formation in the Universe”. In: ApJ 542 (Oct. 2000), pp. 535–541.

[62] N. Y. Gnedin and J. P. Ostriker. “Reionization of the Universe and the Early Production of Metals”. In: ApJ 486 (Sept. 1997), pp. 581–598.

[63] A. W. Graham. “Breaking the Law: The M $_{bh}$-M $_{spheroid}$ Relations for Core-Sérsic and Sérsic Galaxies”. In: ApJ 746, 113 (Feb. 2012), p. 113.

[64] A. W. Graham and N. Scott. “The M $_{BH}$-L $_{spheroid}$ Relation at High and Low Masses, the Quadratic Growth of Black Holes, and Intermediate-mass Black Hole Candidates”. In: ApJ 764, 151 (Feb. 2013), p. 151.

[65] J. E. Greene and L. C. Ho. “A New Sample of Low-Mass Black Holes in Active Galaxies”. In: ApJ 670 (Nov. 2007), pp. 92–104.

[66] J. E. Greene and L. C. Ho. “Active Galactic Nuclei with Candidate Intermediate-Mass Black Holes”. In: ApJ 610 (Aug. 2004), pp. 722–736.

[67] J. E. Greene, L. C. Ho, and A. J. Barth. “Black Holes in Pseudobulges and Spheroidals: A Change in the Black Hole-Bulge Scaling Relations at Low Mass”. In: ApJ 688 (Nov. 2008), pp. 159–179.

[68] Q. Guo, S. White, M. Boylan-Kolchin, et al. “From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology”. In: MNRAS 413 (May 2011), pp. 101–131.

[69] Q. Guo, S. White, C. Li, et al. “How do galaxies populate dark matter haloes?” In: MNRAS 404 (May 2010), pp. 1111–1120.

[70] N. Häring and H.-W. Rix. “On the Black Hole Mass-Bulge Mass Relation”. In: ApJL 604 (Apr. 2004), pp. L89–L92.

[71] Y. Q. He, X. Y. Xia, C. N. Hao, et al. “Photometric Properties and Luminosity Function of Nearby Massive Early-type Galaxies”. In: ApJ 773, 37 (Aug. 2013), p. 37.

[72] A. P. Hearin and D. F. Watson. “The dark side of galaxy colour”. In: MNRAS 435 (Oct. 2013), pp. 1313–1324.

[73] A. Henry, C. L. Martin, K. Finlator, et al. “The Metallicity Evolution of Low-mass Galaxies: New Constraints at Intermediate Redshift”. In: ApJ 769, 148 (June 2013), p. 148.

[74] M. Hilz, T. Naab, and J. P. Ostriker. “How do minor mergers promote inside-out growth of ellipticals, transforming the size, density profile and dark matter fraction?” In: MNRAS 429 (Mar. 2013), pp. 2924–2933.

[75] S. Ikeuchi. “The baryon clump within an extended dark matter region”. In: Ap&SS 118 (Jan. 1986), pp. 509–514.

[76] L. P. Jenkins, A. E. Hornschemeier, B. Mobasher, et al. “Uncovering the Near-IR Dwarf Galaxy Population of the Coma Cluster with Spitzer IRAC”. In: ApJ 666 (Sept. 2007), pp. 846–862.

[77] F. Jiang and F. C. van den Bosch. “Generating merger trees for dark matter haloes: a comparison of methods”. In: MNRAS 440 (May 2014), pp. 193–207.

[78] Y.-F. Jiang, J. E. Greene, and L. C. Ho. “Black Hole Mass and Bulge Luminosity for Low-mass Black Holes”. In: ApJL 737, L45 (Aug. 2011), p. L45.

[79] Y.-F. Jiang, J. E. Greene, L. C. Ho, et al. “The Host Galaxies of Low-mass Black Holes”. In: ApJ 742, 68 (Dec. 2011), p. 68.

[80] Jimmy, K.-V. Tran, S. Brough, et al. “Angular Momenta, Dynamical Masses, and Mergers of Brightest Cluster Galaxies”. In: ApJ 778, 171 (Dec. 2013), p. 171.

