
GALAXY OBSERVABLES DURING THE EPOCH OF REIONIZATION FROM SEMI-ANALYTIC MODELS

CHUANWU LIU

Doctor of Philosophy

July 29, 2017

SCHOOL OF PHYSICS
THE UNIVERSITY OF MELBOURNE

*Submitted in Total Fulfillment
of the Requirements of the Degree of Doctor of Philosophy*

Abstract

The aim of this thesis is to study galaxy observables including UV luminosity and disc sizes at high redshifts using the new semi-analytic galaxy formation model MERAXES. We present calculations of UV luminosities and treatments of dust extinction and Ly α absorption. We show that we successfully reproduce the observed UV luminosity functions (LFs) at $z \sim 5\text{--}10$. We find that the faint-end slope of UV LFs remains steep below current detection limits until at least $M_{\text{UV}} \sim -14$ indicating a large number of undetected faint galaxies during the epoch of reionization (EoR). We find that these undetected galaxies contribute significantly to the reionization. For example, galaxies with luminosities in the range $-17 < M_{\text{UV}} < -13$ emit more flux than the $M_{\text{UV}} < -17$ galaxies at $z > 7$. These predictions will be tested by next generation telescopes including *JWST*. We find that model galaxies with $M_{\text{UV}} \lesssim -14$ are distributed around the empirical luminosity–SFR relation with 0.1–0.3 dex, and that this scatter may bias the luminosity–SFR conversion for galaxies with low SFR ($< 0.01M_{\odot}/\text{yr}$). We also show that the mean luminosity–stellar mass relation and the mean luminosity–halo mass relation from the model are in good agreement with observations, and can be fitted using singular power laws as $M_{*} \propto L_{\text{UV}}^{1.2}$ and $M_{\text{vir}} \propto L_{\text{UV}}^{0.88}$ respectively. These relations retain their form until at least $M_{\text{UV}} \sim -14$ and may break at fainter luminosities due to the hydrogen cooling limit in low-mass haloes. We show that galaxies with $M_{\text{UV}} \sim -20$ are hosted in dark matter haloes of mass $M_{\text{vir}} \sim 10^{11.0 \pm 0.1} M_{\odot}$ at $z \sim 6$. This mass decreases linearly towards higher redshift. We find a size–luminosity relation of $R_{\text{e}} \propto L_{\text{UV}}^{0.25}$, in good agreement with observations including the recently discovered luminous galaxy GN-z11. We also find that galaxy sizes can be an important observational probe for galaxy formation during the EoR, allowing us to discriminate between different star formation and feedback models.

Declaration

This is to certify that:

1. this thesis comprises only my original work towards the PhD except where indicated in the Preface;
2. due acknowledgement has been made in the text to all other material used;
3. this thesis is fewer than 100 000 words in length, exclusive of tables, figures, bibliographies and appendices.

Chuanwu Liu

Preface

This thesis comprises my own work, except where explicitly mentioned in the text. Some of the work have been published in refereed journals within the paper series of the *Dark-Ages Reionization and Galaxy formation Observables from Numerical Simulations (DRAGONS)* project, including:

- Chapter 2, 3 and 4 have been published as:

Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B. Mesinger A., Wyithe J. S. B., 2016

Dark-ages reionization and galaxy formation simulation-IV. UV luminosity functions of high-redshift galaxies

Monthly Notices of the Royal Astronomical Society, Volume 462, Issue 1, p.235-249 (2016)

- Chapter 5 has been published as:

Liu C., Mutch S. J., Poole G. B., Angel P. W., Duffy A. R., Geil P. M., Mesinger A., Wyithe J. S. B.

Dark-ages reionization and galaxy formation simulation-VII. The sizes of high-redshift galaxies

Monthly Notices of the Royal Astronomical Society, Volume 465, Issue 3, p.3134-3142 (2017)

Moderate alterations and additional materials have been added to these chapters in order to maintain consistency of the style.

- Chapter 6 has mentioned the conclusion from one of co-authored publications:

Mutch S. J., Liu C., Poole G. B., Geil P. M., Duffy A. R., Trenti M., Oesch P. A., Illingworth G. D., Mesinger A., Wyithe J. S. B

Dark-ages reionization and galaxy formation simulation-VI. The origins and fate of the highest known redshift galaxy

Monthly Notices of the Royal Astronomical Society, Volume 463, Issue 6, p.3556-3562 (2016)

This doctoral research was carried out at the University of Melbourne under the joint supervision of Prof. Stuart Wyithe and Dr. Simon Mutch.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Stuart Wyithe, who provided unfailing guidance and support throughout every step of my thesis research. Stuart paid great patience on reading and correcting my thesis and papers. This thesis and all my PhD research would not have been possible without his support.

Equally deep gratitude are also to my co-supervisor Simon Mutch who provided unlimited support as far as he could in the past years. He spent a lot of time on answering my questions and giving me detailed instructions. It was so many times that I walked into his office and disturbed his own work. Simon also invested a large amount of time on checking my paper and thesis writing.

I would like to thank Gregory Poole and Alan Duffy who informally involved in my PhD supervision. Both of them gave me essential support and encouragement, which made me feel confident in my research.

I would like to thank other *DRAGONS* members: Paul Geil, Paul Angel, Yuxiang Qin and Andrei Mesinger who gave me many valuable comments on my work. I want to thank Hansik Kim, Manodeep Sinha, Jaehong Park, Antonios Katsianis and Edoardo Tescari who shared knowledge with me. Hansik Kim also provided me financial support when my scholarship was suspended.

Thanks also go to other individuals who helped me in one way or another: Andy Martin who was a member of my PhD panel; Rychard Bouwens who kindly wrote reference letters for my job application; Kim Dorrell and Kirsty Waring who provided assistance for DRAGONS group; Rachel Webster, Andrew Melatos and Michele Trenti who maintained a great scientific environment in Astrogroup.

I enjoyed my student life with many interesting friends in Astrogroup including Sarah Schon, Daniela Carrasco Nunez, Camila Correa, Daniel Neri-Larios, Jack Line, Clare Kenyon, Stephanie Bernard, George Howitt, Lili Sun and Suk Yee Yong.

Finally, I want to thank my wife and other family members, whose support and love extend to my whole life. I can not express my grateful using language.

Contents

1	Introduction	1
1.1	Cosmological background	1
1.1.1	Extragalactic astronomy	1
1.1.2	Expanding universe	2
1.1.3	Relativistic cosmology	4
1.1.4	Cosmological distances	6
1.1.5	Chronology of the Universe	7
1.2	The epoch of reionization	9
1.2.1	Observing reionization	9
1.2.2	Thomson scattering of CMB photons	11
1.2.3	Sources for reionization	13
1.3	Structure Formation	14
1.3.1	Density perturbations	15
1.3.2	Dark matter haloes	17
1.3.3	Internal properties of haloes	19
1.3.4	Cosmological simulations	20
1.3.5	Semi-analytic models	23
1.4	Galaxies and Observables	24
1.4.1	Jeans length	24
1.4.2	Hierarchical galaxy formation	24
1.4.3	Galaxy observables	28
1.5	High- z galaxies and the DRAGONS project	33
1.6	Thesis Structure	34
2	Modeling galaxy observables	35
2.1	Introduction	35
2.2	<i>Tiamat</i> N -body simulation	35
2.2.1	Volume and resolution	35
2.2.2	Merger tree construction	36

2.3	The semi-analytic model — MERAXES	38
2.3.1	Gas infall and cooling	38
2.3.2	Star formation	40
2.3.3	Delayed supernova feedback	40
2.3.4	Metal enrichment	42
2.3.5	Stellar mass recycling	43
2.3.6	Halo infall and gas stripping	43
2.3.7	Mergers	44
2.3.8	Reionization	44
2.3.9	Model calibration	47
2.4	Modeling luminosities	47
2.4.1	Stellar population synthesis	47
2.4.2	Lya absorption	48
2.4.3	Dust attenuation	50
2.5	Summary	53
3	UV luminosity functions	55
3.1	Introduction	55
3.2	Lyman-break selection	56
3.3	UV LFs	58
3.4	UV flux from galaxies below the detection limit	62
3.5	Conclusions	64
4	Probing intrinsic galaxy and halo properties	67
4.1	Introduction	67
4.2	UV luminosity–SFR relation	68
4.3	The SFR functions	69
4.4	UV luminosity–stellar mass relation	71
4.5	UV luminosity–halo mass relation	75
4.6	Conclusions	78
5	The sizes of galaxies	81
5.1	Introduction	81
5.2	Modeling disk sizes	83
5.3	Size–luminosity relation	86
5.4	Resolving galaxies with <i>HST</i> , <i>JWST</i> and <i>GMT</i>	88
5.5	Mass–size relation	89
5.6	Redshift evolution of sizes	90
5.7	Measuring galaxy sizes	92

5.8	Conclusions	93
6	Summary and future work	95
6.1	Summary	95
6.2	Future work	97
A	UV LFs with different metallicities	111

List of Abbreviations

A&A Astronomical and Astrophysics

AGN Active Galactic Nuclei

AJ The Astronomical Journal

ApJ The Astrophysical Journal

ApJL The Astrophysical Journal Letters

ApJS The Astrophysical Journal Supplements

CMB Cosmic Microwave Background

DRAGONS Dark-Ages Reionization And Galaxy Formation Observables from Numerical Simulations

EoR Epoch of Reionization

FoF Friend-of-Friend

HOD Halo Occupation Distribution

HST Hubble Space Telescope

HAM Halo Abundance Matching

HERA Hydrogen Epoch of Reionization Array

IGM Inter-Galactic Medium

IMF Initial Mass Function

ISM Inter-Stellar Medium

JWST James Webb Space Telescope

Λ CDM Lambda–Cold Dark Matter

LF Luminosity Function

MNRAS Monthly Notices of the Royal Astronomical Society

NFW Navarro-Frenk-White

SED Spectral Energy Distribution

SFR Star-Formation Rate

SHAM Sub-Halo Abundance Matching

SKA The Square Kilometre Array

UV Ultra-Violet

UVB Ultra-Violet Background

List of Tables

2.1	Summary of parameters in the <i>Tiamat</i> N-body simulations	36
2.2	Summary of fiducial parameters in MERAXES	47
3.1	The fraction of UV flux at 1600Å above the luminosity limits.	63
4.1	The best-fit slopes and intercepts of the median $\log_{10} M_* - M_{\text{UV}}$ relation (Equation 4.3) for galaxies with $M_{\text{UV}} \leq -14$	72
4.2	The best fitting slopes and the intercepts of the $\log_{10} M_* - M_{\text{UV}}$ relation for galaxies with $M_{\text{UV}} \leq -14$ based on the <i>Tiny Tiamat</i> N-body simulation.	75
4.3	The best fitting slopes and the intercepts of the median $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation (Equation 4.5) for galaxies with $M_{\text{UV}} \leq -14$	75
4.4	The best fitting slopes and the intercepts of the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation for galaxies with $M_{\text{UV}} \leq -14$ based on the <i>Tiny Tiamat</i> N-body simulation.	79
5.1	Observed evolution of galaxy sizes, $R_e \propto (1+z)^m$ from literature, where $L_{z=3}^*$ corresponds to UV magnitude $M_{\text{UV}} = -21.0$	82
5.2	The best-fitting parameters R_0 and β (Equation 5.4) for the model galaxies with UV magnitudes $M_{\text{UV}} < -14$ at $z \sim 5-10$	87

List of Figures

1.1	The anisotropies of the CMB observed by Planck Collaboration et al. (2014)	3
1.2	The spectra of 19 SDSS quasars at $5.74 < z < 6.42$ and the measured Gunn-Peterson optical depth from Fan et al. (2006b)	10
1.3	An illustration of a merger tree	19
1.4	Dark matter structures from the <i>Millennium</i> N-body simulation at $t \sim 4.7$ Gyr	21
1.5	Observed UV LFs at $z \sim 4\text{--}8$ from Bouwens et al. (2015b)	29
1.6	Observed SFR densities for galaxies with UV luminosity brighter than -17 mag from Bouwens et al. (2015b)	31
2.1	The illustration of horizontal-form merger trees compared to vertical-form trees	37
2.2	An observed SED from a model star-forming galaxy at $z = 6$	49
2.3	The average dust attenuation at 1600Å as a function of (dust-attenuated) luminosity at $z \sim 5\text{--}10$	50
2.4	The comparison between the Calzetti et al. (2000) reddening curve and the Charlot & Fall (2000) reddening curve	53
3.1	LBG selection criteria used to select star-forming galaxies at $z \sim 5\text{--}10$	57
3.2	Model UV LFs at $z \sim 5\text{--}10$ from MERAXES, compared with observations	59
3.3	Model UV LFs from MERAXES for all model galaxies extended to low luminosities, illustrating the predicted flattening fainter than $M_{1600} \sim -14$	60
3.4	Comparison for UV LFs of all model galaxies based on <i>Tiamat</i> and <i>Tiny Tiamat</i> simulations at $z \sim 6, 8$ and 10	61
3.5	The cumulative fraction of 1600Å UV flux from the model galaxies brighter the luminosity limit M_{1600}	63
4.1	The instantaneous SFR of galaxies as a function of their (unattenuated) intrinsic UV luminosity M_{1600}^i at $z \sim 5\text{--}10$	68

4.2	SFR functions of galaxies at $z \sim 5\text{--}10$	70
4.3	Stellar mass of galaxies as a function of UV luminosity M_{1500}	72
4.4	The best fitting slopes and the intercepts for the luminosity–stellar mass relation of galaxies	73
4.5	Stellar mass as a function of UV luminosity for model galaxies based on <i>Tiny Tiamat</i>	74
4.6	Dark matter halo (FoF group) mass as a function of UV luminosity	76
4.7	The best fitting slopes and the intercepts for the luminosity–halo mass relation of galaxies	77
4.8	Dark matter halo (FoF group) mass as a function of UV luminosity for model galaxies based on <i>Tiny Tiamat</i>	79
5.1	Effective radius of galaxies as a function of UV luminosity at $z \sim 5\text{--}10$.	85
5.2	Size-mass relation of model galaxies at $z = 5, 6, 8, 10$	89
5.3	The redshift evolution of the mean effective radius for galaxies	91
5.4	Confidence ellipses of the fitted power law of the galaxy size evolution .	92
6.1	The median galaxy mass–metallicity relation from MERAXES	97
A.1	The calculated UV LFs with constant metallicity of $Z = 0.001, 0.004$ and 0.008 at $z \sim 6, 8$ and 10	112
A.2	The calculated UV LFs using evolving metallicity compared with the UV LFs with assumed constant metallicity	112

Chapter 1

Introduction

This chapter gives a brief background introduction to cosmology, the theory of structure formation, current research challenges associated with the epoch of reionization, and the numerical modeling of high-redshift (high- z) galaxy formation and evolution.

1.1 Cosmological background

1.1.1 *Extragalactic astronomy*

Human beings have explored the Universe for thousands of years. We have observed millions of light sources in the sky, including planets, stars and the Milky Way. In 1610, Galileo Galilei resolved the light band of the Milky Way into huge numbers of faint stars using his telescope. The Milky Way Galaxy was later established to be a rotating system of a huge number of stars, held together by gravitational forces.

Some objects in the sky with a fuzzy appearance, rather than point-like sources, were discovered. These objects were initially called nebulae. After the systematic works of Charles Messier and William Herschel, astronomers discovered a large number of these nebulae. However, the nature of nebular objects was, for a long-time, controversial, and leading to the ‘Great Debate’ between Harlow Shapley and Heber Curtis about the nature of spiral nebulae in 1920. Shapley believed that the Milky Way was the entire Universe, and spiral nebulae were within our Milky Way. Curtis on the other side was arguing that the spiral nebulae were other galaxies out side of our Milky Way.

Using Cepheid variables¹ in several spiral nebulae, Hubble (1929b) measured the distance of these spiral nebulae and found that they were too distant to be within our Milky Way. He suggested that these were actually other galaxies. This discovery was the beginning of extragalactic astronomy. After that, more and more galaxies with diverse morphologies and rich properties were found. Thanks to modern technology, we are able to explore deeper and deeper into the Universe using not only ground based optical telescopes, but also other observational means such as radio antenna arrays, space optical telescopes, X-ray/Gamma-ray satellites, neutrino detectors and gravitational wave detectors.

¹A special type of variable star first discovered in the constellation Cepheus. A strong relationship between a Cepheid variable’s luminosity and the associated changing period established Cepheids as important indicators of galactic distances (Leavitt & Pickering, 1912).

Today, we know that galaxies are gravitationally bound systems, which contain not only a large number of stars in the form of main sequence, brown dwarfs, white dwarfs, neutron stars and compact systems of stars like binaries and star clusters, but also include large amount of interstellar medium in the form of gas and dust. Observations also strongly suggest that many galaxies have super-massive black holes in their centre. According to the currently favored Λ cold dark matter (Λ CDM) cosmology, galaxies have dark matter haloes far beyond their visible boundaries.

Since galaxies are much brighter than individual stars and can be detected from large distances, they are an important probe for measuring the structure of the Universe. The formation and evolution of galaxies also play an important role in modern cosmology.

1.1.2 *Expanding universe*

One of the fundamental observational properties of a galaxy is its spectral redshift:²

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}} = \frac{\lambda_{\text{obs}}}{\lambda_{\text{em}}} - 1, \quad (1.1)$$

where λ_{obs} and λ_{em} are spectral wavelengths in the observer-frame and rest-frame respectively. Redshift is usually calculated by measuring the shift of characteristic spectral lines. For high- z galaxies with only photometric measurements, best-fitting template spectral energy distributions (SEDs) are commonly used.

Through measuring the redshifts of several relatively nearby extragalactic galaxies, Hubble (1929a) found that galaxies recede from us with velocities v proportional to their proper distances d

$$v = H_0 d. \quad (1.2)$$

Here H_0 is the *Hubble constant* at $z = 0$ which is usually presented using the dimensionless Hubble parameter h as

$$H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (1.3)$$

The recent results from Planck Collaboration et al. (2016) indicate that $h = 0.678 \pm 0.009$. The linear relation of Equation 1.2 is called Hubble's law which indicates that our Universe is expanding. This is called *Hubble expansion* or Hubble flow. The discovery of Hubble expansion marked the birth of modern cosmology.

The inverse of H_0 is called the *Hubble time*,

$$t_H \equiv \frac{1}{H_0} = 9.778 \times 10^9 h^{-1} \text{ yr}, \quad (1.4)$$

which is an estimate of the age of the Universe if the Universe had always expanded at

²An observed redshift contains both cosmological and Doppler shift.