[81] Y. P. Jing, H. J. Mo, and G. Börner. “Spatial Correlation Function and Pairwise Velocity Dispersion of Galaxies: Cold Dark Matter Models versus the Las Campanas Survey”. In: ApJ 494 (Feb. 1998), pp. 1–12.

[82] X. Kang and F. C. van den Bosch. “New Constraints on the Efficiencies of Ram Pressure Stripping and the Tidal Disruption of Satellite Galaxies”. In: ApJL 676 (Apr. 2008), pp. L101–L104.

[83] R. Kannan, G. S. Stinson, A. V. Macciò, et al. “Galaxy formation with local photoionization feedback - I. Methods”. In: MNRAS 437 (Jan. 2014), pp. 2882–2893.

[84] N. Katz. “Dissipational galaxy formation. II - Effects of star formation”. In: ApJ 391 (June 1992), pp. 502–517.

[85] N. Katz and J. E. Gunn. “Dissipational galaxy formation. I - Effects of gasdynamics”. In: ApJ 377 (Aug. 1991), pp. 365–381.

[86] N. Katz, L. Hernquist, and D. H. Weinberg. “Galaxies and gas in a cold dark matter universe”. In: ApJL 399 (Nov. 1992), pp. L109–L112.

[87] R. C. Kennicutt Jr. “The Global Schmidt Law in Star-forming Galaxies”. In: ApJ 498 (May 1998), pp. 541–552.

[88] D. Kereš, N. Katz, M. Fardal, et al. “Galaxies in a simulated ΛCDM Universe - I. Cold mode and hot cores”. In: MNRAS 395 (May 2009), pp. 160–179.

[89] D. Kereš, N. Katz, D. H. Weinberg, et al. “How do galaxies get their gas?” In: MNRAS 363 (Oct. 2005), pp. 2–28.

[90] L. J. Kewley and S. L. Ellison. “Metallicity Calibrations and the Mass-Metallicity Relation for Star-forming Galaxies”. In: ApJ 681 (July 2008), pp. 1183–1204.

[91] C. Kobayashi, H. Umeda, K. Nomoto, et al. “Galactic Chemical Evolution: Carbon through Zinc”. In: ApJ 653 (Dec. 2006), pp. 1145–1171.

[92] E. Komatsu, K. M. Smith, J. Dunkley, et al. “Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation”. In: ApJS 192, 18 (Feb. 2011), p. 18.

[93] J. Kormendy and L. C. Ho. “Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies”. In: ARA&A 51 (Aug. 2013), pp. 511–653.

[94] J. Kormendy and D. Richstone. “Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei”. In: ARA&A 33 (1995), p. 581.

[95] J. Kormendy, R. Bender, and M. E. Cornell. “Supermassive black holes do not correlate with galaxy disks or pseudobulges”. In: Nature 469 (Jan. 2011), pp. 374–376.

[96] A. V. Kravtsov. “The Size-Virial Radius Relation of Galaxies”. In: ApJL 764, L31 (Feb. 2013), p. L31.

[97] M. R. Krumholz and A. Dekel. “Metallicity-dependent Quenching of Star Formation at High Redshift in Small Galaxies”. In: ApJ 753, 16 (July 2012), p. 16.

[98] M. R. Krumholz, C. F. McKee, and J. Tumlinson. “The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Photodissociation Fronts in Finite Clouds”. In: ApJ 689 (Dec. 2008), pp. 865–882.

[99] M. R. Krumholz, C. F. McKee, and J. Tumlinson. “The Atomic-to-Molecular Transition in Galaxies. II: H I and H_2 Column Densities”. In: ApJ 693 (Mar. 2009), pp. 216–235.

[100] A. Leauthaud, J. Tinker, K. Bundy, et al. “New Constraints on the Evolution of the Stellar-to-dark Matter Connection: A Combined Analysis of Galaxy-Galaxy Lensing, Clustering, and Stellar Mass Functions from z = 0.2 to z =1”. In: ApJ 744, 159 (Jan. 2012), p. 159.

[101] S. N. Leitner. “On the Last 10 Billion Years of Stellar Mass Growth in Star-forming Galaxies”. In: ApJ 745, 149 (Feb. 2012), p. 149.

[102] C. Li, G. Kauffmann, T. M. Heckman, et al. “Interaction-induced star formation in a complete sample of 10^5 nearby star-forming galaxies”. In: MNRAS 385 (Apr. 2008), pp. 1903–1914.

[103] S. J. Lilly, C. M. Carollo, A. Pipino, et al. “Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos”. In: ApJ 772, 119 (Aug. 2013), p. 119.

[104] J. Loveday, P. Norberg, I. K. Baldry, et al. “Galaxy and Mass Assembly (GAMA): ugriz galaxy luminosity functions”. In: MNRAS 420 (Feb. 2012), pp. 1239–1262.

[105] Y. Lu and H. J. Mo. “The accretion and cooling of pre-heated gas in dark matter haloes”. In: MNRAS 377 (May 2007), pp. 617–629.

[106] Y. Lu, H. J. Mo, M. D. Weinberg, et al. “A Bayesian approach to the semi-analytic model of galaxy formation: methodology”. In: MNRAS 416 (Sept. 2011), pp. 1949–1964.

[107] Y. Lu, H. J. Mo, Z. Lu, et al. “Bayesian inferences of galaxy formation from the K-band luminosity and H I mass functions of galaxies: constraining star formation and feedback”. In: MNRAS 443 (Sept. 2014), pp. 1252–1266.

[108] Y. Lu, H. J. Mo, and R. H. Wechsler. “Formation of disc galaxies in preheated media: a preventative feedback model”. In: MNRAS 446 (Jan. 2015), pp. 1907–1923.

[109] Y. Lu, D. Kereš, N. Katz, et al. “On the algorithms of radiative cooling in semi-analytic models”. In: MNRAS 416 (Sept. 2011), pp. 660–679.

[110] Y. Lu, R. H. Wechsler, R. S. Somerville, et al. “Semi-analytic Models for the CANDELS Survey: Comparison of Predictions for Intrinsic Galaxy Properties”. In: ApJ 795, 123 (Nov. 2014), p. 123.

[111] Z. Lu, H. J. Mo, Y. Lu, et al. “An empirical model for the star formation history in dark matter haloes”. In: MNRAS 439 (Apr. 2014), pp. 1294–1312.

[112] Z. Lu, H. J. Mo, and Y. Lu. “Galaxy ecosystems: gas contents, inflows and outflows”. In: MNRAS 450 (June 2015), pp. 606–617.

[113] M.-M. Mac Low and A. Ferrara. “Starburst-driven Mass Loss from Dwarf Galaxies: Efficiency and Metal Ejection”. In: ApJ 513 (Mar. 1999), pp. 142–155.

[114] P. Madau, F. Haardt, and M. J. Rees. "Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources". In: ApJ 514 (Apr. 1999), pp. 648–659.

[115] R. Maiolino, T. Nagao, A. Grazian, et al. "AMAZE. I. The evolution of the mass-metallicity relation at z $>$ 3". In: A&A 488 (Sept. 2008), pp. 463–479.

[116] R. Mandelbaum, U. Seljak, G. Kauffmann, et al. "Galaxy halo masses and satellite fractions from galaxy-galaxy lensing in the Sloan Digital Sky Survey: stellar mass, luminosity, morphology and environment dependencies". In: MNRAS 368 (May 2006), pp. 715–731.

[117] F. Mannucci, G. Cresci, R. Maiolino, et al. "A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies". In: MNRAS 408 (Nov. 2010), pp. 2115–2127.

[118] D. Marchesini, M. Stefanon, G. B. Brammer, et al. "The Evolution of the Rest-frame V-band Luminosity Function from z = 4: A Constant Faint-end Slope over the Last 12 Gyr of Cosmic History". In: ApJ 748, 126 (Apr. 2012), p. 126.