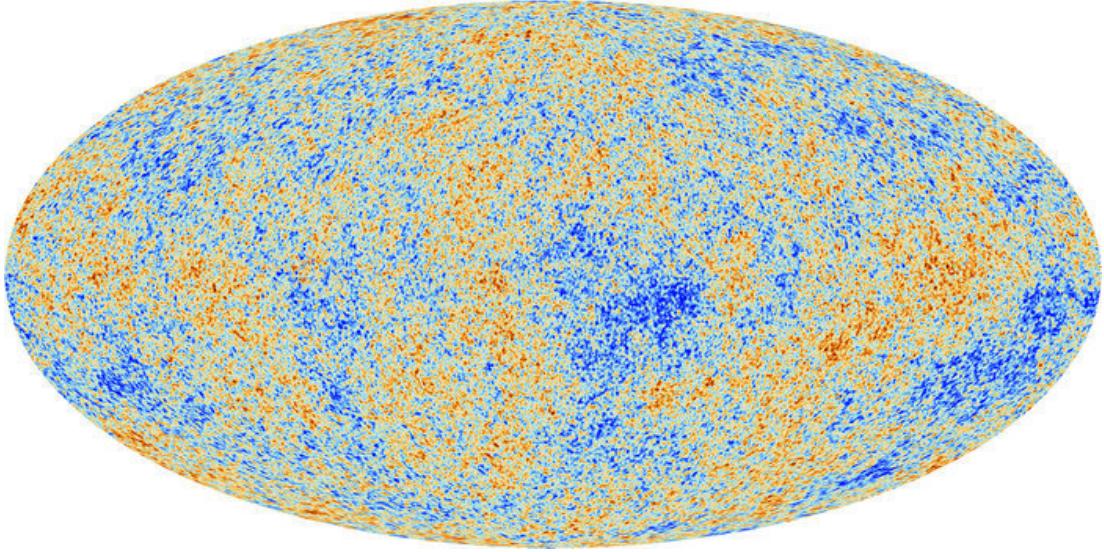


Figure 1.1: The anisotropies of the CMB observed by Planck Collaboration et al. (2014). The CMB photons decoupled with matter after recombination era ($t \sim 38$ kyr) and were red-shifted to 2.73 K. The CMB map shows tiny temperature fluctuations that correspond to slight density fluctuations in the early Universe which seeded the formation of structure at later times. Such a CMB map supports the cosmological principle that the Universe is isotropic and homogeneous on large scales.

its current rate. $t_H = 14.4$ billion years when choosing $h = 0.678$. The speed of light, c , times the Hubble time is the *Hubble distance*,

$$D_H \equiv \frac{c}{H_0} = 3.0 \times 10^3 h^{-1} \text{Mpc}, \quad (1.5)$$

which is used to estimate the characteristic scale of the Universe.

A fundamental hypothesis of cosmology is the cosmological principle, which assumes that on sufficiently large scales, the Universe is spatially isotropic and homogeneous. The cosmological principle can be demonstrated by observations including the mapping of the cosmic microwave background (CMB) radiation as shown in Figure 1.1 (e.g. Planck Collaboration et al., 2014). We can find a cosmological rest frame, for example using the CMB, in which the Universe is isotropic. From large volume surveys of galaxies, the number density of galaxies and the Hubble constant are measured to be the same in different directions, which agrees with the isotropic principle. The number density of galaxies is also observed to be invariant at different places in the Universe, which indicates homogeneity. Observations show that the cosmological principle is valid on cosmic scales (e.g. $\gtrsim 100$ Mpc). However, individual galaxies may have a *peculiar velocity* v_{pec} relative to the rest frame due to the gravitational effect on smaller scales. For example, The Local Group of galaxies has an average velocity of $627 \pm 22 \text{ km s}^{-1}$

relative to the CMB (Lineweaver et al., 1996). The distribution of mass traced by observed galaxies are found to be clustered on scales $\lesssim 10h^{-1}$ Mpc.

In an isotropic and homogeneous universe, we can define a *cosmic time* which is synchronized everywhere. At the same cosmic time, the local cosmological properties at every position of the Universe are the same. In an expanding universe, we can define a *comoving coordinate* \mathbf{x} which expands (comoving) with the expanding cosmic background. The proper coordinate of a position, \mathbf{r} , is therefore

$$\mathbf{r} = a(t)\mathbf{x}, \quad (1.6)$$

where $a(t)$ is the scale factor describes the scale of space at cosmic time t .

The observed spectral redshift of a galaxy contains both Hubble expansion and its peculiar motion. The observed redshift of an object due to the expansion of the Universe is related to the scale factor by

$$z = \frac{a(t_0)}{a(t_{\text{em}})} - 1, \quad (1.7)$$

where t_{em} is the time of the light when it was emitted from the object, and t_0 is present time that the light was observed. If we define the present value of scale factor $a_0 = 1$, the scale factor of the Universe at redshift z is

$$a(z) = \frac{1}{1+z}. \quad (1.8)$$

For two objects 1 and 2 with redshifts z_1 and z_2 ($z_1 < z_2$) relative to us, we can define a redshift z_{12} of object 2 relative to object 1, which is given by

$$1 + z_{12} = \frac{a(t_1)}{a(t_2)} = \frac{1 + z_2}{1 + z_1}. \quad (1.9)$$

Using the scale factor, the Hubble constant at t can be written as

$$H(t) = \frac{\dot{a}(t)}{a(t)}, \quad (1.10)$$

where the over-dot represents the time derivative with respect to cosmic time.

1.1.3 Relativistic cosmology

A 4-dimensional space-time hypersurface which is both homogeneous and isotropic can be described by the *Robertson-Walker Metric*,

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (1.11)$$

where ds is the space-time interval, (r, θ, ϕ) are spherical comoving coordinates, and k is the curvature, which has values of +1, 0 and -1 representing closed, flat and open universes respectively. Observations imply a flat Universe to high precision (Spergel et al., 2003, 2007; Planck Collaboration et al., 2016).

The standard cosmology is built upon Einstein's general relativity, where space-time is curved by the matter and energy in the Universe through Einstein's field equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi GT_{\mu\nu}}{c^4}. \quad (1.12)$$

Here $R_{\mu\nu}$ is the Ricci curvature tensor describing the local curvature, R is the curvature scalar, $g_{\mu\nu}$ is the metric tensor, Λ is the cosmological constant, G is the gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor. For $T_{\mu\nu}$ from a uniform ideal fluid and $g_{\mu\nu}$ from a isotropic and homogeneous universe, which is given by the Robertson-Walker Metric, Einstein's equations yield

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \quad (1.13)$$

where ρ is the mass density of the fluid. This equation is known as the Friedmann equation. The density ρ is contributed by different components: (i) non-relativistic matter including baryons and dark matter, $\rho_m \propto a^{-3}$; (ii) radiation, $\rho_r \propto a^{-4}$; (iii) cosmological constant, $\rho_\Lambda = \Lambda c^2/8\pi G$, which can be considered as an energy component whose density does not change with time.

We can define a critical density which corresponds to $k = 0$:

$$\rho_{\text{crit}}(t) = \frac{3H^2(t)}{8\pi G}, \quad (1.14)$$

then define $\Omega \equiv \rho/\rho_{\text{crit}}$, $\Omega_m \equiv \rho_m/\rho_{\text{crit}}$, $\Omega_r \equiv \rho_r/\rho_{\text{crit}}$ and $\Omega_\Lambda \equiv \rho_\Lambda/\rho_{\text{crit}}$. The Friedmann equation can thus be rewritten as

$$H(z) = H_0 E(z), \quad (1.15)$$

where

$$E(z) = [\Omega_{\Lambda,0} + (1 - \Omega_0)(1 + z)^2 + \Omega_{m,0}(1 + z)^3 + \Omega_{r,0}(1 + z)^4]^{1/2}. \quad (1.16)$$

Here the 0-subscript indicates the value at the present time. From Equation 1.14 and 1.15, we see that $\rho_{\text{crit}} \propto E^2(z)$. We can calculate Ω_Λ , Ω_m , Ω_r at a given z using their

present-time values:

$$\Omega_\Lambda(z) = \frac{\Omega_{\Lambda,0}}{E^2(z)}, \quad \Omega_m(z) = \frac{\Omega_{m,0}(1+z)^3}{E^2(z)}, \quad \Omega_r(z) = \frac{\Omega_{r,0}(1+z)^4}{E^2(z)}. \quad (1.17)$$

If we assume the expansion of the Universe starts at $a = 0$, the age of the Universe at redshift z can be calculated by

$$t(z) = \int_0^{a(z)} \frac{da}{\dot{a}} = \frac{1}{H_0} \int_z^\infty \frac{dz'}{(1+z')E(z')}. \quad (1.18)$$

The *lookback time* of object at redshift z can be therefore calculate by integrating the redshift from 0 to z .

1.1.4 Cosmological distances

These are several types of distances defined in cosmology (e.g. Mukhanov, 2005; Weinberg, 2008). A proper or ruler distance d is the distance for any two objects at the same time t measured by a super-observer. A comoving distance can be obtained by dividing the proper distance by the scale factor at that time. In the real Universe, distances are usually measured by observing light, where a photon emitted at a distance r_{em} and time t_{e} can be received at the observer at the origin. This *line-of-sight* distance between two objects is measured along null-geodesics ($ds^2 = 0$) of Equation 1.11,

$$c^2 dt^2 = a^2(t) \frac{dr^2}{1 - kr^2}. \quad (1.19)$$

Using the Equation 1.7, 1.10, and 1.15, the line-of-sight comoving distance in a flat universe is

$$r_{\text{em}} = \int_0^z \frac{c}{H(z')} dz' = D_H \int_0^z \frac{dz'}{E(z')}. \quad (1.20)$$

We can also define the *angular diameter distance* d_A of an object as the ratio of its physical size D to its angular size. The physical size D of the object is the proper distance between two parts of the object with the same r_{em} and ϕ at the cosmic time t_{em} , and can be calculated by integrating the apparent angle $d\delta$ over the transverse direction.

$$D = a(t_{\text{em}})r_{\text{em}}d\theta = a(t_{\text{em}})r_{\text{em}}\delta, \quad (1.21)$$

where δ is the angular size of the object. Therefore we have

$$d_A = \frac{D}{\delta} = a(t_{\text{em}})r_{\text{em}} = \frac{r_{\text{em}}}{1+z}. \quad (1.22)$$

The *luminosity distance* is defined as

$$d_L = \sqrt{\frac{L_{\text{em}}}{4\pi F_{\text{obs}}}}, \quad (1.23)$$

where L_{em} is the bolometric luminosity (integrated over all frequencies) of the object and F_{obs} is the observed flux. Due to the expansion of the Universe, the observed flux is changed by the time dilation and the redshift of light, where the former one results in a $1/(1+z)$ times lower arrival rate of photons and the latter one leads to a $1/(1+z)$ times energy reduction of individual photons (e.g. Weinberg, 2008). We therefore have:

$$d_L = (1+z)r_{\text{em}} = (1+z)^2 d_A. \quad (1.24)$$

1.1.5 Chronology of the Universe

The space-time evolution of the Universe determines its thermal history. In the standard Hot Big Bang cosmology (e.g. Mukhanov, 2005; Weinberg, 2008), The expansion of the Universe implies that it was extremely hot and dense at its beginning. Here we give a brief chronology of the physical Universe since the Big Bang, where t and T represent the age and temperature of the Universe respectively³.

- Plank epoch, $t < 10^{-43}$ s, $T > 10^{32}$ K. The description of this stage is beyond current Physics.
- Grand unification, $t \sim 10^{-43}\text{--}10^{-36}$ s, $T > 10^{27}$ K. The strong nuclear force, the weak nuclear and the electromagnetic force are unified as an electronuclear force. Gravity is separated from them.
- Inflation, $t \sim 10^{-36}\text{--}10^{-32}$ s. The Universe undergoes an extreme rapid expansion (Guth, 1981), which increases the scale of the Universe by a factor of at least 10^{26} .
- Electroweak epoch, $t \sim 10^{-36}\text{--}10^{-12}$ s. The strong nuclear force is separated from the electroweak interaction.
- Quark epoch, $t \sim 10^{-12}\text{--}10^{-6}$ s. Quarks, electrons, and neutrinos form in large numbers, but the temperature is too high to allow quarks to combine into hadrons (baryons and mesons). The four fundamental forces have separated.
- Hadron epoch, $t \sim 10^{-6}\text{--}1$ s, $T \sim 10^{10}\text{--}10^9$ K. The temperature of the Universe is cool enough to allow quarks to bind into hadrons. As the temperate continues to fall, hadrons and anti-hadrons are eliminated in annihilation reactions. Baryon asymmetry starts from this stage.

³In a adiabatic universe, $T \propto (1+z)$.

- Lepton epoch, $t \sim 1\text{--}10\text{ s}$, $T \sim 10^9\text{ K}$. After the majority of hadrons and anti-hadrons annihilated at the end of the hadron epoch, leptons and anti-leptons dominate the mass of the Universe.
- Big Bang nucleosynthesis, $t \sim 10\text{--}10^3\text{ s}$, $T \sim 10^{11}\text{--}10^9\text{ K}$. The temperature of the Universe allows protons and neutrons to combine through nuclear fusion to form the nuclei of hydrogen, helium and lithium. (e.g. Wagoner et al., 1967; Wagoner, 1973; Boesgaard & Steigman, 1985; Walker et al., 1991). The Big Bang nucleosynthesis ends when the temperature of the Universe has fallen to the threshold where nuclear fusion cannot continue.
- Photon epoch, $t \sim 10\text{--}10^{13}\text{ s}$, $T \sim 10^9\text{--}10^3\text{ K}$. After the lepton epoch when most leptons and anti-leptons have annihilated, the energy of the Universe is dominated by photons. The Big Bang nucleosynthesis happens during the first 20 minutes. Then, the Universe is filled with a hot dense plasma of nuclei, electrons and photons. As photons frequently interact with electrons, protons and nuclei, the Universe is opaque.
- Recombination, $t \sim 3.8 \times 10^5\text{ yr}$, $z \sim 1100$, $T \sim 4000\text{ K}$. Atomic nuclei and electrons bind to form neutral atoms (Peebles, 1968; Zeldovich et al., 1968). Since most electrons are captured and bound into atoms, photons can travel freely without encounter with electrons. Recombination results in radiation decoupling from matter. The Universe becomes transparent and observable from this stage. Due to the redshift, these free photons left from recombination are observed as the $T = 2.7\text{ K}$ CMB photons today.
- Dark Ages, $t \sim 0.38\text{--}150\text{ Myr}$, $T \sim 40000\text{--}60\text{ K}$. During the time between recombination and the formation of first stars, the Universe is highly neutral and the CMB photons are in the infrared band. The Universe is dominated by dark matter, and the non-linear growth of density perturbations results in the formation of dark matter haloes.
- Reionization, $t \sim 150\text{--}1000\text{ Myr}$, $T \sim 60\text{--}19\text{ K}$. During the redshift in the range of $z \sim 20\text{--}6$, the first light sources, including Population III stars, quasars and protogalaxies form, and the radiation from these light sources ionizes the neutral hydrogen in the intergalactic medium (IGM).
- Galaxy formation and evolution, $t \gtrsim 1\text{ Gyr}$. At $z \lesssim 6$, the hydrogen in the IGM is completely ionized but galaxy formation and evolution continue, where protogalaxies form and merge into massive galaxies, and protoclusters evolve into galaxy clusters and superclusters.

Since the fraction of different components (i.e., radiation, matter and dark energy, etc.) in the Universe evolve differently during the cosmic expansion (i.e., $\rho_r \propto a^{-4}$, $\rho_m \propto a^{-3}$ and ρ_Λ is a constant), the Universe is dominated by different components at different times. After inflation and until about 47 kyr the dynamics of the Universe is dominated by radiation, which is called the radiation-dominated era. After that, until $t \sim 10$ Gyr, the Universe is dominated by matter, which is called the matter-dominated era. Then the Universe is dominated by dark energy at $z \lesssim 2$.

1.2 The epoch of reionization

Before the recombination era, the Universe was filled with a hot dense plasma of nuclei, electrons and photons. Recombination happened at $z \sim 1100$ or $t \sim 380$ kyr, when the temperature of the Universe dropped to $T \approx 13.6$ eV ≈ 4000 K. During the recombination era, electrons and protons started to combine to form hydrogen atoms, and photons decoupled from baryons. The epoch after recombination and before the first ionizing sources formed is called the ‘dark ages’, during which the Universe was filled with dark matter particles, CMB photons, and neutral hydrogen (HI) and helium (HeI) gas. The dark ages ended when the first shining objects (stars or quasars) formed and started to reionize the neutral IGM. The epoch from the formation of first ionizing sources to the completion of reionization is called the epoch of reionization (EoR), which is at $z \sim 30$ –6 (e.g. Barkana & Loeb, 2001; Fan et al., 2006b).