[119] D. Marchesini, P. G. van Dokkum, N. M. Förster Schreiber, et al. "The Evolution of the Stellar Mass Function of Galaxies from z = 4.0 and the First Comprehensive Analysis of its Uncertainties: Evidence for Mass-Dependent Evolution". In: ApJ 701 (Aug. 2009), pp. 1765–1796.

[120] C. L. Martin. "Properties of Galactic Outflows: Measurements of the Feedback from Star Formation". In: ApJ 513 (Mar. 1999), pp. 156–160.

[121] D. H. McIntosh, Y. Guo, J. Hertzberg, et al. "Ongoing assembly of massive galaxies by major merging in large groups and clusters from the SDSS". In: MNRAS 388 (Aug. 2008), pp. 1537–1556.

[122] R. J. McLure and J. S. Dunlop. "On the black hole-bulge mass relation in active and inactive galaxies". In: MNRAS 331 (Apr. 2002), pp. 795–804.

[123] M. L. Milne, C. J. Pritchet, G. B. Poole, et al. "The Faint End of the Luminosity Function in the Core of the Coma Cluster". In: AJ 133 (Jan. 2007), pp. 177–185.

[124] H. Mo, F. C. van den Bosch, and S. White. Galaxy Formation and Evolution. May 2010.

[125] H. J. Mo and S. Mao. "Galaxy formation in pre-heated intergalactic media". In: MNRAS 333 (July 2002), pp. 768–778.

[126] H. J. Mo and S. Mao. "Galaxy formation in pre-processed dark haloes". In: MNRAS 353 (Sept. 2004), pp. 829–840.

[127] H. J. Mo and J. Miralda-Escude. "Gaseous Galactic Halos and Quasi-stellar Object Absorption-Line Systems". In: ApJ 469 (Oct. 1996), p. 589.

[128] H. J. Mo and S. D. M. White. "An analytic model for the spatial clustering of dark matter haloes". In: MNRAS 282 (Sept. 1996), pp. 347–361.

[129] H. J. Mo, X. Yang, F. C. van den Bosch, et al. "Pre-heating by pre-virialization and its impact on galaxy formation". In: MNRAS 363 (Nov. 2005), pp. 1155–1166.

[130] H. J. Mo, S. Mao, and S. D. M. White. "The formation of galactic discs". In: MNRAS 295 (Apr. 1998), pp. 319–336.

[131] B. Moore, N. Katz, G. Lake, et al. "Galaxy harassment and the evolution of clusters of galaxies". In: Nature 379 (Feb. 1996), pp. 613–616.

[132] B. P. Moster, R. S. Somerville, C. Maulbetsch, et al. "Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift". In: ApJ 710 (Feb. 2010), pp. 903–923.

[133] B. P. Moster, T. Naab, and S. D. M. White. "Galactic star formation and accretion histories from matching galaxies to dark matter haloes". In: MNRAS 428 (Feb. 2013), pp. 3121–3138.

[134] J. Moustakas, A. L. Coil, J. Aird, et al. "PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from z = 0-1". In: ApJ 767, 50 (Apr. 2013), p. 50.

[135] S. J. Mutch, G. B. Poole, and D. J. Croton. "Constraining the last 7 billion years of galaxy evolution in semi-analytic models". In: MNRAS 428 (Jan. 2013), pp. 2001–2016.

[136] S. J. Mutch, D. J. Croton, and G. B. Poole. "The simplest model of galaxy formation - I. A formation history model of galaxy stellar mass growth". In: MNRAS 435 (Nov. 2013), pp. 2445–2459.

[137] J. F. Navarro, C. S. Frenk, and S. D. M. White. "The Structure of Cold Dark Matter Halos". In: ApJ 462 (May 1996), p. 563.

[138] K. G. Noeske, S. M. Faber, B. J. Weiner, et al. "Star Formation in AEGIS Field Galaxies since z=1.1: Staged Galaxy Formation and a Model of Mass-dependent Gas Exhaustion". In: ApJL 660 (May 2007), pp. L47–L50.