1.2.1 Observing reionization

Gunn-Peterson test

Neutral hydrogen in the IGM can easily absorb Ly α photons which have wavelength $\lambda_\alpha = 1216$ Å. Due to redshift, shorter rest-frame wavelength photons can be absorbed when they redshift to 1216 Å. This results in an absorption trough in the spectrum of a background source. The number density of neutral hydrogen in the IGM can be therefore investigated using Ly α absorption on the spectra of distant light sources. Gunn & Peterson (1965) first proposed using the Ly α absorption feature in the spectra of distant quasars as a probe to the ionizing state of the IGM. The Gunn-Peterson optical depth to Ly α photons is

$$\tau_{\text{GP}} = \frac{\pi e^2}{m_e c} f_\alpha \lambda_\alpha H^{-1}(z) n_{\text{HI}}, \quad (1.25)$$

where e and m_e are the charge and mass of an electron respectively, f_α is the oscillator strength of the Ly α transition, $H(z) = H_0 E(z)$ is the Hubble constant (Equation 1.16) and n_{HI} is the number density of neutral hydrogen. The proper number density of neutral hydrogen atoms can be subsequently estimated using the Gunn-Peterson optical

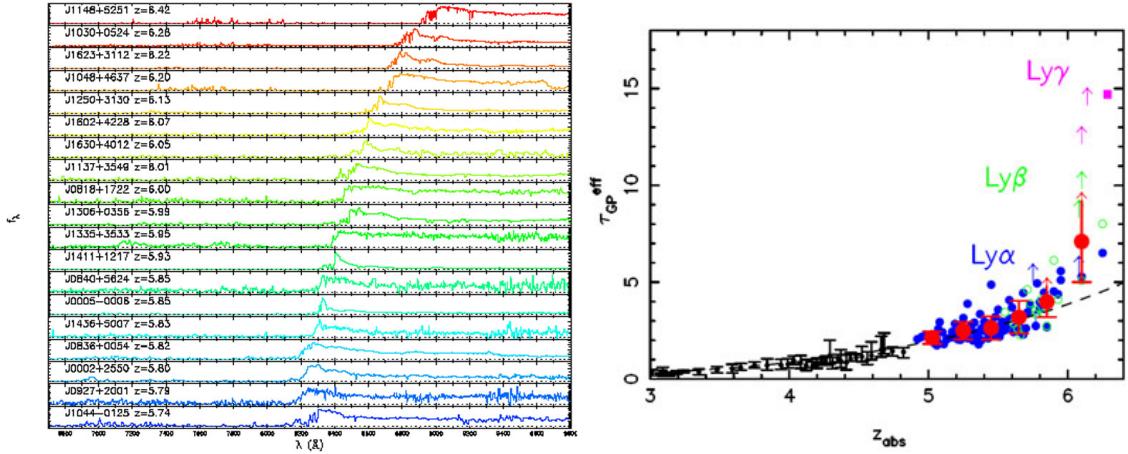


Figure 1.2: The spectra of 19 SDSS quasars at $5.74 < z < 6.42$ (left panel) and the measured Gunn-Peterson optical depth (right panel) from Fan et al. (2006b). The measured Gunn-Peterson optical depth increases rapidly at $z \gtrsim 5.7$ indicating a rapid increase of the neutral fraction in the IGM.

depth

$$n_{\text{HI}}(z) \sim 2.42 \times 10^{-11} \tau_{\text{GP}} h E(z) \text{ cm}^{-3}. \quad (1.26)$$

The observed τ can be measured from the ratio of observed flux, f_λ^{obs} , in the continuum to the intrinsic flux in the absence of absorption, f_λ^{int} . The effective optical depth $\tau_{\text{GP}}^{\text{eff}}$ is averaged over a region of the spectrum as

$$e^{-\tau_{\text{GP}}} = \langle \frac{f_\lambda^{\text{obs}}}{f_\lambda^{\text{int}}} \rangle. \quad (1.27)$$

Using a sample of high- z quasars from the Sloan Digital Sky Survey (SDSS), Fan et al. (2006b) obtained a best fit effective optical depth

$$\tau_{\text{GP}}^{\text{eff}} = (0.85 \pm 0.06) \left(\frac{1+z}{5} \right)^{4.3 \pm 0.3} \quad \text{at } z < 5.5. \quad (1.28)$$

This indicates that the IGM is highly ionized at $z < 5.5$. Fan et al. (2006b) also found that $\tau_{\text{GP}}^{\text{eff}}$ had an accelerated evolution

$$\tau_{\text{GP}}^{\text{eff}} \propto (1+z)^{10.9} \quad \text{at } z > 5.7. \quad (1.29)$$

This corresponds to a rapid increase of neutral hydrogen at $z > 5.7$. These observations indicate that $z \sim 6$ is about the end of the reionization era. Figure 1.2 shows the spectra of Gunn-Peterson quasars at $5.74 < z < 6.42$ and the measured Gunn-Peterson optical depth from Fan et al. (2006b). At $z < 5$, when global reionization has already completed, the Ly α absorption is resolved to a series of dispersed absorption lines in high-resolution

spectra called the Ly α forest (e.g. Cen et al., 1994; Rauch, 1998; Hernquist et al., 1996). The forest corresponds to a series of gas clouds containing neutral hydrogen along the line-of-sight. Observations show that the density of Ly α lines in the Ly α forest above a given threshold increases strongly with redshift (e.g. Rauch, 1998).

Due to the large optical depth of Ly α absorption, the Gunn-Peterson test is only sensitive near the end of the reionization era when the IGM is mostly ionized, and quickly saturates for a higher neutral fraction in the earlier Universe. When the Ly α flux is too weak to be detected, other absorption lines with smaller optical depths such as Ly β and Ly γ can be used instead (e.g. Fan et al., 2006b).

1.2.2 Thomson scattering of CMB photons

Another independent test of the reionization state is using the anisotropy of the CMB. Once the Universe is ionized, CMB photons can be scattered by free electrons through *Thomson scattering*. The Thomson optical depth along the line-of-sight up to a redshift z is

$$\tau_e(z) = \sigma_T \int_0^z n'_e \frac{cdl(z')}{dz'} dz', \quad (1.30)$$

where $\sigma_T = 6.652 \times 10^{-25} \text{ cm}^2$ is the Thomson scattering cross-section, dl is the proper depth corresponding to dz , and $n_e(z)$ is the number density of free electrons in proper units at redshift z along the line-of-sight. Assuming a homogeneous IGM distribution and a uniform ionizing fraction at z , $n_e(z)$ can be related to the number density of hydrogen and helium and their ionizing fractions:

$$n_e(z) = n_H(z)X_{\text{HII}}(z) + n_{\text{He}}(z)[X_{\text{HeII}}(z) + 2X_{\text{HeIII}}(z)], \quad (1.31)$$

where $n_H(z)$ and $n_{\text{He}}(z)$ are the number densities of hydrogen and helium atoms in proper units at redshift z respectively, and X_{HII} , X_{HeII} and X_{HeIII} are the fraction of hydrogen and helium in the states of HI, HeII and HeIII. Thus, the optical depth of Thomson scattering can be used to constrain reionization. The scattering effect can be observed in the CMB temperature and polarization spectra on scales comparable to the horizon size at the reionization era. A recent measurement from Planck Collaboration et al. (2016) gives $\tau_e = 0.066 \pm 0.016$.

21-cm line

The Gunn-Peterson test and CMB polarization measurements have their own limitations. The large optical depth of Ly α absorption makes the Gunn-Peterson trough unable to measure the reionization fraction precisely in the early stage of the EoR. The reionization structure probed through the Gunn-Peterson method is also limited by the number of available distant quasars. The drawback of CMB polarization is that it is an integrated measurement along the line-of-sight and does not offer discriminatory infor-

mation at different redshifts. Another proposed probe for reionization as well structure formation is using the hyperfine 21-cm line which is associated with the spin-flip transition in the ground state of neutral hydrogen atom. The energy difference between these two spin states is $E_{10} = 5.874 \times 10^{-6}$ eV, which corresponds to a radio photon with a rest-frame wavelength $\lambda_{10} = 21.106$ cm or frequency $\nu_{10} = 1420.4$ MHz, and an equivalent temperature $T_{10} = E_{10}/k_B = 68$ mK. The population ratio of the two spin states is determined by the spin temperature, T_s , through

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} \exp\left[-\frac{T_{10}}{T_s}\right]. \quad (1.32)$$

where $g_1/g_0 = 3$. The 21-cm optical depth for background photons in neutral hydrogen is

$$\tau_\nu(z) = \frac{3c^3 h_P A_{10} n_{\text{HI}}(z)}{32\pi\nu^2 k_B T_s(z) H(z)}, \quad (1.33)$$

where h_P is the Planck constant, k_B is the Boltzmann's constant, and $A_{10} = 2.85 \times 10^{-15} \text{ s}^{-1}$ is the spontaneous decay rate of the spin-flip transition. CMB photons can be used as the background photons of the 21-cm absorption. Since we are interested in the photons with frequencies, ν , much smaller than the peak frequency of CMB blackbody, the brightness temperature of the background photons, T_ν , is related to the specific intensity, I_ν , through the Ralyleigh-Jeans limit as

$$I_\nu = \frac{2k_B T_\nu}{c^2}. \quad (1.34)$$

Transferring through a medium containing neutral hydrogen, the CMB photons with temperature $T_R(\nu)$ can be observed with the brightness temperature

$$T_b(\nu) = T_S(1 - e^{-\tau_\nu}) + T_R(\nu)e^{-\tau_\nu}. \quad (1.35)$$

Since the neutral hydrogen is at high redshifts ($z \gg 1$) and this transition has a small optical depth ($\tau \ll 1$), this yields a differential brightness temperature (Pritchard & Loeb, 2012):

$$\delta T_b = \frac{T_S - T_R}{1+z}(1 - e^{-\tau_\nu}) \approx \frac{T_S - T_R}{1+z}\tau_\nu \quad (1.36)$$

$$\approx 27x_{\text{HI}}(1 + \delta_b) \left(\frac{\Omega_{b,0}h}{0.023}\right) \left(\frac{\Omega_m}{0.15}\right)^{-1/2} \left(\frac{1+z}{10}\right)^{1/2} \left(\frac{T_S - T_R}{T_S}\right) \text{ mK}, \quad (1.37)$$

where x_{HI} is the neutral fraction of hydrogen.

There are actually three processes altogether to determine T_s : (i) interactions with background radio photons from the CMB, and radio loud point sources; (ii) collisions with other atoms and electrons; (iii) the resonant scattering of Ly α photons that cause

a spin-flip via an intermediate excited state. CMB photons rapidly drive the spin temperature towards the temperature of the CMB, T_γ . Two other processes break this coupling and make the spin temperature different from the CMB. The collisional excitation/de-excitation couples the spin temperature to the kinetic temperature of gas, T_K , but it is weak in a neutral, cold medium. In addition, the Wouthuysen-Field mechanism (Wouthuysen, 1952; Field, 1958) is a coupling process in which a hydrogen atom absorbs a Ly α photon and jumps from the $n = 1$ to the $n = 2$ energy level and then spontaneously decays back to the $n = 1$ energy level but with a different spin state. The Wouthuysen-Field coupling drives the spin temperature towards the brightness temperature of the radiation field, T_α . Thus, T_s is determined by the competition of these three processes:

$$T_s^{-1} = \frac{T_\gamma^{-1} + x_\alpha T_\alpha^{-1} + x_c T_K^{-1}}{1 + x_\alpha + x_c}, \quad (1.38)$$

where x_α and x_c are the coupling coefficients for Wouthuysen-Field and collisional process respectively. During the reionization, Ultraviolet (UV) photons from ionizing sources can provide Ly α photons for the Wouthuysen-Field process. Thus, mapping the fluctuations of 21-cm intensity and spectra may provide information about the reionization state and gas distribution during the EoR (e.g. Wyithe & Loeb, 2007, 2008; Furlanetto, 2016). However, the 21-cm signal from the reionization era is very difficult to detect because it is weak, and is buried behind a much brighter foreground of extragalactic radio sources. Nevertheless, the 21-cm astronomy has great potential and could be detected by new radio telescopes such as SKA and HERA in the future (e.g. Pritchard & Loeb, 2012).

1.2.3 Sources for reionization

Photoionization is the dominant process that reionizes the IGM during the EoR. To complete reionization of hydrogen at $z \sim 6$, a significant flux of UV photons with energy larger than 13.6 eV is required. Several candidate sources for the UV photons have been proposed (e.g. Madau et al., 1999; Barkana & Loeb, 2001; Wyithe & Loeb, 2003)

- Quasars: A quasar is a luminous and compact region surrounding a super-massive black hole in the centre of massive galaxies. Quasars are very efficient at converting mass to energy. Observations show that quasars are the dominant ionizing sources of UV background photons at $z \leq 2$ (Haardt & Madau, 1996). However, the number density of luminous quasars drops exponentially towards higher redshifts (e.g. Richards et al., 2006), with the number of quasars ($M_{1450} \lesssim -27$) at $z \sim 6$ being 30 times smaller than at $z \sim 3$ (Fan et al., 2004). The small observed number of luminous quasars at $z \gtrsim 6$ likely could not reionize the IGM.
- Galaxies: Star-forming galaxies contain a large fraction of young stellar popula-

tions, in which massive (OB) stars can emit a significant number of UV photons during their short main-sequence life time. These massive stars are short-lived, so their number densities are determined by the current star-formation rate (SFR) density of galaxies. From the observed cosmic star-formation history and the emissivity of galaxies, galaxies are expected to be dominant UV contributors at $z > 3$ (Haardt & Madau, 1996). However, galaxies usually contain dust which obscures the UV radiation and re-emits the energy in infrared band. If we assume high redshift galaxies contain similar amounts of dust with their low-redshift analogs (e.g. Ma et al., 2015), then galaxies above current detection limits are not sufficiently numerous to maintain reionization. Therefore we expect there are more faint galaxies which exist below current detection limits (e.g. Robertson et al., 2013, 2015; Bouwens et al., 2015a). Studying the abundance and ionizing photon contribution of undetected galaxies is one of the main tasks in this thesis (see Chapter 3)

- Population III stars: These are extremely metal-poor stars which form in the very early Universe. Since these stars form in metal-free environments, they may have very high stellar masses (e.g. up to $300 M_{\odot}$) and provide UV photons at very high redshifts. Population III stars may play an important role in reionization at $z > 10$ (e.g. Cen, 2003; Sokasian et al., 2004). However, there is currently no direct observational evidence for Population III stars, their abundance is instead inferred from simulation and theory.

1.3 Structure Formation

Our Universe is homogeneous on cosmological scales, but it has structures on smaller scales. The structure of matter can be traced by the spatial distribution of observed galaxies, although most of the matter is unseen dark matter.

The existence of dark matter is an important discovery in modern cosmology. The most straightforward supporting evidence of dark matter comes from rotation curves of disk galaxies (Rubin et al., 1980, 1982). If the visible ordinary (baryonic) mass are considered as the total mass of galaxies, their rotation curves will drop as $v_c \propto r^{-1/2}$ at large radii according to Newtonian dynamics. However, flat rotation curves are observed for galaxies, which indicate extended dark matter haloes around galaxies (e.g. Persic et al., 1996; de Blok et al., 2008). Current observations show that dark matter contributes about 85 per cent of the total mass of the Universe (e.g. Planck Collaboration et al., 2016). Dark matter particles are often considered to be a kind of weakly interacting massive particle, although their physical nature is still a mystery. According to the thermal velocities of candidate dark matter particles, they can be classified into "hot" dark matter (HDM), "warm" dark matter (WDM) and "cold" dark

matter (CDM). Supported by observations and cosmological simulations, the Λ CDM cosmology has been firmly established as the current standard model of cosmology (White & Rees, 1978; Peebles, 1982; Blumenthal et al., 1984; Davis et al., 1985).

In Λ CDM cosmology, structures form from small initial perturbations through gravitational instability (Jeans, 1902). This section gives a brief introduction to the structure formation model in the Λ CDM cosmology (see Mo et al., 2010, for more details).

1.3.1 Density perturbations

In current structure formation models, the initial perturbations are assumed to be adiabatic and to originate from quantum fluctuations through inflation (Guth, 1981). In adiabatic initial conditions, matter and radiation are tightly coupled and perturbed in the same way. The total density spatially varies but the ratio of photons to baryons is spatially constant. Before photons decouple from matter, small scale baryonic perturbations are damped by photon diffusion through Silk damping (Silk, 1968). After recombination, perturbations with scales above the Jeans length can grow through gravitational instability.

The density perturbation field at position \mathbf{x} and time t is defined using the relative overdensity

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (1.39)$$

where the $\rho(\mathbf{x}, t)$ is density of mass at (\mathbf{x}, t) and $\bar{\rho}(t)$ is the mean background density at t . For structures with sizes much smaller than the horizon size, and with small perturbations ($\delta \ll 1$), the evolution of the Universe can be described by a linearised non-relativistic fluid. The time evolution of the fluid is given by the equation of continuity describing mass conservation, the Euler equations which are the equation of motion, and the Poisson equation describing the gravitational field. In comoving coordinates these are

$$\text{Continuity: } \frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla \cdot [(1 + \delta) \mathbf{v}] = 0; \quad (1.40)$$

$$\text{Euler: } \frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a} \mathbf{v} + \frac{1}{a} (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{a} \nabla \phi; \quad (1.41)$$

$$\text{Poisson: } \nabla^2 \phi = 4\pi G \bar{\rho} a^2 \delta, \quad \text{where } \phi \equiv \Phi - \frac{2}{3} \pi G \bar{\rho} a^2 x^2. \quad (1.42)$$

The linear approximation is obtained by keeping only the first order terms in δ and \mathbf{v} :

$$\frac{\partial \delta}{\partial t} + \frac{1}{a} \nabla \cdot \mathbf{v} = 0, \quad \frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a} \mathbf{v} + \frac{1}{a} \nabla \phi = 0. \quad (1.43)$$

Solving the above equations, we obtain

$$\frac{\partial^2 \delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t} = 4\pi G\bar{\rho}\delta. \quad (1.44)$$

The solution for this equation has both a growing mode and a decaying mode. We may neglect the decaying mode after a relatively long time from the initial perturbations, leaving only the growing mode:

$$\delta(\mathbf{x}, t_2) = \delta(\mathbf{x}, t_1) \frac{D(t_2)}{D(t_1)}. \quad (1.45)$$

Here $D(t)$ is the *growth factor*. It can be calculated at redshift z as using

$$D(z) = \frac{H(z)}{H_0} \int_z^\infty \frac{(1+z')dz'}{H^3 z'} \quad (1.46)$$

This implies $D \propto a(t) \propto t^{2/3}$ for $\Omega_m \approx 1$ during the matter dominated era ($t \sim 47$ kyr to 10 Gyr). Linear perturbations grow roughly as a power law with time rather than an exponential, so they do not forget their initial conditions.

The perturbation fields can be written using their Fourier transforms in comoving units.

$$\delta(\mathbf{x}, t) = \int \delta_{\mathbf{k}}(t) \exp(2\pi i \mathbf{k} \cdot \mathbf{x}) d^3k; \quad \delta_{\mathbf{k}} = \int \delta(\mathbf{x}, t) \exp(-2\pi i \mathbf{k} \cdot \mathbf{x}) d^3x, \quad (1.47)$$

where $k = 2\pi/\lambda$ is the wavenumber on the scale of λ .

The initial perturbations are usually assumed to be Gaussian in which different \mathbf{k} -modes are statistically uncorrelated. In a spatially isotropic and homogeneous universe, the statistics of a Gaussian random perturbation field is completely determined by its power spectrum

$$P(k, t) = \langle |\delta(k, t)|^2 \rangle = P_i(k) T^2(k) D^2(t), \quad (1.48)$$

where $T(k)$ is the transfer function which has the different form for scales relative to the horizon scale at the time of matter-radiation equality, and P_i is the initial power spectrum of the field. In the standard cosmological model, the initial power spectrum is assumed to be a power law $P_i(k) \propto k^{n_s}$, where n_s is the power spectral index. Determining the shape and amplitude of the linear power spectrum is one of the most important tasks of observational cosmology. Recent observations give $n_s = 0.968 \pm 0.006$ (Planck Collaboration et al., 2016).