[139] K. G. Noeske, B. J. Weiner, S. M. Faber, et al. "Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies". In: ApJL 660 (May 2007), pp. L43–L46.

[140] M. A. Norris, C. G. Escudero, F. R. Faifer, et al. "An extended star formation history in an ultra-compact dwarf". In: MNRAS 451 (Aug. 2015), pp. 3615–3626.

[141] B. D. Oppenheimer and R. Davé. "Mass, metal, and energy feedback in cosmological simulations". In: MNRAS 387 (June 2008), pp. 577–600.

[142] B. D. Oppenheimer, R. Davé, D. Kereš, et al. "Feedback and recycled wind accretion: assembling the z = 0 galaxy mass function". In: MNRAS 406 (Aug. 2010), pp. 2325–2338.

[143] C. Pacifici, S. Charlot, J. Blaizot, et al. "Relative merits of different types of rest-frame optical observations to constrain galaxy physical parameters". In: MNRAS 421 (Apr. 2012), pp. 2002–2024.

[144] E. Papastergis, A. Cattaneo, S. Huang, et al. "A Direct Measurement of the Baryonic Mass Function of Galaxies and Implications for the Galactic Baryon Fraction". In: ApJ 759, 138 (Nov. 2012), p. 138.

[145] H. Parkinson, S. Cole, and J. Helly. "Generating dark matter halo merger trees". In: MNRAS 383 (Jan. 2008), pp. 557–564.

[146] J. A. Peacock and R. E. Smith. "Halo occupation numbers and galaxy bias". In: MNRAS 318 (Nov. 2000), pp. 1144–1156.

[147] M. S. Peeples and F. Shankar. "Constraints on star formation driven galaxy winds from the mass-metallicity relation at z= 0". In: MNRAS 417 (Nov. 2011), pp. 2962–2981.

[148] M. S. Peeples, J. K. Werk, J. Tumlinson, et al. "A Budget and Accounting of Metals at z ˜ 0: Results from the COS-Halos Survey". In: ApJ 786, 54 (May 2014), p. 54.

[149] M. Peimbert and R. Costero. "Chemical Abundances in Galactic HII Regions". In: Boletin de los Observatorios Tonantzintla y Tacubaya 5 (May 1969), pp. 3–22.

[150] Y.-j. Peng, S. J. Lilly, K. Kovač, et al. "Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function". In: ApJ 721 (Sept. 2010), pp. 193–221.

[151] P. G. Pérez-González, G. H. Rieke, V. Villar, et al. “The Stellar Mass Assembly of Galaxies from z = 0 to z = 4: Analysis of a Sample Selected in the Rest-Frame Near-Infrared with Spitzer”. In: ApJ 675 (Mar. 2008), pp. 234–261.

[152] M. Pettini and B. E. J. Pagel. “[OIII]/[NII] as an abundance indicator at high redshift”. In: MNRAS 348 (Mar. 2004), pp. L59–L63.

[153] L. S. Pilyugin and T. X. Thuan. “Oxygen Abundance Determination in H II Regions: The Strong Line Intensities-Abundance Calibration Revisited”. In: ApJ 631 (Sept. 2005), pp. 231–243.

[154] A. Pontzen and F. Governato. “How supernova feedback turns dark matter cusps into cores”. In: MNRAS 421 (Apr. 2012), pp. 3464–3471.

[155] P. Popesso, A. Biviano, H. Böhringer, et al. “RASS-SDSS Galaxy cluster survey. IV. A ubiquitous dwarf galaxy population in clusters”. In: A&A 445 (Jan. 2006), pp. 29–42.

[156] G. Popping, K. I. Caputi, R. S. Somerville, et al. “An indirect measurement of gas evolution in galaxies at $0.5 < z < 2.0$”. In: MNRAS 425 (Sept. 2012), pp. 2386–2400.

[157] G. Popping, R. S. Somerville, and S. C. Trager. “Evolution of the atomic and molecular gas content of galaxies”. In: MNRAS 442 (Aug. 2014), pp. 2398–2418.