The power spectrum is related to the variance of the density field smoothed in

randomly placed spheres of radii, R , by

$$\sigma^2(R) = \frac{1}{2\pi^2} \int P(k) \widetilde{W}_R^2(k) k^2 dk, \quad (1.49)$$

where

$$\widetilde{W}_R(k) = \frac{3}{(kR^2)} [\sin(kR) - kR \cos(kR)] \quad (1.50)$$

is the Fourier transform of the real-space spherical top-hat window function

$$\widetilde{W}_R(r) = \begin{cases} 3/(4\pi R^3) & \text{if } r < R \\ 0 & \text{otherwise.} \end{cases} \quad (1.51)$$

We therefore can normalize the theoretical power spectrum using $\sigma_8 \equiv \sigma(R=8h^{-1}\text{Mpc})$. Recent observations of CMB give $\sigma_8 = 0.815 \pm 0.009$ (Planck Collaboration et al., 2016). In observations, the variance of matter fluctuations, σ_m , can be derived from the variance of the distribution of galaxies, σ_{gal} , with a *bias parameter*, b , using

$$\sigma_m(R=8h^{-1}\text{Mpc}) = \frac{\sigma_{\text{gal}}(R=8h^{-1}\text{Mpc})}{b}. \quad (1.52)$$

The derived linear bias between matter fluctuations and the galaxy distribution arises from the fact that galaxies form preferentially in high density regions.

1.3.2 Dark matter haloes

At early times, the density perturbation is small, and the growth of the perturbation is in the linear regime. Once the local overdensity grows to $|\delta| \sim 1$, the fluctuation growth transfers to the more complex nonlinear regime. Nonlinear dynamics is difficult to solve analytically and usually relies on numerical solutions. For a spherically symmetric mass distribution, an analytic solution can still be found. According to Birkhoff's theorem, the gravitational acceleration of a spherically symmetric system depends only on the mass enclosed within the sphere. The motion of the system can be described by the Newtonian equation

$$\frac{dr}{dt^2} = H_0^2 \Omega_{\Lambda,0} r - \frac{GM}{r^2}, \quad (1.53)$$

where r is the radius of the system, and M is the mass enclosed within the radius. The first term on the right-hand side is a correction for the gravitational acceleration due to the cosmological constant. The initial velocity field is given by the Hubble flow. Initially, the overdensity grows in accordance with linear theory. However, the system will reach a maximum radius at which $dr/dt = 0$ and then start to collapse to a point. The overdensity at the time of collapse predicted by linear perturbation theory is $\delta_c \approx 1.686$ (e.g. Weinberg, 2008).

For a collapsing dark matter system, even a slight perturbation can break the symmetry and prevent the system from collapsing to a point. During the collapse, dark matter particles will gain kinetic energy by losing potential energy. Dark matter particles are considered to be collisionless systems, and so do not dissipate energy. Further, the tidal interaction between dark matter substructures will convert the kinetic energy to random motion. The dark matter system eventually relaxes to a quasi-static structure supported by random motions. An overdense region will collapse until it approximately achieves the state of *virial equilibrium* where the total gravitational potential energy U of dark matter particles balances their total kinetic energy K such that $2K + U = 0$. These nonlinear, quasi-equilibrium dark matter structures are called *dark matter haloes*. According to the spherical collapse model (e.g. Gunn, 1977; Bertschinger, 1985), the mean density contrast of dark matter haloes at virialization in a flat, matter dominated universe is $\Delta_c = 18\pi^2$ ($\Delta_c = 200$ is sometimes adopted for estimation). For a flat universe with $\Omega_m + \Omega_\Lambda = 1$, a useful approximation is

$$\Delta_c \approx 18\pi^2 + 82x - 39x^2, \quad (1.54)$$

where $x = \Omega(z) - 1$ (Bryan & Norman, 1998).

It is possible to estimate the mass distribution of dark matter haloes by combining the Gaussian random perturbations, the linear gravitational growth and spherical collapse. We can calculate the smoothed density field δ_s from the Gaussian random field using spherical filters with different characteristic radius R . If the density within the sphere is greater than $\delta_c(t)$, an associated mass of M will collapse and form a halo. Press & Schechter (1974) gave the number density $n(M, t)$ of dark matter haloes with masses between m and $m + dm$:

$$n(M, t)dm = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma} \exp\left(-\frac{\delta_c^2}{2\sigma^2}\right) \left| \frac{d \ln \sigma}{d \ln M} \right| dm. \quad (1.55)$$

Within this equation, a factor of 2 has been adopted to include the mass contribution of underdense regions enclosed in larger collapsed objects (i.e., the ‘cloud-in-cloud’ problem). Equation 1.55 is known as the Press-Schechter (PS) mass function. An alternative derivation of the PS formalism was presented by Bond et al. (1991) using the excursion set formalism, which is known as the extended Press-Schechter (EPS) formalism.

Both analytic calculations and simulations show that small dark matter haloes at early times can merge into a bigger halo at a later time. Alternatively, a halo at the present time t_0 had a number of progenitor haloes at $t < t_0$, and the mass of the present halo was distributed over many small branches at $t < t_0$. The assembly history of dark matter haloes are usually illustrated using dark matter halo *merger trees*. An illustration of a merger tree is shown in Figure 1.3. Dark matter merger trees are widely adopted

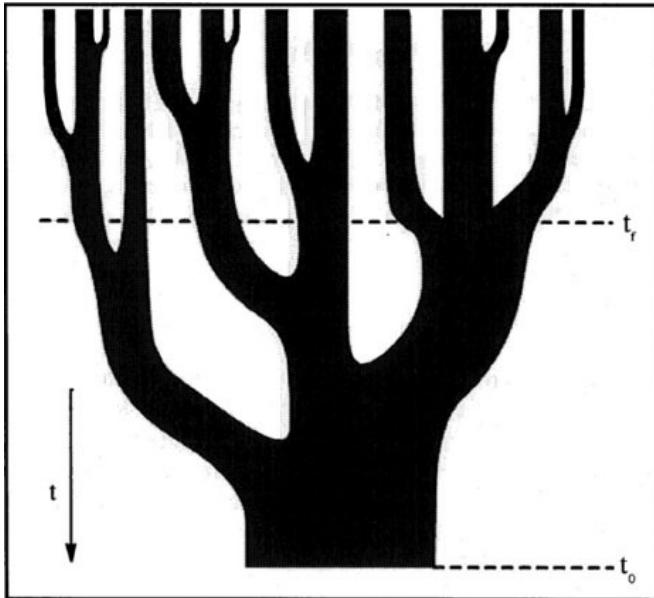


Figure 1.3: An illustration of a merger tree. Time increases from top to bottom, and the widths of the tree branches represent the mass of individual haloes. A horizontal slice through the tree, such as $t = t_f$ gives the distribution of the masses of progenitor haloes at a given time (figure adapted from Lacey & Cole, 1993).

in cosmological numerical simulations, and they play an important role in structure evolution as well as galaxy formation and evolution. In hierarchical galaxy formation models, dark matter haloes provide the environment for galaxy formation, and galaxies assemble along the merger trees of their host haloes.

1.3.3 Internal properties of haloes

As mentioned above, a dark matter halo is defined as a collapsed matter sphere in virialization which has a mean density of Δ_c times the background mass density. The radius of the virialized system is also called *virial radius*, represented as R_{vir} . The mass enclosed within the virial radius is called the *virial mass*:

$$M_{\text{vir}} = \frac{4\pi R_{\text{vir}}^3}{3} \Delta_c \bar{\rho}, \quad (1.56)$$

where $\bar{\rho}$ is the mean density of the Universe. The circular velocity of a dark matter halo at R_{vir} is also called the *virial velocity*:

$$V_{\text{vir}} = \sqrt{\frac{GM_{\text{vir}}}{R_{\text{vir}}}}. \quad (1.57)$$

These virial parameters are widely used in numerical simulations to characterize dark matter haloes even those with non-spherical shape.

Inside a dark matter halo, the internal mass distribution is described by a density profile, $\rho(r)$, which is function of radius r . The simplest model usually assumes dark matter haloes are truncated singular isothermal spheres which have $\rho(r) \propto r^{-2}$. However, dark matter haloes are complex, nonlinear, quasi-equilibrium systems, so that detailed dynamics need to use numerical solutions. Using N-body simulations with CDM cosmologies, Navarro et al. (1996) found the density profile can be described by

$$\rho(r) = \rho_{\text{crit}} \frac{\delta_{\text{char}}}{(r/r_s)(1+r/r_s)^2}, \quad (1.58)$$

where δ_{char} is a characteristic overdensity, r_s is a scale radius, and δ_{char} and r_s vary from halo to halo. This density profile is known as the NFW profile. The logarithmic slope of the NFW profile changes from -1 near the centre to -3 at large radii, and only resembles isothermal -2 at $r \sim r_s$. Navarro et al. (1997) found that the density profile of dark matter haloes seems universal for CDM-like cosmologies since it is independent of the halo mass, the initial density perturbation spectrum and the values of cosmological parameters. The ratio of the virial radius and the NFW scale radius of a dark matter halo measures the concentration of the density profile, and so that is called the concentration parameter, represented as $c = r_{\text{vir}}/r_s$.

Another important property of a dark matter halo is its angular momentum. The collapse of dark matter is usually asymmetric, and produces haloes with significant angular momentum. This can be parameterized through a dimensionless *spin parameter*, traditionally defined as (e.g. Efstathiou & Jones, 1979; Fall & Efstathiou, 1980)

$$\lambda = \frac{J|E|^{1/2}}{GM^{5/2}}, \quad (1.59)$$

where J , E , and M are the total angular momentum, energy and mass of the dark matter halo respectively. These quantities are conserved for an isolated, dissipationless dynamical system. In numerical simulations, the calculation of the energy of the halo is computationally difficult. An alternative and more practical spin parameter was given by Bullock et al. (2001):

$$\lambda = \frac{J}{\sqrt{2}M_{\text{vir}}V_{\text{vir}}R_{\text{vir}}}. \quad (1.60)$$

The Bullock et al. (2001) spin parameter is very similar to the traditional spin parameter for NFW haloes, and it reduces to the traditional spin parameter when measured at the virial radius for a truncated singular isothermal halo.

1.3.4 Cosmological simulations

The non-linear processes of structure formation is complicated, and effects such as violent relaxation, dynamical friction and tidal torques are naturally included. A detailed

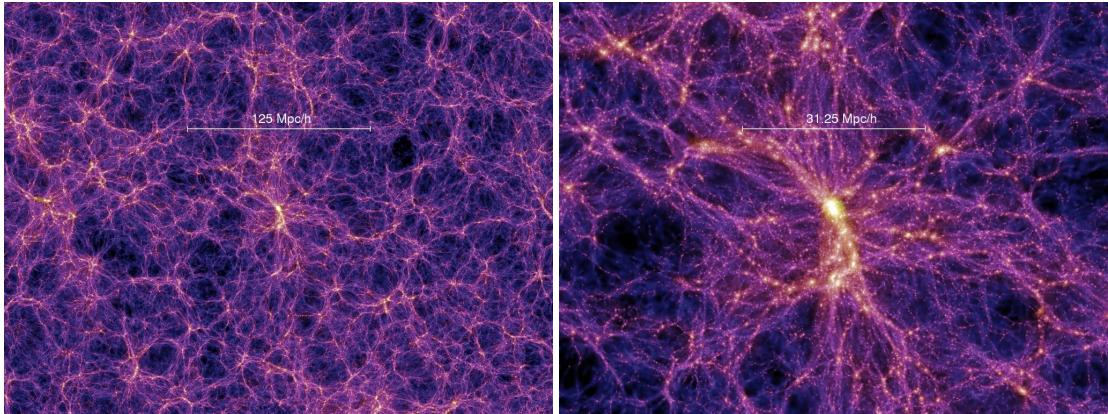


Figure 1.4: Dark matter structures from the *Millennium* N-body simulation at $t \sim 4.7$ Gyr on different scales (Springel, 2005). Different colours represent different mass densities. Simulations show that dark matter haloes are clustered on scales of $\lesssim 100$ Mpc. On larger scales, the distribution of dark matter shows a web structure which is called the cosmic web.

and precise description must therefore involve a numerical approach. Thanks to powerful modern computers, large-scale simulations have become an indispensable tool in cosmological studies. In numerical simulations, the mass distribution of a cosmic region is represented by particles or grids. The change of the mass distribution is calculated by tracing the applied forces and motion of every particle or mass element. The mass represented by a single particle is called the mass resolution of the simulation. N -body simulations and hydrodynamical simulations are two kinds of widely used cosmological simulation. The former calculates only the gravitational interaction of collisionless dark matter particles, whilst the latter simultaneously calculates the gravitational interaction of dark matter and baryonic particles as well as solves the hydrodynamical equations for the gas.

The initial conditions (in terms of both position and velocity) of the simulation particles are usually set using the observed set of cosmological parameters (e.g. $h, \Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s$) with a cosmological model such as Λ CDM. The boundary conditions are usually chosen to be periodic. Due to the limitation of resolution and numerical parameters (e.g. gravitational softening), the simulation results must be checked and calibrated by comparing the statistics with observations or other higher resolution simulations.

The scale of a simulation can be represented by its particle number, N , and the simulation volume. Current cosmological simulations usually simulate billions of particles within a cubic region with side length, L , in units of Mpc. Since N -body simulations only calculate the gravitational interaction between dark matter particles, they can be run with much larger numbers of particles compared to hydrodynamical simulations. However, hydrodynamical simulations can calculate the baryonic effects on

dark matter structure formation, and simulate the formation and evolution of galaxies. There are already many successful cosmological simulations. For instance, the *Millennium* N -body simulation has $N = 2160^3$ and $L = 500 h^{-1} \text{Mpc}$ with mass resolution of $8.6 \times 10^8 h^{-1} M_\odot$ (Springel, 2005). The *EAGLE* hydrodynamical simulation has $N = 1504^3$ and $L = 100 \text{Mpc}$, with dark matter mass resolution of $9.7 \times 10^6 M_\odot$ and baryonic matter resolution of $1.8 \times 10^6 M_\odot$ (Schaye et al., 2015). The *Illustris* hydrodynamical simulation ran with $N = 2 \times 1820^3$ and $L = 106.5 \text{Mpc}$, and with dark matter mass resolution of $6.26 \times 10^6 M_\odot$ and initial baryonic matter mass resolution of $1.26 \times 10^6 M_\odot$ (Vogelsberger et al., 2014; Genel et al., 2014).

In N -body simulations, the most straightforward way to calculate the gravitational force on a particle i is the directly summation of particle-particle interactions

$$F_i = - \sum_{i \neq j} \frac{Gm^2(\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^3}. \quad (1.61)$$

where m and \mathbf{r} are the mass and position of simulation particles. To prevent close encounters between simulation particles which may result in unrealistic large-angle scattering events, a force softening algorithm has been introduced. This is usually accomplished by adding a finite softening length, ϵ , to the particle separation when calculate gravitational forces (i.e. change $\mathbf{r}_j - \mathbf{r}_i$ in Equation 1.61 to be $\mathbf{r}_j - \mathbf{r}_i + \epsilon$).

The algorithm for evaluating the forces through pairwise calculation as shown in Equation 1.61 is called the Particle-Particle (PP) algorithm. The computational cost of the PP algorithm is proportional to N^2 (i.e. has computational complexity $O(N^2)$). To reduce the computational expenditure, a number of algorithms have been used (see Hockney & Eastwood, 1988):

1. Tree algorithm. This hierarchical algorithm groups the particles and replaces the force from each group by its centre of mass. The force from each group can also be calculated using a multipole expansion, and only a small number of multipole components need to be kept for distant groups. This algorithm can reduce the computational complexity to $O(N \log N)$ (Barnes & Hut, 1986).
2. Particle-Mesh (PM) algorithm. This algorithm divides the computational box to a grid of meshes and calculates the forces at each meshpoint from the gravitational potential by solving the Poisson equation. The forces on particles are then evaluated by interpolating the forces at meshpoints to the positions of particles. PM algorithm can reduce the computational complexity to $O(N)$.
3. Particle-Particle/Particle-Mesh (P^3M) algorithm. This is a hybrid algorithm which calculates the short-range forces using the PP method and calculate the long-range forces using the PM method.

Cosmological simulations calculate the spatial and velocity distribution of simulation particles. The collapsed structures (dark matter haloes and/or galaxies) can be identified by a collection of particles which are locally dense and gravitationally bound. A collapsed object typically needs more than ~ 100 simulation particles to be reliably identified, and needs more than ~ 1000 particles to provide a reliable morphology and density profile.

Figure 1.4 shows the dark matter structures from the *Millennium* N-body simulation (Springel, 2005) at $t \sim 4.7$ Gyr on different scales. We can see that dark matter haloes are clustered on scales of $\lesssim 100$ Mpc. On larger scales, the distribution of dark matter shows a complex web of filaments and sheets surrounding large empty voids called the *cosmic web*.

1.3.5 Semi-analytic models

As mentioned in Section 1.3.4, hydrodynamical simulations can be used to study the formation and evolution of galaxies by tracing the density clumps of baryonic particles. However, they are computationally expensive because they have to solve the gravitational and gas dynamics simultaneously. A galaxy has stellar mass which is about two magnitudes smaller than its host dark matter halo, so that to obtain a realistic description of galaxies, a higher mass resolution for baryonic particles is needed. This also increases the computational expenditure and limits the scales of hydrodynamical simulations. Motivated by the hierarchical model of structure formation, we can separate the large-scale structure formation, dominated by dark matter, from galaxy formation within these dark matter structures to save the computational cost. Such efforts are referred to as semi-analytic models.

A typical semi-analytic model consists of numerical and analytic parts (see Baugh, 2006, for a detailed description):

1. N-body simulations are used to calculate the formation and merger history of dark matter haloes and generate the dark matter halo merger trees;
2. The interplay amongst the baryonic, stellar and gas components are modelled analytically upon these dark matter halo merger trees;
3. The observable features, such as the luminosities and colours are evaluated by stellar population synthesis (e.g. Bruzual & Charlot, 2003; Gonzalez-Perez et al., 2014).

When compared with hydrodynamical simulations, semi-analytic models are far less computationally expensive, since they process the dark-matter-dominated large-scale structure formation and computationally intensive baryon-dominated galaxy formation

separately. On the other hand, semi-analytic models sacrifice a detailed description of the gas physics, and rely on empirical or idealized laws for galaxy formation properties.

1.4 Galaxies and Observables

1.4.1 Jeans length

We can apply the equations of motion of perturbation growth to a pure baryonic universe. However, the force of baryonic pressure must be taken account. If the evolution is adiabatic, the *Jeans length* in a baryonic gas with mean density $\bar{\rho}$ and sound speed c_s is

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\bar{\rho}}}. \quad (1.62)$$

The Jeans length expresses the distance that a sound wave can travel in a gravitational free-fall time. For the perturbations with scales smaller than the Jeans length ($\lambda < \lambda_J$), the sound wave has enough time to react the perturbations and builds up pressure to resist to the gravitational infall, so that the perturbations will not grow. The perturbations with scales larger than the Jeans length ($\lambda > \lambda_J$) can grow through gravitational instability. We can define the mass within a sphere of radius $\lambda_J/2$ as the *Jeans mass*.

$$M_J = \frac{\pi}{6} \bar{\rho} \lambda_J^3. \quad (1.63)$$

For a matter dominated universe containing both baryonic and dark matter, the density perturbations of baryons and dark matter are expected to be equal in the linear regime, except for perturbations smaller than the Jeans length. In the non-linear regime, the dark matter dominates the large scale structure formation, and baryons react to the gravitational field of dark matter. As the density fluctuations grow and dark matter turns around from Hubble expansion and collapses, the hydrostatic forces start to play an important role in the dynamics. Gas can be shocked and heated, and can cool, fragment and form stars. Formed stars will further release mass, energy and chemical elements into the surrounding gas and affect further star formation. These processes result in a different distribution of baryons from dark matter.

1.4.2 Hierarchical galaxy formation

In hierarchical structure formation models (White & Rees, 1978; White & Frenk, 1991), density perturbations and large-scale structure formation are dominated by collisionless and dissipationless dark matter particles, and galaxies subsequently form within these dark matter haloes. Below we provide a brief overview of hierarchical galaxy formation, which underpins semi-analytic models (e.g. White & Frenk, 1991; Kauffmann et al., 1993; Cole et al., 1994; Kauffmann et al., 1999; Cole et al., 2000; Croton et al., 2006; Bower et al., 2006; Somerville et al., 2008; Lacey et al., 2011, 2015).

Baryonic infall

Within a dark matter halo, the baryonic gas collapses and is assumed to be heated by accretion shocks. This produces a hot gaseous halo with temperature equal to the virial temperature of the dark matter halo

$$T_{\text{vir}} = \frac{1}{2} \frac{\mu m_{\text{H}}}{k_{\text{B}}} V_{\text{vir}}^2, \quad (1.64)$$

where m_{H} is the mass of a hydrogen atom and μm_{H} is the mean molecular weight of the gas. Although universal baryon fraction within haloes, $f_b = \Omega_b/\Omega_m$, is often assumed, processes such as photo-evaporation due to the UV background and gas ejection from feedback can modify the amount of gas infall within haloes.

Gas cooling

Gas can cool from the hot halo through a series of cooling mechanisms depending on the temperature, density and chemical composition of the gas (e.g. Kauffmann et al., 1993; Baugh, 2006):

- (i) Inverse Compton scattering (Compton cooling) of CMB photons. This is effective if the temperature of background photons is much smaller than the temperature of electrons ($T_{\gamma} \ll T_e$). For $\Omega_{m,0} = 0.3$, the time scale of Compton cooling exceeds the age of the Universe at $z \lesssim 6$. Therefore Compton cooling can only be important for gas in the early Universe.
- (ii) The radiation from the combination of electrons and ions or the decay of excited atoms (atomic cooling). Collisions between atoms and electrons can ionize or excite the hydrogen atoms. The subsequent recombination or de-excitations will emit photons and cool the gas. This mechanism is temperature sensitive, and is important for haloes with virial temperature in the range of $10^4 \text{ K} < T_{\text{vir}} < 10^6 \text{ K}$.
- (iii) The bremsstrahlung radiation due to the acceleration of electrons as they encounter atomic nuclei in an ionized plasma. This is the dominant radiation process in massive haloes with $T_{\text{vir}} > 10^7 \text{ K}$.
- (iv) The radiation from the excitation of rotational/vibrational levels of molecules (molecular cooling). This cooling mechanism is less efficient and only important in haloes with $T_{\text{vir}} < 10^4 \text{ K}$.

Among these mechanisms, (ii) and (iii) are the primary cooling mechanisms relevant to the formation of galaxies. In massive haloes with $T_{\text{vir}} > 10^7 \text{ K}$, gas is fully collisionally ionized and the dominant cooling mechanism is bremsstrahlung radiation. For haloes with temperature in the range of $10^4 \text{ K} < T_{\text{vir}} < 10^6 \text{ K}$, the atomic cooling is

the most efficient cooling mechanism. Below $\sim 10^4$ K, gas is almost completely neutral, bremsstrahlung and atomic cooling are strongly suppressed. Although molecular cooling is still possible, it is less efficient and the molecular hydrogen can easily be photo-dissociated by trace amounts of star formation. In the standard model of galaxy formation, only the gas in haloes with $T_{\text{vir}} \gtrsim 10^4$ K can efficiently cool, which is associated with a halo mass of $M_{\text{vir}} \sim 10^8 M_{\odot}$ (Barkana & Loeb, 2001). The dynamical process of the cooling gas depends on the competition between the cooling time and the local dynamical time. As the gas cools, the pressure of the gas cannot support the gravity and it falls in to the centre of the potential well of the halo. If the cooling time is shorter than the dynamical time of the halo, gas will rapidly cool. Otherwise, if the cooling time is sufficiently long then a quasi-static hot gas halo will form.

Star formation

As cold gas falls into the potential well of the dark matter halo, it eventually becomes the interstellar medium (ISM) of galaxies. The cold gas can form gas clouds or gravitationally supported discs if the angular momentum is conserved. For these cold gas objects, the baryonic self-gravity becomes important, or dominant, due to the relatively small scales. With more and more cold gas accreted, the cold gas will further collapse and fragment, and stars will form. Disc instabilities and turbulences induced by sources such as supernova (SN) explosions and galaxy mergers can also trigger star formation. An important empirical relation relating the gas to the global SFR is the *Kennicutt-Schmidt law* (Schmidt, 1959; Kennicutt, 1998b). The relation was first proposed by Schmidt (1959), who found a relation between the star formation rate per unit area of a galaxy, $\dot{\Sigma}_*$, and surface density of cold gas, Σ_{gas} to be a power law:

$$\dot{\Sigma}_* \propto \Sigma_{\text{gas}}^N. \quad (1.65)$$

Kennicutt (1998b) verified this relation using both normal spiral and starburst galaxies and found a best fit of $N \sim 1.4 \pm 0.15$. In analytic models, the simplest way to estimate the global SFR is

$$\dot{m}_* \propto \frac{m_{\text{cold}}}{t_{\text{SF}}}, \quad (1.66)$$

where m_{cold} is the amount of cold gas which is available to form stars, and t_{SF} is the time scale of star formation, which is usually taken to be the dynamical time of the galaxy.

The mass distribution of formed stars is determined by the initial mass function (IMF). The empirical IMF is often expressed in terms of single or multi-band power laws (e.g. Salpeter, 1955; Kroupa, 2001, 2002; Chabrier, 2003, 2005). The IMF is important in a cosmological sense because it determines the number density of massive stars which

can emit ultraviolet (UV) flux to ionize the IGM, and the number of supernovae which provide feedback, thus altering further star formation.

Feedback

Observations show that only a small fraction of baryons are in the form of stars, and that the observed luminosity function of galaxies drops at the both bright and faint ends. Therefore, a mechanism is needed to suppress the gas cooling rate or the star formation rate. Such mechanisms are called feedback. Generally, two feedback mechanisms play an important role in suppressing star formation in galaxies. The first one is supernova feedback. Massive stars ($m_* > 8M_\odot$) end in a violent supernova explosion. The radiation and blast waves from these supernovae may heat the interstellar medium or blow it out of the galaxy. Supernova feedback is particularly important in low-mass galaxies (e.g. Kauffmann et al., 1999; Somerville & Primack, 1999). Another important feedback is provided by active galactic nuclei (AGN) which are powered by accretion onto supermassive black holes (Rees, 1984) in the centres of massive galaxies. AGN are very luminous and can suppress gas cooling and galaxy formation. Studies show that AGN feedback may be dominant in massive galaxies (e.g. Croton et al., 2006). Due to the feedback, gas can be reheated or blown into the hot gaseous halo, from which it may subsequently cool again. If the feedback is strong enough, the reheated gas can be ejected from the halo. The ejected gas may later be reincorporated into its current halo, or could be re-accreted when the halo merges with a more massive system. Apart from the internal feedback within galaxies, the UV background (UVB) radiation produced by ionizing sources can also heat the ISM and reduce the infall of gas into the potential wells of small haloes (e.g. Efstathiou, 1992; Gnedin & Ostriker, 1997; Sobacchi & Mesinger, 2013a,b).

Galaxy interactions

In the hierarchical scenario of structure formation, dark matter haloes and their hosted galaxies frequently interact with each other. Galaxy mergers are one of the most important interactions. In semi-analytic models, after halo progenitors merge, it is common for the more massive galaxy to be placed at the centre of the newly formed halo. Any hot gas that cools would fall directly into the central galaxy. The less-massive galaxy progenitor becomes a satellite orbiting the central galaxy. Due to dynamical friction, the satellite will eventually merge with the central galaxy. The dynamical friction time can be used to estimate the time scale of the orbital decay of the satellite.

Galaxy mergers play an important role in the build up of galaxy mass, both through hierarchical mass assembly and induced star formation. Galaxy mergers result in strong turbulence in the ISM, and often trigger a burst of star formation. According to the mass ratio of galaxy progenitors $q = M_1/M_2$ (where $M_1 < M_2$), galaxy mergers are classified

into major mergers (often $q \geq 1/3$) and minor mergers ($q < 1/3$). Major mergers are more violent and can significantly change the morphology and the density profile of galaxies. However, minor mergers are more prevalent and could be more important for galaxy mass growth compared to major mergers.

1.4.3 Galaxy observables

Stellar mass plays a fundamental role in galaxy formation, but it is not directly observable. To constrain galaxy formation models, we need to investigate the observable properties (observables) of galaxies. Galaxies have rich observables including luminosity, colour, morphology, size, metallicity and clustering. These observables provide not only constraints on models, but are also important probes of large-scale structure and cosmology. Galaxy observables in the early Universe are also important for constraining models of reionization and understanding which objects are important ionizing photon emitters (see Section 1.2).

Observing galaxies

We observe a galaxy through its radiation. This radiation can be characterized by a spectral energy distribution (SED), f_λ . For historical reasons, the luminosity of a galaxy, L , is conventionally represented by its magnitude. One of the most commonly used monochromatic magnitude systems is the AB magnitude (Oke & Gunn, 1983):

$$M_{\text{AB}} = -2.5 \log_{10} f_\nu - 48.6, \quad (1.67)$$

where f_ν is the photon flux in units of $\text{erg s}^{-1} \text{Hz}^{-1} \text{cm}^{-2}$ at frequency ν . In the high- z galaxy surveys, broadband photometry is commonly used to obtain a higher signal-to-noise ratio. The AB magnitudes observed through a photometric filter can be defined as (Fukugita et al., 1996):

$$M_{\text{AB}} = -2.5 \frac{\int d(\log_{10} \nu) f_\nu S_\nu}{\int d(\log_{10} \nu) S_\nu} - 48.6, \quad (1.68)$$

where S_ν is the filter function which describes the system quantum efficiency. A filter function can be characterized by its effective wavelength and bandwidth. The SEDs of photometrically observed galaxies are usually obtained by fitting the photometric luminosities with the standard library of SEDs obtained from observed galaxies or through stellar population synthesis models. The photometric colours of galaxies are usually evaluated by comparing the magnitudes in different bands, which corresponds to the slope of the SED. In the meanwhile, the colours or the slope of SEDs are determined by both the star-formation history and the dust content of galaxies.

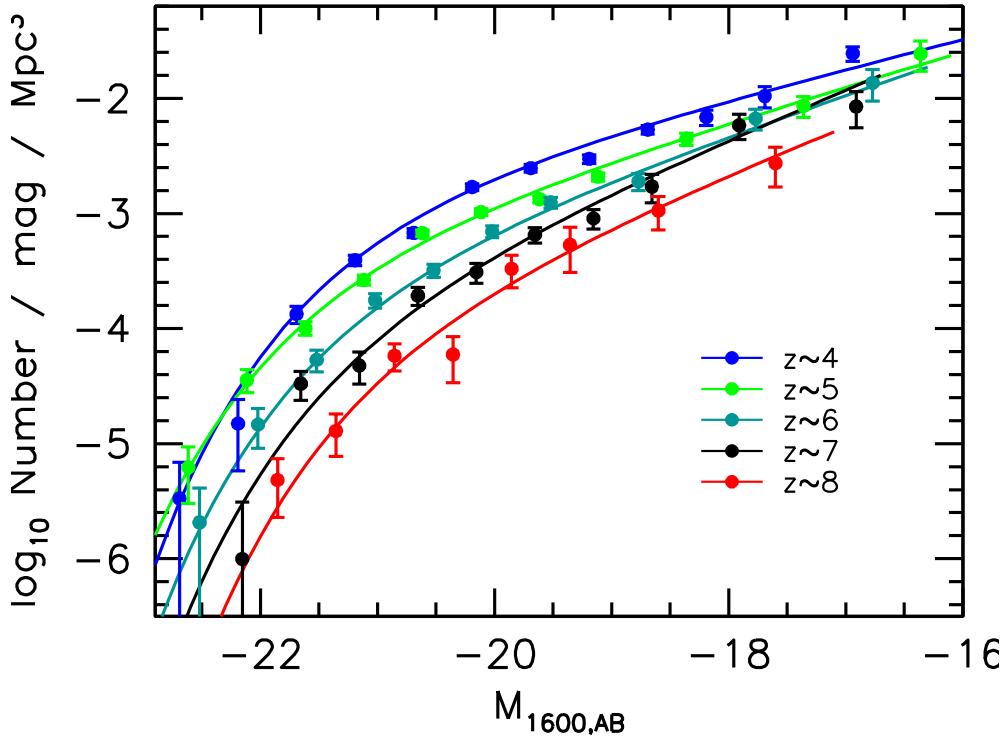


Figure 1.5: Observed UV LFs (solid circles) and associated Schechter fits (solid lines) at $z \sim 4\text{--}8$ from Bouwens et al. (2015b). $M_{1600,\text{AB}}$ represents the observed AB magnitude at wavelength 1600 Å. The observed LFs show the clear evidence for the buildup of luminosity and number density of galaxies across the cosmic time.

Lyman-break technique

The SEDs of star-forming galaxies usually show a break (or dropout) feature at wavelengths shorter than the Lyman limit ($\lambda \sim 912\text{\AA}$) due to neutral hydrogen around the star-forming region in the line-of-sight. At $z \gtrsim 6$, similar to the Gunn-Peterson trough (see Section 1.2), the high density of neutral hydrogen in the IGM can also cause a significant dropout feature at wavelengths shorter than Ly α . Galaxies with SEDs showing Lyman-break features are called *Lyman-break galaxies*, and are characterized as having a relatively high star-formation rate. Lyman-break galaxies can be selected using photometric colours in different bands. This method is called the *Lyman-break technique*. For high- z galaxies, the Lyman-breaks are shifted to larger wavelength depending the redshifts of galaxies. The photometric redshifts of these galaxies can be efficiently identified using the related colour selection criteria (Steidel et al., 1996, 1999).

Luminosity functions

One of the fundamental statistics of galaxies is the luminosity function (LF), $\phi(L)$, which describes the spatial number density of galaxies per unit luminosity. The galaxy LF is commonly fitted by a Schechter function (Schechter, 1976):

$$\phi(L)dL = \left(\frac{\phi_*}{L_*}\right) \left(\frac{L}{L_*}\right)^\alpha \exp\left(-\frac{L}{L_*}\right) dL, \quad (1.69)$$

where ϕ_* is the density normalization, L_* is a characteristic luminosity where the power-law form of the function cuts off, and α is the faint-end slope. Equation 1.69 can be presented using magnitudes as

$$\phi(M)dM = 0.4 \ln 10 \phi_* 10^{0.4(\alpha+1)(M_* - M)} \exp\left[-10^{0.4(M_* - M)}\right] dM, \quad (1.70)$$

where M_* is the magnitude corresponding to L_* .

The observed LFs show a drop out at both bright and faint ends with respect to a pure power law. In semi-analytic models, the faint-end of the luminosity can be matched with a combination of supernova feedback and the suppression of gas cooling in low-mass haloes by the photoionization background (e.g. Benson et al., 2002b). The drop out at bright-end of LF may be the consequence of the strong feedback from both AGN and supernovae (e.g. Croton et al., 2006).

To characterize the UV flux from galaxies during the EoR, we are particularly interested in their UV LFs (rest-frame). At $z > 6$, intrinsic UV luminosities are shifted to the infrared band in the observer-frame, so that advanced infrared cameras are needed in high- z observations. For comparison with different observations and simulations, the observed UV luminosities (M_{UV}) are usually represented using some specific wavelength. For example, M_{1500} , M_{1600} correspond to the absolute luminosity at rest-frame 1500 and 1600 Å, respectively. These are estimated from fitting template SEDs or converted from the observed luminosities using the estimated slope of the UV continuum. Figure 1.5 shows the observed UV LFs at 1600 Å at redshift $z \sim 4\text{--}8$ from Bouwens et al. (2015b). We can see the clear evidence for the buildup of luminosity and number density of galaxies across the cosmic time. Successfully modelling the observed LFs is one of the main goals of galaxy formation models.