[158] L. Portinari, C. Chiosi, and A. Bressan. “Galactic chemical enrichment with new metallicity dependent stellar yields”. In: A&A 334 (June 1998), pp. 505–539.

[159] W. H. Press and P. Schechter. “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation”. In: ApJ 187 (Feb. 1974), pp. 425–438.

[160] R. M. Reddick, R. H. Wechsler, J. L. Tinker, et al. “The Connection between Galaxies and Dark Matter Structures in the Local Universe”. In: ApJ 771, 30 (July 2013), p. 30.

[161] M. J. Rees. “Lyman absorption lines in quasar spectra - Evidence for gravitationally-confined gas in dark minihaloes”. In: MNRAS 218 (Jan. 1986), 25P–30P.

[162] A. E. Reines, J. E. Greene, and M. Geha. “Dwarf Galaxies with Optical Signatures of Active Massive Black Holes”. In: ApJ 775, 116 (Oct. 2013), p. 116.

[163] E. Roediger and M. Brüggen. “Ram pressure stripping of disc galaxies orbiting in clusters - I. Mass and radius of the remaining gas disc”. In: MNRAS 380 (Oct. 2007), pp. 1399–1408.

[164] E. Roediger and M. Brüggen. "Ram pressure stripping of disc galaxies: the role of the inclination angle". In: MNRAS 369 (June 2006), pp. 567–580.

[165] E. E. Salpeter. "The Luminosity Function and Stellar Evolution." In: ApJ 121 (Jan. 1955), p. 161.

[166] P. Santini, D. J. Rosario, L. Shao, et al. "Enhanced star formation rates in AGN hosts with respect to inactive galaxies from PEP-Herschel observations". In: A&A 540, A109 (Apr. 2012), A109.

[167] P. Santini, A. Fontana, A. Grazian, et al. "The evolving slope of the stellar mass function at $0.6 < z < 4.5$ from deep WFC3 data". In: A&A 538, A33 (Feb. 2012), A33.

[168] R. Scoccimarro, R. K. Sheth, L. Hui, et al. "How Many Galaxies Fit in a Halo? Constraints on Galaxy Formation Efficiency from Spatial Clustering". In: ApJ 546 (Jan. 2001), pp. 20–34.

[169] N. Scott, A. W. Graham, and J. Schombert. "The Supermassive Black Hole Mass-Spheroid Stellar Mass Relation for Sérsic and Core-Sérsic Galaxies". In: ApJ 768, 76 (May 2013), p. 76.

[170] U. Seljak. "Analytic model for galaxy and dark matter clustering". In: MNRAS 318 (Oct. 2000), pp. 203–213.

[171] F. Shankar, A. Lapi, P. Salucci, et al. "New Relationships between Galaxy Properties and Host Halo Mass, and the Role of Feedbacks in Galaxy Formation". In: ApJ 643 (May 2006), pp. 14–25.

[172] F. Shankar, D. H. Weinberg, and J. Miralda-Escudé. "Self-Consistent Models of the AGN and Black Hole Populations: Duty Cycles, Accretion Rates, and the Mean Radiative Efficiency". In: ApJ 690 (Jan. 2009), pp. 20–41.

[173] R. K. Sheth, H. J. Mo, and G. Tormen. "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes". In: MNRAS 323 (May 2001), pp. 1–12.

[174] L. Simard, J. T. Mendel, D. R. Patton, et al. "A Catalog of Bulge+disk Decompositions and Updated Photometry for 1.12 Million Galaxies in the Sloan Digital Sky Survey". In: ApJS 196, 11 (Sept. 2011), p. 11.

[175] J. Skilling. "Nested Sampling for General Bayesian Computation". In: Bayesian Analysis (2006), pp. 833–860.

[176] R. Smit, R. J. Bouwens, M. Franx, et al. “The Star Formation Rate Function for Redshift z ˜ 4-7 Galaxies: Evidence for a Uniform Buildup of Star-forming Galaxies during the First 3 Gyr of Cosmic Time”. In: ApJ 756, 14 (Sept. 2012), p. 14.