Probing the cosmic star-formation history

As a direct observable, UV luminosity of galaxies is a valuable tracer of other intrinsic galaxy properties. For example, the observed galaxy luminosity at different redshifts can be used to estimate the global star-formation activity as a function of the cosmic time. The cosmic star-formation history is usually characterized by the star-formation rate density $\dot{\rho}_*(z)$, which is defined as the total star formation per unit time per unit

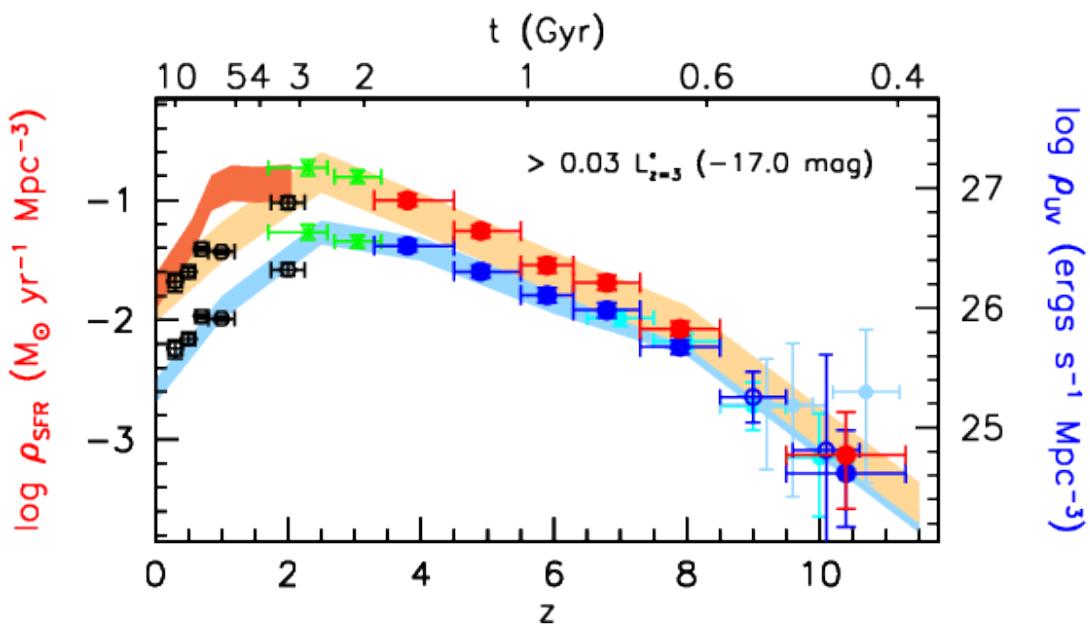


Figure 1.6: Observed SFR densities for galaxies with UV luminosity brighter than -17mag . The upper and lower set of data points from Schiminovich et al. (2005, black hexagons) at $z \sim 0\text{--}2$, Reddy & Steidel (2009, green crosses) at $z \sim 2\text{--}3$ and Bouwens et al. (2015b) and shaded regions show the SFR and UV luminosity densities with and without dust correction respectively. The instantaneous star-formation rate is derived using relation between UV luminosity and star-formation rate from Kennicutt (1998a) and Madau et al. (1998). Observed data from McLure et al. (2013, cyan circles) at $z \sim 7\text{--}8$, Ellis et al. (2013, cyan circles) at $z \sim 9\text{--}10$, CLASH (Zheng et al., 2012; Coe et al., 2013; Bouwens et al., 2014a, light blue circles), Oesch et al. (2013, 2014, blue open circles) and Daddi et al. (2009); Magnelli et al. (2009, 2011, dark red shaded region) are also shown. Figure is adapted from Bouwens et al. (2015b).

volume at redshift z . If we assume the SFR of a galaxy is a function of its luminosity in a specific band $m_* = m_*(L)$, the SFR density can be estimated from the observed UV LFs of galaxies

$$\dot{\rho}_*(z) = \int_0^\infty m_*(L)\phi(L, z)dL. \quad (1.71)$$

Figure 1.6 shows the SFR densities from Bouwens et al. (2015b) calculated using the dust-corrected UV luminosity of galaxies. We can see that the cosmic star-formation history shows a peak at redshift $z \sim 2\text{--}3$ and declines at lower and higher redshifts. Integrating the cosmic star-formation history over time, we can obtain the buildup of the global stellar population of galaxies. We should note that the estimation of the cosmic star-formation history may be biased by some aspects, including (i) the actual lower limit of the integration in Equation 1.71 is always above zero and is determined by the detection limits of telescopes; (ii) The effect of dust attenuation which obscures the light from stars needs to be carefully modelled according to the properties or environment of galaxies.

Probing structure formation

Galaxies are not found to be randomly distributed in space. Instead, they are clustered and grouped. Since galaxies are assumed to form within dark matter haloes, their spatial distribution can therefore probe the large-scale structure of dark matter. One of the commonly used statistics to characterize the degree of clustering of galaxies is the two-point correlation function. The correlation function of galaxies is measured using either their spatial or angular distribution. The spatial two-point correlation function can be defined as the excess number of galaxies at a separation, r , relative to that expected from a random galaxy distribution:

$$\xi(r) = \frac{DD(r)\Delta r}{RR(r)\Delta r} - 1, \quad (1.72)$$

where $DD(r)\Delta r$ and $RR(r)\Delta r$ are the numbers of galaxy pairs with separations in the range $r \pm \Delta r/2$ for the sample and for the expected number if galaxies are randomly distributed. On small scales with $0.1h^{-1}\text{Mpc} \leq r \leq 10h^{-1}\text{Mpc}$, the spatial correction function is well described by a power law

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}, \quad (1.73)$$

with $\gamma \sim 1.8$ and a correction length, $r_0 \sim 5h^{-1}\text{Mpc}$. The power spectrum $P(k)$ is related to the two-point correlation function by

$$\xi(r) = \frac{1}{2\pi^2} \int k^2 P(k) \frac{\sin(kr)}{kr} dk. \quad (1.74)$$

In addition to the distribution of galaxies, the distribution of galaxy groups and clusters can also be used as a probe for the large scale structure. However, the spatial distribution of galaxies, galaxy groups and galaxy clusters depends not only on the distribution of dark matter but also on the how galaxies form and assemble within different dark matter haloes. Semi-analytic models based on hierarchical galaxy formation are one of the approaches to associate the unseen dark matter to observed galaxy population, as described in the next chapters of the thesis.

1.5 High- z galaxies and the DRAGONS project

Many observational and theoretical works have studied the galaxy formation and the process of reionization in the early Universe. A significant number of galaxies at $z > 6$ have been observed and statistically studied in the last decade (Bouwens et al., 2007, 2010; Oesch et al., 2010b; McLure et al., 2010; Bouwens et al., 2011; Finkelstein et al., 2012; Schenker et al., 2013; McLure et al., 2013; Duncan et al., 2014; Schmidt et al., 2014; Bouwens et al., 2014a, 2015b), and individual candidate galaxies up to $z \sim 11$ have been discovered (e.g. Oesch et al., 2016). However, as models predict the existence of faint galaxies below current detection limits (e.g. $M_{\text{UV}} \sim -17$ at $z \sim 6$), observations with higher sensitivity and spatial resolution are required in order to fully study the formation and evolution of high- z galaxies. These requirements will be satisfied by next generation telescopes including the *James Webb Space Telescope (JWST)*, which will observe galaxies ≥ 2 magnitudes fainter than current detection limits. *JWST* will also observe the galaxy population at $z > 10$ in greater numbers and observe candidates approaching the first galaxies in the Universe. To better understand these new observations, and to make observational predictions, new simulations and models will be needed for studying the galaxy formation at very high redshifts. However, due to the much smaller mass and dynamical time scales of dark matter haloes and galaxies at high redshifts, the galaxy formation physics such as supernova feedback, photoionization suppression and mass recycling must be carefully modelled and studied.

At the same time, observations of the CMB and high- z sources, such as quasars, Lyman- α emitters, Lyman-break galaxies and gamma-ray bursts, have provided constraints on the completion and duration of reionization. However, the detailed reionization history and the reionization morphology are still unknown. As predicted by theory, these may be constrained using 21-cm signals (see Section 1.2). Next generation radio instruments such as *The Square Kilometre Array (SKA)* and *Hydrogen Epoch of Reionization Array (HERA)* will allow high-resolution 21-cm tomographic maps of ionizing structure to be made (e.g. Mellema et al., 2015; Wyithe et al., 2015). To model the structure and evolution of reionization, many numerical works have been carried out. N -body simulations with post-processed radiative transfer methods based on halo occupation provide a description of the reionization structure but not a self-consistent

calculation of ionizing source properties (e.g. McQuinn et al., 2007). Hydrodynamical simulations of galaxy formation including radiative transfer have been used to calculate the ionization structure and the feedback effect on galaxy formation self-consistently (e.g. So et al., 2014; Norman et al., 2015; Pawlik et al., 2015; Ocvirk et al., 2016). However, the high computational expense of these simulations limits their scales and resolution. Efficiently and self-consistently modelling the reionization structure on sufficient scales becomes essential for the study of reionization.

To study the high- z galaxy formation and the structure of reionization, the *Dark-Ages, Reionization And Galaxy-formation Observables from Numerical Simulations (DRAGONS)* project⁴ is carried out. *DRAGONS* combines high (mass and temporal) resolution N -body simulations, a semi-analytic model of galaxy formation and a semi-numerical model of reionization structure. The goal of DRAGONS is to self-consistently model observations of high- z galaxies and the structure of reionization.

1.6 Thesis Structure

The work described in this thesis has been carried out as part of the *DRAGONS* project. The aim of which is to study the high- z galaxy observables including UV luminosities and disc sizes, as well as investigate the ionizing photon budget of reionization. The models, calculations and related results will be presented as follows:

- Chapter 2 introduces the dark matter N -body simulations and the semi-analytic model of galaxy formation used in this thesis. Calculations of multi-band photometric luminosities of galaxies including the Ly α absorption and dust extinction are also described.
- Chapter 3 studies the UV LFs of model galaxies. Using the predicted UV LFs, we investigate the faint-end slope of UV LFs below current detection limits and calculate the UV photon contribution of undetected galaxies.
- Chapter 4 studies the UV luminosity dependence of galaxy/halo intrinsic properties including the luminosity–SFR relation, the luminosity–stellar mass relation and the luminosity–halo mass relation.
- Chapter 5 discusses the modelling of galaxy sizes and studies the size evolution of galaxies. The effects of supernova feedback on galaxy sizes is also studied.
- Chapter 6 summarises the key findings of the thesis and presents ideas for future studies.

⁴<http://dragons.ph.unimelb.edu.au/>

Chapter 2

Modeling galaxy observables

2.1 Introduction

Modeling galaxy formation during the EoR is challenging. In addition to the fact that high- z galaxies are faint leading to a lack of observational constraints, there are at least two main issues: first, the simulations of dark matter halo merger trees in the early Universe require much higher mass and temporal resolutions due to the lower virial masses and shorter dynamical time scales involved. Second, the typical mass of high- z galaxies is much lower than the mass of their low- z counterparts, so that feedback from supernovae and the UV photoionizing background play an important role in preventing galaxy formation within low-mass haloes and need to be accurately modeled. Therefore, we cannot directly implement existing galaxy formation models, which are designed and calibrated at low redshifts to high-redshift studies without significant modification. For this reason, within the *DRAgOnS* project, the *N*-body simulation and the semi-analytic model have been specially designed for studies of reionization.

The galaxy formation model used in this work is MERAXES, introduced by Mutch et al. (2016a), which is a new semi-analytic model with updated physics based on Croton et al. (2006). MERAXES is implemented on dark matter halo merger trees generated from the N-body simulations *Tiamat* described in Poole et al. (2016).

In this chapter, we give a description of *Tiamat*, MERAXES and the calculation of galaxy luminosities. The chapter is organised as follows: Section 2.2 introduces the *Tiamat* N-body simulation. Section 2.3 describes the MERAXES semi-analytic model. Section 2.4 describes the stellar population synthesis method, Ly α absorption and dust extinction we have used for luminosity calculations. We employ a standard spatially-flat Λ CDM cosmology based on the *Planck* 2015 data (Planck Collaboration et al., 2016) throughout: $(h, \Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s) = (0.678, 0.308, 0.0484, 0.692, 0.815, 0.968)$. All galaxy magnitudes are presented in the AB system (Oke & Gunn, 1983).

2.2 *Tiamat* *N*-body simulation

2.2.1 Volume and resolution

Tiamat is a suite of collisionless dark matter *N*-body simulations carried out by Poole et al. (2016) and Angel et al. (2016). The fiducial *Tiamat* simulation includes $N = 2160^3$

Table 2.1: Summary of parameters in the *Tiamat* N -body simulations. N is the number of simulation particles, L is the side length of the simulation box, m_{DM} is the mass of each dark matter particle, ϵ is the gravitational softening length, and η is the integration accuracy parameter.

Simulation run	N	$L[h^{-1}\text{Mpc}]$	$m_{\text{DM}}[h^{-1}\text{M}_\odot]$	$\epsilon[h^{-1}\text{kpc}]$	η
<i>Tiamat</i>	2160^3	67.8	2.64×10^6	0.63	0.025
<i>Medi Tiamat</i>	1080^3	22.6	7.83×10^5	0.42	0.025
<i>Tiny Tiamat</i>	1080^3	10.0	6.79×10^5	0.19	0.025

particles in a periodic box with side length of $L = 67.8 h^{-1}\text{Mpc}$ (comoving). This simulation volume is approximately sufficient to capture the evolving structure and statistics of the reionization field (e.g. Furlanetto, 2016). The mass of each particle is $m_{\text{DM}} = 2.64 \times 10^6 h^{-1}\text{M}_\odot$, which allows us to identify the low mass dark matter haloes ($M_{\text{vir}} \sim 10^8 \text{ M}_\odot$) with virial temperatures close to the hydrogen cooling limit of ($T_{\text{vir}} \sim 10^4 \text{ K}$) across the redshifts relevant to reionization (e.g Barkana & Loeb, 2001). To investigate the effect of mass resolution in *Tiamat*, several small volume but higher mass resolution simulations have also been run, which include *Medi Tiamat* and *Tiny Tiamat*. The *Medi Tiamat* and *Tiny Tiamat* run have the same particle number of $N = 1080^3$, but the former was run in a volume of $L = 22.6 h^{-1}\text{Mpc}$ cube with each particle representing $m_{\text{DM}} = 7.83 \times 10^5 h^{-1}\text{M}_\odot$, while the latter was run in a $L = 10.0 h^{-1}\text{Mpc}$ cube box with a mass of each particle representing $m_{\text{DM}} = 6.9 \times 10^4 h^{-1}\text{M}_\odot$.

All these N -body simulations were performed using standard ΛCDM cosmology with parameters give in Section 2.1. Initial conditions were generated using the second-order Lagrangian perturbation code *2LPTic* (Crocce et al., 2006) with fixes to ensure correct behaviour if particle or displacement field grid cell counts exceed $2^{32} - 1$ (Poole et al., 2015). The initial power spectrum was generated using CAMB (Lewis et al., 2000).

The *Tiamat* simulations were carried out using the cosmological simulation code GADGET-2 (Springel et al., 2001b; Springel, 2005), with RAM conserving modifications listed in Poole et al. (2015). The gravitational softening length, ϵ , (see Section 1.3.4) was set to $0.02\bar{d}$, where $\bar{d} = L/\sqrt[3]{N}$ is the average separation of simulation particles. The integration accuracy parameter, η , was set to 0.025 in each simulation. The relevant numerical parameters of the *Tiamat* simulations are summarized in Table 2.1.

2.2.2 Merger tree construction

In *Tiamat*, dark matter halo finding was carried out using the SUBFIND code (Springel et al., 2001a). This code first identifies dark matter collapsed regions (haloes) using a friends-of-friends (FoF) algorithm with a linking length criterion of 0.2 times of the mean inter-particle separation. The self-bound substructures (subhaloes) are subsequently identified within these FoF groups as locally overdense collections of dark

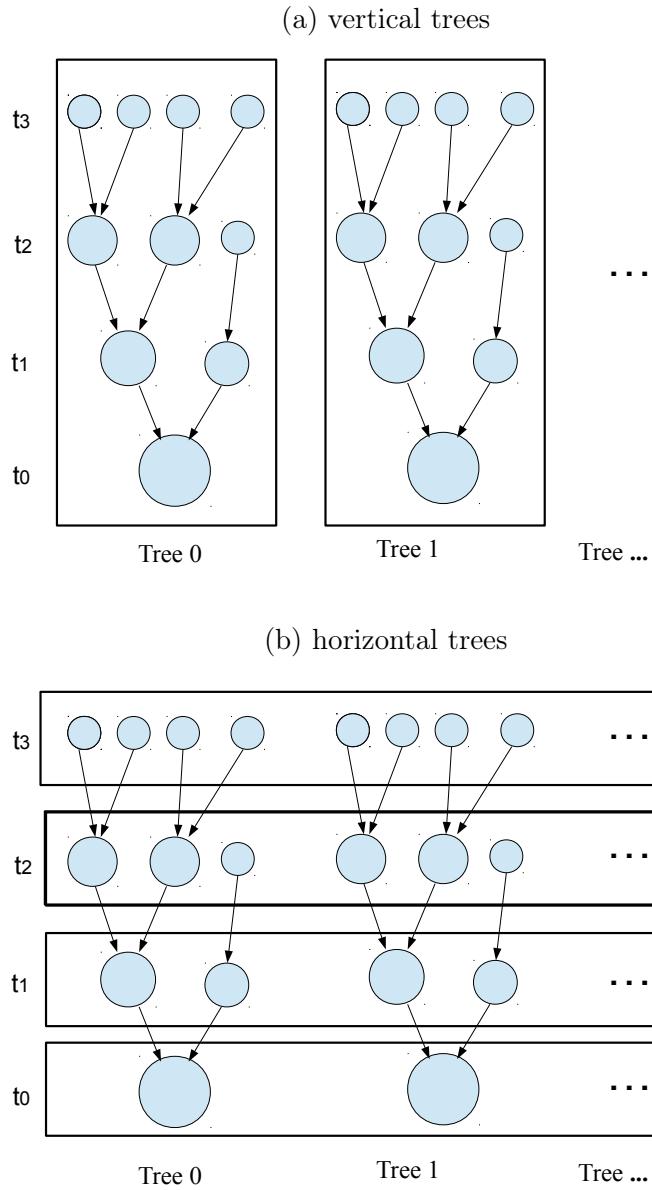


Figure 2.1: Illustration of horizontal-form merger trees (lower panel) compared to vertical-form merger trees (upper panel). In vertical trees, a halo at the latest time (t_0) and all its progenitor haloes at early times (t_1, t_2 , etc.) are stored as one tree so that one has to locate a halo at a given redshift from its latest descendant halo (i.e. “root”). In horizontal trees, all haloes at the same snapshot are stored simultaneously, so that one can locate and process all haloes at a given snapshot using only the snapshot number rather than knowing their future evolution.

matter particles, after removing unbound particles through a gravitational unbinding procedure. A FoF group typically contains a central halo holding most of the virial mass and a group of lower-mass subhaloes which trace the undigested parts of merger events.

All *Tiamat* runs outputs 100 snapshots from $z = 35$ to 5 with a constant temporal resolution of ~ 11 Myr per snapshot. This high snapshot cadence resolves the dynamical time of galaxy discs at $z > 6$, and is comparable to the lifetime of massive stars. Dark matter halo merger trees constructed from *Tiamat* are stored and processed in a “horizontal” form. This form is different with the commonly used “vertical” form (e.g. Springel et al., 2005), in which a halo at the latest time and all its progenitor haloes at early times are stored as one tree so that one has to locate a halo at a given redshift from its latest descendant halo (i.e. “root”). In the horizontal form, all haloes from the same snapshot are stored simultaneously, so that one can locate and process all haloes at a given snapshot using only the snapshot number rather than knowing their future evolution. Another key point is that in the vertical tree paradigm, haloes which are spatially co-located are assumed to interact with each other. However, ionizing bubbles can be larger $10 h^{-1} \text{Mpc}$ and across over trees during the EoR, so that the horizontal form of dark matter merger tree is essential. The horizontal form of the dark matter merger trees enables MERAXES to temporally and spatially couple a semi-numerical reionization algorithm 21CMFAST (Mesinger et al., 2011) to study the structure of reionization and its evolution.

This tree form also allows MERAXES to implement a self-consistent calculation of feedback from the photoionizing background on galaxy formation in low mass haloes, where the galaxy formation at a given snapshot depends on the ionizing photons from the local reionization background at the same snapshot.¹

2.3 The semi-analytic model — MERAXES

This section provides some basic characteristics of the MERAXES semi-analytic model introduced by Mutch et al. (2016a).

2.3.1 Gas infall and cooling

In each FoF group, the baryonic mass is initially assumed to be a universal fraction of the total halo mass, $f_b = \Omega_b/\Omega_m$, and the baryonic mass is assumed to be in the form of primordial gas. However, the infalling baryons which remain bound to the FoF group and participate in galaxy formation may be reduced by a number of factors including the photo-evaporation due to the UVB radiation, in which ionizing photons from both local and external sources can heat the IGM, increasing the local Jeans mass

¹UVB feedback in MERAXES is actually evaluated using the reionization background at the last snapshot rather the current snapshot.

and reducing the gas infall in low-mass haloes. MERAXES therefore includes a baryonic fraction modifier, f_{mod} (see Section 2.3.8), which represents the reduced baryonic gas fraction which captured by an FoF group relative to the global baryon fraction.

Gas infalling into a halo is assumed to be shocked to the virial temperature of the halo

$$T_{\text{vir}} = 35.9 \left(\frac{V_{\text{vir}}}{\text{km/s}} \right)^2 \text{ K}, \quad (2.1)$$

where V_{vir} is the virial velocity (see Section 1.3.3 for the definition of virial parameters). The hot gas can subsequently cool via a number of mechanisms, with a cooling time at radius r of

$$t_{\text{cool}}(r) = \frac{3}{2} \frac{\bar{\mu} m_p k T}{\rho_{\text{hot}}(r) \Lambda(T, Z)}, \quad (2.2)$$

where $\bar{\mu} m_p$ is the mean particle mass of the gas, k_B is the Boltzmann constant, ρ_{hot} is the hot gas density, and $\Lambda(T, Z)$ is the cooling function (Sutherland & Dopita, 1993) which depends on both the temperature, T , and the metallicity, Z , of the gas. The density distribution of hot gas in a halo is assumed to have a simple isothermal spherical profile:

$$\rho_{\text{hot}}(r) = \frac{m_{\text{hot}}}{4\pi R_{\text{vir}}^2 r^2}. \quad (2.3)$$

The hot gas may subsequently cool via a number of cooling mechanisms (see Section 1.4.2). The cooling radius, r_{cool} , is defined as the radius at which t_{cool} is equal to the dynamical (free-fall) time of the halo, $t_{\text{dyn}} = R_{\text{vir}}/V_{\text{vir}}$ (Croton et al., 2006). The gas enclosed within r_{cool} has sufficient time to cool and flow to the centre. There are three cooling regimes for an individual halo according to the comparison of r_{cool} and R_{vir} :

1. For haloes which have $r_{\text{cool}} > R_{\text{vir}}$, the hot gas directly cools into the central regions of haloes. The cold gas is assumed to flow directly into the central galaxy of the FoF group over the dynamical time, so that the cooling rate is

$$\dot{m}_{\text{cool}} = \frac{m_{\text{hot}}}{t_{\text{dyn}}}. \quad (2.4)$$

2. For haloes which have $r_{\text{cool}} < R_{\text{vir}}$, a quasi-static hot atmosphere will form. The mass cooling rate is determined by the mass differential enclosed within the evolving cooling radius:

$$\begin{aligned} \dot{m}_{\text{cool}} &= 4\pi \rho_{\text{hot}}(r_{\text{cool}}) r_{\text{cool}}^2 \dot{r}_{\text{cool}} \\ &= \frac{1}{2} m_{\text{hot}} \frac{r_{\text{cool}} V_{\text{vir}}}{R_{\text{vir}}^2}. \end{aligned} \quad (2.5)$$

3. When $T_{\text{vir}} \leq 10^4 \text{ K}$, no cooling occurs. This is motivated by the fact that the hot

gas can not efficiently cool below 10^4 K (Barkana & Loeb, 2001).

2.3.2 Star formation

Cold gas flowing onto a galaxy is assumed to settle into a rotationally supported galaxy disc with size $r_{\text{disk}} = (3\lambda/\sqrt{2})R_{\text{vir}}$, where λ is the spin parameter of the halo (Mo et al., 1998). Based on the observational work of Kennicutt (1998b); Bullock et al. (2001), the global SFR of a spiral galaxy can be related to the surface density of cold gas above a given threshold. In our model, we adopt a critical surface density for the disc, above which gas cannot maintain stability and will start forming stars. The critical density at a radius r is adopted from Kauffmann (1996),

$$\Sigma_{\text{crit}}(r) = \Sigma_{\text{norm}} \left(\frac{V_{\text{vir}}}{\text{km s}^{-1}} \right) \left(\frac{r}{\text{kpc}} \right)^{-1} \text{M}_{\odot} \text{pc}^{-2}, \quad (2.6)$$

where $\Sigma_{\text{norm}} = 0.2$ is a free parameter in MERAXES. Stars are assumed to form within a maximum radius set to $3R_d$ based on the properties of the Milky Way (van den Bergh, 2000). By integrating Σ_{crit} to R_{disc} , we obtain the critical mass of the disc,

$$m_{\text{crit}} = 2\pi \Sigma_{\text{norm}} \left(\frac{V_{\text{vir}}}{\text{km s}^{-1}} \right) \left(\frac{R_{\text{disc}}}{\text{kpc}} \right) 10^6 \text{M}_{\odot}. \quad (2.7)$$

If the mass of cold gas in the disc, m_{cold} , exceeds this threshold mass then stars will form with a star-formation rate given by

$$\dot{m}_* = \alpha_{\text{SF}} \frac{m_{\text{cold}} - m_{\text{crit}}}{t_{\text{dyn}}^{\text{disc}}}, \quad (2.8)$$

where $\alpha_{\text{SF}} = 0.03$ is a free parameter describing the star formation efficiency and $t_{\text{dyn}}^{\text{disc}} = R_d/V_{\text{vir}}$ is the dynamical time of the disc.

The assumed initial stellar mass function (IMF) in MERAXES is a standard Salpeter (1955) IMF of the form $\phi(m) \propto m^{-2.35}$ in the mass range of $0.1\text{M}_{\odot} \leq m \leq 120\text{M}_{\odot}$.

2.3.3 Delayed supernova feedback

MERAXES includes internal galaxy feedback from type II supernovae. Stars with mass greater than 8M_{\odot} end their lives as type II supernovae and release mass and energy. In many semi-analytic models (e.g. those based on the Millennium Simulation; Springel et al., 2005), the time separation between each simulation snapshot is large (~ 200 – 250 Myr), and an assumption of instantaneous supernova feedback is used, with energy and mass released as soon as the progenitor stars are formed. However, the *Tiamat* simulation has a much higher time resolution of ~ 11 Myr in order to resolve the shorter galaxy dynamical time at high redshift. An 8M_{\odot} star, which lives ~ 40 Myr (e.g. Portinari et al., 1998) will therefore explode ~ 3 – 4 snapshots after it formed. For this

reason, MERAXES implements a delayed supernova feedback scheme, where a supernova may explode several snapshots after the star-formation episode.

In MERAXES, the mass and energy released by a type II supernova are estimated as averaged values per supernova, and the total released mass and energy can therefore be calculated by counting the number of supernovae. The number of type II supernovae per unit stellar mass, η_{SNII} is calculated by integrating the IMF of stars from a lower-mass cut 8 M_\odot to a maximum stellar mass of 120 M_\odot :

$$\eta_{\text{SNII}} = \int_{8\text{M}_\odot}^{120\text{M}_\odot} \phi(m) dm = 7.43 \times 10^{-3} \text{M}_\odot^{-1}, \quad (2.9)$$

where ϕ is defined by the assumed IMF via:

$$\int_{0.1\text{M}_\odot}^{120\text{M}_\odot} \phi(m) dm = 1. \quad (2.10)$$

For a burst of stars with mass Δm_* , the total energy from supernova feedback deposited into the ISM in the galaxy is

$$\Delta E_{\text{total}} = \epsilon_{\text{energy}} E_{\text{nova}} \eta_{\text{SNII}} \Delta m_*, \quad (2.11)$$

where $E_{\text{nova}} = 10^{51}$ erg is the average energy released by a type-II supernova, and ϵ_{energy} is the efficiency with this energy couples to the surrounding ISM. The mass of surrounding gas which is reheated by this energy deposition (not the released mass from supernovae) is

$$\Delta m_{\text{total}} = \epsilon_{\text{mass}} \Delta m_*, \quad (2.12)$$

where ϵ_{mass} is a free parameter commonly referred to as the mass loading factor. MERAXES follows Guo et al. (2013) for parameterizing ϵ_{energy} :

$$\epsilon_{\text{energy}} = \alpha_{\text{energy}} \left[0.5 + \left(\frac{V_{\text{max}}}{V_{\text{energy}}} \right)^{-\beta_{\text{energy}}} \right], \quad (2.13)$$

and parameterizing ϵ_{mass} :

$$\epsilon_{\text{mass}} = \min \left\{ \alpha_{\text{mass}} \left[0.5 + \left(\frac{V_{\text{max}}}{V_{\text{mass}}} \right)^{-\beta_{\text{mass}}} \right], \epsilon_{\text{mass}}^{\text{max}} \right\}, \quad (2.14)$$

where α_{energy} , V_{energy} , β_{energy} , α_{mass} , V_{mass} and β_{mass} are taken to be free parameters. In MERAXES, $\beta_{\text{mass}} = 2$ is set for a standard energy-driven wind (Murray et al., 2005). MERAXES has also imposed constraints to make $0 \leq \epsilon_{\text{energy}} \leq 1$ for it is an efficiency, and to let $\epsilon_{\text{mass}} \leq \epsilon_{\text{mass}}^{\text{max}} = 10$ based on expectations for typical high- z galaxies (e.g. Martin, 1999; Uhlig et al., 2012).

The snapshot cadence of the simulation is 11 Myr, so that the ~ 40 Myr life time of the least massive star ($8 M_{\odot}$) for type II supernovae corresponds to ~ 4 snapshots. The total energy (E_{reheated}) and the total mass (m_{reheated}) released in a given snapshot are calculated by tracing the number of stars formed during last 4 snapshots that will explode in the current snapshot. To calculate the main-sequence time of massive stars with different masses, MERAXES uses a functional fit to the $Z = 0.004$ core burning lifetime tabulated by Portinari et al. (1998) and assumes that stars will go supernova immediately upon expending their H and He cores.

The eventual fate of reheated materials depend on both $\Delta E_{\text{reheated}}$ and $\Delta m_{\text{reheated}}$. Assuming that the cold gas with mass m_{heated} will be adiabatically heated to T_{vir} , the associated change in thermal energy is

$$\Delta E_{\text{hot}} = \frac{1}{2} \Delta m_{\text{reheated}} V_{\text{vir}}^2, \quad (2.15)$$

where V_{vir} is the virial velocity of the FoF group. If $\Delta E_{\text{reheated}} > \Delta E_{\text{hot}}$, The gas is assumed to be added to the hot halo of the host FoF group, and the excess energy ($\Delta E_{\text{reheated}} - \Delta E_{\text{hot}}$) is assumed to go into ejecting a fraction of the hot reservoir of the FoF group from the system entirely:

$$\Delta m_{\text{eject}} = \frac{\Delta E_{\text{reheated}} - \Delta E_{\text{hot}}}{0.5 V_{\text{vir}}^2}. \quad (2.16)$$

If $\Delta E_{\text{reheated}} < \Delta E_{\text{hot}}$, the released energy can only reheat a fraction of total reheated mass to the FoF hot reservoir:

$$\Delta m_{\text{hot}} = \frac{\Delta E_{\text{reheated}}}{0.5 V_{\text{vir}}^2}, \quad (2.17)$$

with the rest falling back to the galaxy. The gas and metals which are ejected from the system entirely are placed in an ejected reservoir which is assumed to play no further role in galaxy formation in the host FoF group. However, the ejected material will be re-accreted into a new FoF group hot gas halo when the group mergers with a more massive system.

2.3.4 Metal enrichment

In addition to energy, supernovae also release chemical elements into the ISM of galaxy. MERAXES implements a simple metal enrichment scheme that is common in semi-analytic models (e.g. De Lucia et al., 2004; Somerville et al., 2008; Guo et al., 2011). MERAXES assumes that a fixed yield, Y , of metals is produced per solar mass of stars. These metals will be released into the ISM by supernovae explosions. The released metals at a given snapshot are calculated by tracing the mass of all stars formed during

previous snapshots that explode at this snapshot.

2.3.5 Stellar mass recycling

When stars leave the main sequence, a fraction of stellar mass return to the ISM through stellar winds or supernova explosion, and the released mass will be used for further star formation. This is the stellar mass recycling. The common assumption in many semi-analytic models is the instantaneous recycling approximation (e.g. Cole et al., 2000; Croton et al., 2006), in which a fixed fraction of the mass of newly formed stars is recycled immediately into the ISM of the galaxy. At high redshift, as main sequence lifetimes of most stars become comparable to the age the Universe, more realistic non-instantaneous recycling schemes must be implemented (e.g. Nagashima et al., 2005a,b; Arrigoni et al., 2010; Benson & Bower, 2010; De Lucia et al., 2014).

MERAXES implements an non-instantaneous stellar mass recycling scheme. When a star with mass more massive than $1 M_{\odot}$ leaves its main-sequence, MERAXES assumes that its entire mass is returned to the ISM and no remnants are left. To calculate the number of stars that die in a given snapshot, MERAXES divides stars into two groups according to their mass:

1. Stars more massive than $8 M_{\odot}$ die as supernova explosion. The number of dying stars are calculated by tracing the stars formed in the past 4 snapshots as introduced in delayed supernova feedback.
2. For stars with mass in the range of $1-8 M_{\odot}$. MERAXES assumes that these stars have formed in a single coeval burst which occurred at a time defined by their mass-weighted age.

2.3.6 Halo infall and gas stripping

In N -body simulations, the mass of a FoF group may drop due to tidal stripping by other more massive systems. When this happens, MERAXES assumes that same fraction of the ejected or hot baryonic content of the FoF group is also removed. Depending on the amount of mass which is needed to be removed, the material is first taken from the ejected reservoir, with further mass being removed from the hot gas reservoir if required. No cold gas or stars can be removed because they are deep in the potential wells of haloes. In addition to the long-range tidal stripping, hot gas material in a satellite galaxy can be strongly stripped by the central galaxy. In MERAXES, if FoF groups fall directly into a massive system, it is assumed that the entire hot and ejected reservoirs of these FoF groups fall into the hot reservoir of the new FoF group. Although this rapid stripping represents an extreme and simplistic scenario, we note that such approximation has been made in a number of previous semi-analytic models and defer a more realistic treatment to future works.

2.3.7 Mergers

Mergers play an important role in galaxy formation through both hierarchical mass assembly and induced star formation. Mergers are more important at high redshifts due to their high prevalence. Galaxy mergers are triggered by the merging of their host haloes. The position of an infalling galaxy is initially represented by the most bound particle of its subhalo. When the mass of the subhalo drops below a threshold, the position and velocity of the infalling galaxy are represented by the subhalo at the last time it was identified. At the same time, dynamical friction arguments are utilized to estimate the time of the orbit decay of the infalling galaxy (Binney & Tremaine, 2008):

$$t_{\text{friction}} = 1.17 \frac{V_{\text{vir}} r_{\text{sat}}^2}{G m_{\text{sat}} \ln(1 + M_{\text{vir}}/m_{\text{sat}})}, \quad (2.18)$$

where m_{sat} and r_{sat} are the mass and distance of the infalling galaxy, respectively, M_{vir} and V_{vir} are the virial properties of the halo of the central galaxy. All of these quantities are evaluated at the last time that the infalling subhalo was identified.

Galaxy mergers can also trigger a strong burst of star formation. MERAXES assumes that a fraction of the total cold gas of the newly formed system is consumed during such a burst (Somerville et al., 2001)

$$e_{\text{burst}} = \alpha_{\text{burst}} \left(\frac{m_{\text{small}}}{m_{\text{big}}} \right)^{\gamma_{\text{burst}}}, \quad (2.19)$$

where $m_{\text{small}}/m_{\text{big}}$ is the mass ratio of the merging galaxies, and $\alpha_{\text{burst}} = 0.56$ and $\gamma_{\text{burst}} = 0.7$ are chosen to fit the numerical results of Cox et al. (2004) and Mihos & Hernquist (1994, 1996) for merger mass ratios in the range 0.1–1.0 (Croton et al., 2006). For simplicity, it is assumed that the merger-driven burst occurs within a single snapshot, which is comparable to the disc dynamical time of the majority of galaxies.

2.3.8 Reionization

A key goal of DRAGONS is to connect the reionization structure of the IGM to the formation of galaxies. Previous semi-analytic models have employed a homogeneous reionization prescription which treats the ionizing state of the IGM as a global property at each redshift (e.g. Benson et al., 2002b,a; Somerville et al., 2008). However, during the EoR, the ionizing state of the IGM depends strongly on the property of local ionizing sources and therefore must be modelled spatially. MERAXES self-consistently couples the process of reionization with galaxy formation, where the local ionizing state is determined by the number of photons generated from nearby galaxies and the resulting UVB in turn regulates the infalling baryons for local galaxy formation. To achieve this, MERAXES embeds a modified version of the semi-numerical reionization algorithm,

21CMFAST (Mesinger et al., 2011) and the calculation of the local UVB described in Sobacchi & Mesinger (2013b). This is implemented in the following manner:

1. At a single snapshot, MERAXES models the formation and evolution of galaxies, and constructs halo mass, stellar mass in the entire volume, SFR grids. The grid resolution is 512^3 , which resolves the UVB for galaxy formation.
2. The 21CMFAST algorithm calculates the ionization state and inhomogeneous UVB intensity in each grid cell.
3. MERAXES keeps track of the redshift at which each cell first become ionized, z_{ion} , and the corresponding UVB intensity, and calculates the baryon fraction modifier, f_{mod} , in each grid cell.
4. FoF groups use the calculated f_{mod} of their host grid cells for the next time step in the simulation.

This procedure is repeated until all grid cells are ionized. Through this procedure, the evolution of galaxies and ionizing structure are spatially and temporally coupled.