[177] R. S. Somerville, P. F. Hopkins, T. J. Cox, et al. “A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei”. In: MNRAS 391 (Dec. 2008), pp. 481–506.

[178] V. Springel, T. Di Matteo, and L. Hernquist. “Modelling feedback from stars and black holes in galaxy mergers”. In: MNRAS 361 (Aug. 2005), pp. 776–794.

[179] V. Springel, S. D. M. White, A. Jenkins, et al. “Simulations of the formation, evolution and clustering of galaxies and quasars”. In: Nature 435 (June 2005), pp. 629–636.

[180] G. Stasińska. “Biases in abundance derivations for metal-rich nebulae”. In: A&A 434 (May 2005), pp. 507–520.

[181] G. Stinson, A. Seth, N. Katz, et al. “Star formation and feedback in smoothed particle hydrodynamic simulations - I. Isolated galaxies”. In: MNRAS 373 (Dec. 2006), pp. 1074–1090.

[182] L. J. Tacconi, R. Neri, R. Genzel, et al. “Phibss: Molecular Gas Content and Scaling Relations in z ˜ 1-3 Massive, Main-sequence Star-forming Galaxies”. In: ApJ 768, 74 (May 2013), p. 74.

[183] J. E. Taylor and A. Babul. “The Dynamics of Sinking Satellites around Disk Galaxies: A Poor Man’s Alternative to High-Resolution Numerical Simulations”. In: ApJ 559 (Oct. 2001), pp. 716–735.

[184] J. E. Taylor and A. Babul. “The evolution of substructure in galaxy, group and cluster haloes - I. Basic dynamics”. In: MNRAS 348 (Mar. 2004), pp. 811–830.

[185] C. E. Thornton, A. J. Barth, L. C. Ho, et al. “The Host Galaxy and Central Engine of the Dwarf Active Galactic Nucleus POX 52”. In: ApJ 686 (Oct. 2008), pp. 892–910.

[186] E. Toloba, A. Boselli, R. F. Peletier, et al. “Formation and evolution of dwarf early-type galaxies in the Virgo cluster. II. Kinematic scaling relations (Corrigendum)”. In: A&A 557, C2 (Sept. 2013), p. C2.

[187] A. R. Tomczak, R. F. Quadri, K.-V. H. Tran, et al. “Galaxy Stellar Mass Functions from ZFOURGE/CANDELS: An Excess of Low-mass Galaxies since z = 2 and the Rapid Buildup of Quiescent Galaxies”. In: ApJ 783, 85 (Mar. 2014), p. 85.

[188] C. A. Tremonti, T. M. Heckman, G. Kauffmann, et al. "The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey". In: ApJ 613 (Oct. 2004), pp. 898–913.

[189] A. Vale and J. P. Ostriker. "Linking halo mass to galaxy luminosity". In: MNRAS 353 (Sept. 2004), pp. 189–200.

[190] F. van de Voort, J. Schaye, C. M. Booth, et al. "The rates and modes of gas accretion on to galaxies and their gaseous haloes". In: MNRAS 414 (July 2011), pp. 2458–2478.

[191] F. C. van den Bosch, X. Yang, and H. J. Mo. "Linking early- and late-type galaxies to their dark matter haloes". In: MNRAS 340 (Apr. 2003), pp. 771–792.

[192] F. C. van den Bosch, D. Aquino, X. Yang, et al. "The importance of satellite quenching for the build-up of the red sequence of present-day galaxies". In: MNRAS 387 (June 2008), pp. 79–91.

[193] A. van der Wel, M. Franx, P. G. van Dokkum, et al. "3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since z = 3". In: ApJ 788, 28 (June 2014), p. 28.

[194] P. G. van Dokkum, K. E. Whitaker, G. Brammer, et al. "The Growth of Massive Galaxies Since z = 2". In: ApJ 709 (Feb. 2010), pp. 1018–1041.

[195] L. Wang, D. Farrah, S. J. Oliver, et al. "Connecting stellar mass and star-formation rate to dark matter halo mass out to z $\sim$ 2". In: MNRAS 431 (May 2013), pp. 648–661.

[196] G. A. Wegner. "The double galaxy cluster Abell 2465 - I. Basic properties: optical imaging and spectroscopy". In: MNRAS 413 (May 2011), pp. 1333–1352.

[197] S. M. Weinmann, A. Pasquali, B. D. Oppenheimer, et al. "A fundamental problem in our understanding of low-mass galaxy evolution". In: MNRAS 426 (Nov. 2012), pp. 2797–2812.

[198] S. M. Weinmann, G. Kauffmann, A. von der Linden, et al. "Cluster galaxies die hard". In: MNRAS 406 (Aug. 2010), pp. 2249–2266.

[199] S. M. Weinmann, F. C. van den Bosch, X. Yang, et al. "Properties of galaxy groups in the Sloan Digital Sky Survey - I. The dependence of colour, star formation and morphology on halo mass". In: MNRAS 366 (Feb. 2006), pp. 2–28.

[200] S. M. Weinmann, F. C. van den Bosch, X. Yang, et al. "Properties of galaxy groups in the Sloan Digital Sky Survey - II. Active galactic nucleus feedback and star formation truncation". In: MNRAS 372 (Nov. 2006), pp. 1161–1174.

[201] D. R. Weisz, A. E. Dolphin, E. D. Skillman, et al. "Comparing the ancient star formation histories of the Magellanic Clouds". In: MNRAS 431 (May 2013), pp. 364–371.

[202] D. R. Weisz, J. J. Dalcanton, B. F. Williams, et al. "The ACS Nearby Galaxy Survey Treasury. VIII. The Global Star Formation Histories of 60 Dwarf Galaxies in the Local Volume". In: ApJ 739, 5 (Sept. 2011), p. 5.

[203] A. R. Wetzel, J. L. Tinker, C. Conroy, et al. "Galaxy evolution in groups and clusters: satellite star formation histories and quenching time-scales in a hierarchical Universe". In: MNRAS 432 (June 2013), pp. 336–358.

[204] A. R. Wetzel, J. L. Tinker, and C. Conroy. "Galaxy evolution in groups and clusters: star formation rates, red sequence fractions and the persistent bimodality". In: MNRAS 424 (July 2012), pp. 232–243.

[205] S. D. M. White and C. S. Frenk. "Galaxy formation through hierarchical clustering". In: ApJ 379 (Sept. 1991), pp. 52–79.

[206] S. D. M. White and M. J. Rees. "Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering". In: MNRAS 183 (May 1978), pp. 341–358.

[207] X. Yang, H. J. Mo, and F. C. van den Bosch. "Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies". In: MNRAS 339 (Mar. 2003), pp. 1057–1080.

[208] X. Yang, H. J. Mo, F. C. van den Bosch, et al. "Constraining the Star Formation Histories in Dark Matter Halos. I. Central Galaxies". In: ApJ 770, 115 (June 2013), p. 115.

[209] X. Yang, H. J. Mo, F. C. van den Bosch, et al. "Evolution of the Galaxy-Dark Matter Connection and the Assembly of Galaxies in Dark Matter Halos". In: ApJ 752, 41 (June 2012), p. 41.

[210] X. Yang, H. J. Mo, and F. C. van den Bosch. "Galaxy Groups in the SDSS DR4. II. Halo Occupation Statistics". In: ApJ 676 (Mar. 2008), pp. 248–261.

[211] H. J. Zahid, P. Torrey, M. Vogelsberger, et al. "Empirical constraints for the magnitude and composition of galactic winds". In: Ap&SS 349 (Feb. 2014), pp. 873–879.

[212] A. R. Zentner, A. A. Berlind, J. S. Bullock, et al. “The Physics of Galaxy Clustering. I. A Model for Subhalo Populations”. In: ApJ 624 (May 2005), pp. 505–525.

[213] D. H. Zhao, Y. P. Jing, H. J. Mo, et al. “Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos”. In: ApJ 707 (Dec. 2009), pp. 354–369.