Reionization condition

The basic methodology of 21CMFAST is to utilize an excursion-set formalism to identify HII regions and produce a neutral hydrogen fraction for the simulation volume. The excursion set calculation is started on a scale comparable to the mean free path of ionizing photons at $z \sim 6$. A spherical region of IGM with radius r is identified as an ionized bubble if the integrated number of ionizing photons is greater than the number of absorbing atoms and associated recombinations:

$$\frac{m_*(r)}{m_p} N_\gamma f_{\text{esc}} \geq (1 + \bar{N}_{\text{rec}}) \frac{f_b(1 - \frac{3}{4}Y_{\text{He}}) M_{\text{tot}}(r)}{m_p}. \quad (2.20)$$

Here $m_*(r)$ is the integrated stellar mass within r , m_p is the mass of a proton, so that $m_*(r)/m_p$ is the number of baryons, N_γ is the average number of ionizing photons produced per baryon, f_{esc} is the escape fraction, \bar{N}_{rec} is the mean number of recombination per absorbing atom, $M_{\text{tot}}(r)$ is the total mass within r , Y_{He} is the helium mass fraction, and it is assumed that helium is singly ionized with the same rate as hydrogen. Re-writing Equation 2.20 using the hydrogen ionizing efficiency ξ :

$$\xi \frac{m_*(r)}{M_{\text{tot}}} \geq 1, \quad (2.21)$$

where

$$\xi = 6214 \left(\frac{0.157}{f_b} \right) \left(\frac{N_\gamma}{4000} \right) \left(\frac{f_{\text{esc}}}{0.2} \right) \left(\frac{0.82}{1 - \frac{3}{4}Y_{\text{He}}} \right). \quad (2.22)$$

Here $\bar{N}_{\text{rec}} = 0$ is assumed in the diffuse IGM based on the studies of high- z Ly α forest (e.g. Bolton & Haehnelt, 2007; McQuinn et al., 2011). For parameters in Equation 2.22, Y_{He} and f_b are observationally measured, and N_γ can be estimated using the assumed stellar IMF. f_{esc} is less constrained for high- z galaxies, and a value of 0.2 is chosen in the fiducial model to produce a Thomson electron scattering optical depth consistent with current observations (Planck Collaboration et al., 2016).

UVB feedback

MERAXES includes UVB photo-suppression feedback, which leads to a reduced baryon fraction, f_{mod} , in individual host dark matter halos relative to the global baryon fraction, $f_b = \Omega_b/\Omega_m$ (Sobacchi & Mesinger, 2013b):

$$f_{\text{mod}}(M_{\text{vir}}) = 2^{-M_{\text{crit}}/M_{\text{vir}}}, \quad (2.23)$$

where M_{vir} is the mass of the halo, M_{crit} is the critical halo mass at which $f_{\text{mod}} = 0.5$:

$$M_{\text{crit}} = M_0 J_{21}^a \left(\frac{1+z}{10} \right)^b \left[1 - \left(\frac{1+z}{1+z_{\text{ion}}} \right)^c \right]^d, \quad (2.24)$$

z_{ion} is the redshift at which the halo was first exposed to the UVB and $(M_0, a, b, c, d) = (2.8 \times 10^9 M_\odot, 0.17, -2.1, 2.0, 2.5)$ are parameters found by Sobacchi & Mesinger (2013b) to provide the best fit to their simulation results. The J_{21} term represents the local UVB intensity

$$J(\nu) = J_{21} \left(\frac{\nu}{3.2872 \times 10^{15} \text{ Hz}} \right)^{-\alpha}, \quad (2.25)$$

where $\alpha = 5.0$ is adopted for a stellar-driven UV spectrum (Thoul & Weinberg, 1996). The average UVB intensity to which a galaxy is exposed within an ionized region is given by

$$\bar{J}_{21} = \frac{(1+z)^2}{4\pi} \lambda_{\text{mfp}} h_{\text{P}} \alpha f_{\text{bias}} \bar{\epsilon}, \quad (2.26)$$

where λ_{mfp} is the comoving mean-free path of ionizing photons and it is assumed to have a value equal to the radius of the ionizing bubble, and $f_{\text{bias}} = 2$ is introduced to account for the excess ionizing emissivity due to galaxy clustering. $\bar{\epsilon}$ is the average number of ionizing photons emitted into the IGM per unit time per unit comoving volume, and is calculated using the grid of gross stellar mass $m_{*,\text{gross}}$ (ie. without any decrement due to stellar evolution) formed within r :

$$\bar{\epsilon} = \frac{f_{\text{esc}} N_\gamma}{\frac{4}{3}\pi r^3 m_{\text{p}}} \frac{m_{*,\text{gross}}(r)}{t_{\text{H}}}, \quad (2.27)$$

where t_{H} is the Hubble time at the related snapshot.

Table 2.2: Summary of fiducial parameters in MERAXES. As the reference, *Munich model* values from Croton et al. (2006); Guo et al. (2011, 2013); Mutch et al. (2013) and Henriques et al. (2013, 2015) are also listed although these were calibrated against $z < 3$ observations.

Parameter	Equation	Fiducial value	<i>Munich model</i> values
α_{SFR}	2.8	0.03	0.01–0.07
α_{energy}	2.13	0.5	0.18–0.7
β_{energy}	2.13	2.0	0–3.5
V_{energy}	2.13	70.0	70–336
α_{mass}	2.14	6.0	2.1–10.3
β_{mass}	2.14	0.0	0–3.5
V_{mass}	2.14	70.0	70–430
$\epsilon_{\text{mass}}^{\text{max}}$	2.14	10.0	–
Y	–	0.03	0.03–0.047
f_{esc}	2.20	0.2	–

2.3.9 Model calibration

Using *Tiamat* and MERAXES, the formation and evolution of galaxies can be studied through the growth and assembly history of galaxy stellar mass. The free parameters in MERAXES were calibrated (Mutch et al., 2016a) to replicate the observed stellar mass function at $z \sim 5$ –7 (González et al., 2011; Duncan et al., 2014; Grazian et al., 2015; Song et al., 2016), as well as the *Planck* optical depth to electron scattering measurements (Planck Collaboration et al., 2016). The resulting values of parameters in the fiducial model are listed in Table 2.2. In chapter 5, as a comparison, we will also use a no-supernova feedback model, which is calibrated using the stellar mass density at $z = 5$.

2.4 Modeling luminosities

Although MERAXES calculates a series of intrinsic properties of model galaxies, these are not direct observable. The direct observables of a high- z galaxy is its luminosities and colours. MERAXES does not directly give luminosities or colours for individual galaxies, although these could be estimated using empirical indicators such as the instantaneous SFR. New work for this thesis includes the calculation of galaxy observables including galaxy SEDs and multi-band photometric luminosities for *DRAGONS*. This is achieved by integrating the star formation histories of galaxies with template SEDs from stellar population synthesis models. To calculate observable luminosities we have to include the Ly α absorption by the IGM and the dust extinction by the ISM. This section introduces the calculation of model luminosities.

2.4.1 Stellar population synthesis

Galaxies contain populations of stars with different ages, which form in one or more progenitor galaxies. From a galaxy at a specific redshift z_0 , we trace all progenitors in

the merger tree and calculate their total SFR at each snapshot redshift z_i ($z_i > z_0$). The stars formed at z_i have an age of

$$\tau = t - t' = t(z_0) - t(z_i),$$

where $t(z)$ is the age of the Universe at redshift z . Through this process, we build a star-formation history as a function of time, $\Psi(t)$, for the observed galaxy. Because of the short dynamical time at high redshift, the starburst can result in a rapid change in UV flux during a single snapshot. Rather than begin the burst at the beginning or end of the snapshot, we therefore interpolate over 10 timesteps between each snapshot assuming a constant SFR rate. We find that our results are insensitive to the precise number of sub-steps.

For a “normal” galaxy without a significant AGN, the intrinsic (unattenuated) stellar luminosity at the rest-frame wavelength λ is

$$L_\lambda = \int_0^t \Psi(t') \mathcal{L}_\lambda(t - t') dt', \quad (2.28)$$

where $\mathcal{L}_\lambda(\tau)$ is the luminosity per unit stellar mass of the coeval population with stellar age τ , and Ψ is the star-formation history.

In this thesis, model stellar energy distributions (SEDs) are generated using the public software package STARBURST99 (Leitherer et al., 1999; Vázquez & Leitherer, 2005; Leitherer et al., 2010, 2014) with a Salpeter IMF in the mass range $0.1\text{--}120M_\odot$ in order to be consistent with the calculation of SFR in MERAXES. The Geneva evolutionary tracks with standard mass loss are selected. The metallicity is set to $Z=0.001$ ($0.05Z_\odot$), which is appropriate during the EoR. Although MERAXES computes the evolution in metallicity of the interstellar medium, for simplicity we have taken the simple approach of a single constant metallicity value for star formation. We have checked that assuming metallicity values in the range $0.001 < Z < 0.008$ does not significantly affect our results (see section 3). We do not include nebular components as they would not affect the UV luminosities of our model galaxies.

2.4.2 Ly α absorption

The spectrum of UV radiation from a high- z galaxy passing through intergalactic gas clouds which contain neutral hydrogen shows a series of Ly α absorption lines at wavelengths shorter than $\lambda=1216(1+z)$ Å. The number and optical depth of Ly α absorption are observed to significantly increase towards high redshifts, and the SED of high- z star-forming galaxies shows a significant dropout at $1216(1+z)$ Å. These are LBGs and can be identified using colour-colour selection criteria (Steidel et al., 1996, 1999).

To mimic the LBG selection process, we calculate the Lyman-dropout feature for

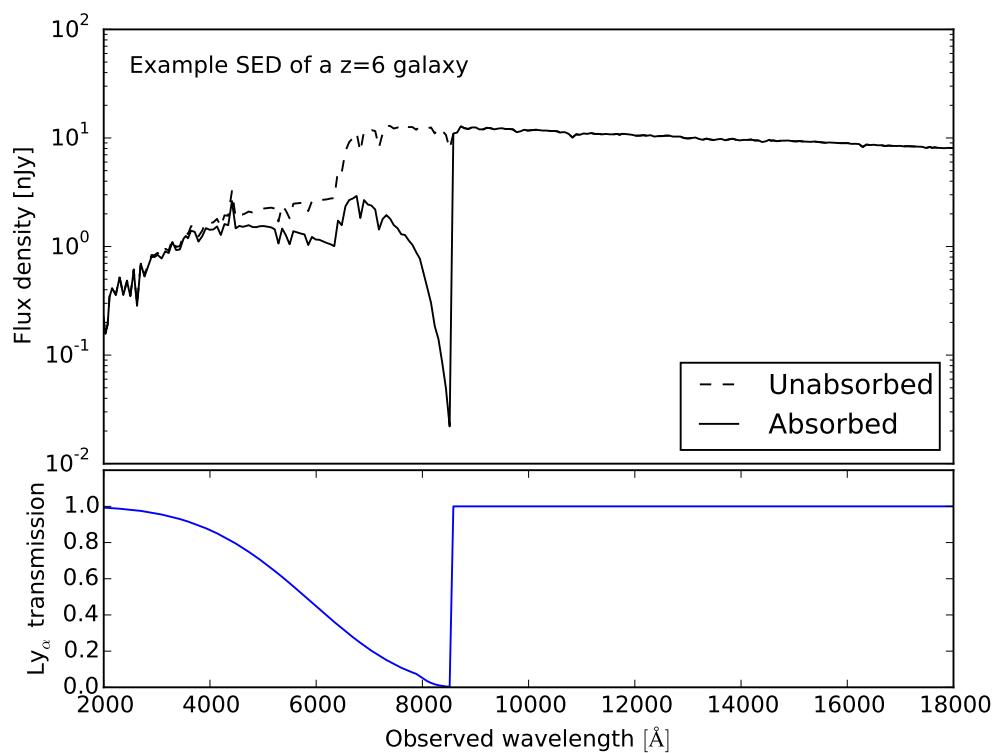


Figure 2.2: An observed SED from a model star-forming galaxy at $z = 6$. Up panel shows the observer-frame SED with (solid line) and without (dashed line) Ly α absorption. Lower panel shows the Ly α transmission curve.

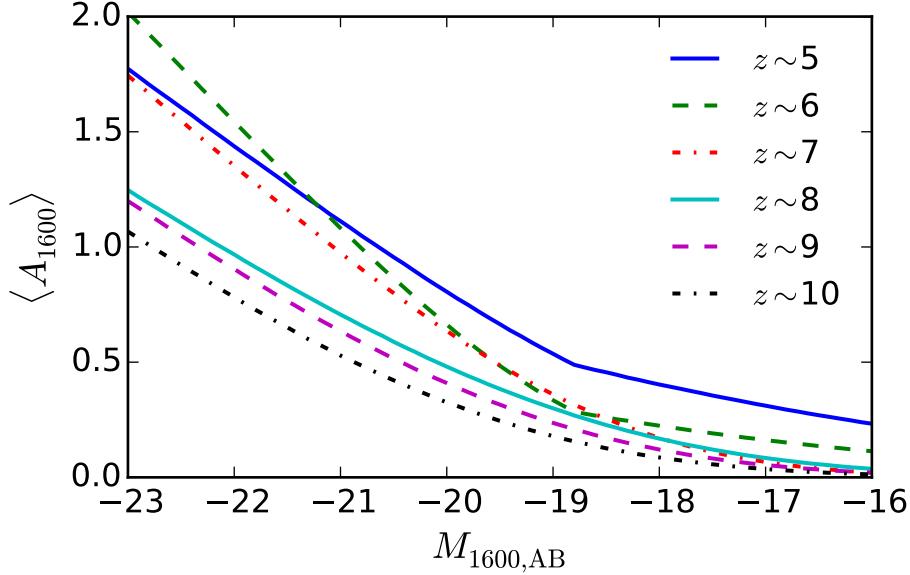


Figure 2.3: The average dust attenuation at 1600Å as a function of (dust-attenuated) luminosity at $z \sim 5\text{--}10$. The linear relation between A_{1600} and the slope of UV continuum β from Meurer et al. (1999) is used. Rest frame observations for β are from Bouwens et al. (2014b).

our model galaxies using a model of IGM absorption. We adopt the best fit effective optical depth of Ly- α absorption from Fan et al. (2006b), listed in Equation 1.28 at $z < 5.5$ and Equation 1.29 at $z = 5.5\text{--}6.3$. For simplicity, we adopt this relation for all redshifts at $z \geq 5.5$. Although this extrapolation is unphysical, the observed Ly α flux vanishes at $z > 6$, so that this assumption does not bias the LBG selection. Fig. 2.2 shows an example SED of a galaxy at $z = 6$, which is assumed to have a stellar mass of $M_* = 10^8 M_\odot$ and a constant SFR over the past 100 Myr. In this figure, the attenuated SED (solid line) is transformed from the unattenuated SED (dashed line) using the Ly α transmission. The position of the Ly α absorption trough in the transmission curve will be shifted to different observer-frame wavelength depending on the redshift of the source. The attenuated SED of model LBGs is identified by the selection criteria shown in Chapter 3.

2.4.3 Dust attenuation

To compare our luminosities with observations, we need to add the effect of dust attenuation. The rest-frame UV continuum for a galaxy is assumed to have the form

$$f_\lambda \propto \lambda^\beta, \quad (2.29)$$

where f_λ is the flux density per wavelength interval and β is the power-law index. For high- z galaxies, β can be estimated through photometric SED fitting (Bouwens et al., 2012, 2014b). UV flux can be strongly attenuated by dust grains within galaxies. This is parameterized as

$$F_o(\lambda) = F_i 10^{-0.4A_\lambda}, \quad (2.30)$$

where F_i and F_o are the intrinsic and observed continuum flux densities respectively, and A_λ is the change in magnitude at rest-frame wavelength λ . The amount of dust attenuation is wavelength dependent with larger optical depths for shorter wavelengths. Dust attenuation therefore reddens the spectrum by steepening the observed spectral slope.

Dust attenuation can be estimated through a variety of indicators such as emission line ratios (e.g. Balmer series), the slope of the rest-frame continuum, the ratio between infrared and UV radiation (Meurer et al., 1999), the stellar mass of galaxies (Pannella et al., 2009; Heinis et al., 2014), and the SFR of galaxies (Reddy et al., 2006). In this work, we adopt a luminosity-dependent dust attenuation model (Bouwens et al., 2012; Smit et al., 2012; Bouwens et al., 2014b) which is summarized below.

Assuming a constant star-formation history, stellar population synthesis shows that galaxies have similar intrinsic UV continuum slopes (e.g. Leitherer & Heckman, 1995). The dust-attenuated UV continuum slope, β , of galaxies will therefore reflect the amount of dust attenuation. Meurer et al. (1999) established a relation between UV dust attenuation and observed UV continuum β :

$$A_{1600} = 4.43 + 1.99\beta, \quad (2.31)$$

where A_{1600} is the dust attenuation at 1600Å(rest-frame). This relation is calibrated by comparison with starburst galaxies in the local Universe, assuming that high- z galaxies have the same spectral properties as local galaxies (Meurer et al., 1999).

The key observable for determining the dust-attenuation at high redshifts is the value of β for high- z galaxies. Observational studies of high- z galaxies show that β is larger for galaxies with higher redshifts and lower luminosities (Bouwens et al., 2012, 2014b). Bouwens et al. (2014b) studied this relation using a large sample (>4000 sources) of galaxies at $z \sim 4-8$. They found a piece-wise linear relation between the mean of β and

M_{UV} for galaxies at $z \sim 4\text{--}6$:

$$\beta = \begin{cases} \frac{d\beta}{dM_{\text{UV}}}(M_{\text{UV,AB}} + 18.8) + \beta_{M_{\text{UV}}=-18.8}, & M_{\text{UV,AB}} \leq -18.8, \\ -0.08(M_{\text{UV,AB}} + 18.8) + \beta_{M_{\text{UV}}=-18.8}, & M_{\text{UV,AB}} > -18.8, \end{cases} \quad (2.32)$$

where $d\beta/dM_{\text{UV}}$ and $\beta_{M_{\text{UV}}=-18.8}$ are from Table 4 of Bouwens et al. (2014b). We use this piece-wise relation for our model galaxies at $z \sim 5$ and 6.

For galaxies at $z \sim 7$ and 8, we use the linear relation

$$\beta = \frac{d\beta}{dM_{\text{UV}}}(M_{\text{UV,AB}} + 19.5) + \beta_{M_{\text{UV}}=-19.5}, \quad (2.33)$$

where $d\beta/dM_{\text{UV}}$ and $\beta_{M_{\text{UV}}=-19.5}$ are from Table 3 of Bouwens et al. (2014b).

Measurements of β at $z \gtrsim 9$ are limited (e.g. Wilkins et al., 2016). We assume that the linear mean β -luminosity dependence in Equation 2.33 remains valid at $z \sim 9$ and 10. We estimate $\beta_{M_{\text{UV}}=-19.5} = -2.19$ and -2.16 for $z \sim 9$ and 10 respectively by linearly fitting the observations (Bouwens et al., 2014b) at $z \sim 4\text{--}8$. We set $d\beta/dM_{\text{UV}} = -0.16$ for $z \sim 9$ and 10, which equals the mean at $z \sim 4\text{--}8$. The uncertainty in this relation is large. However, galaxies at $z > 9$ are usually faint and dust will not significantly attenuate the UV continuum.

We assume β is normally distributed around the mean value with a standard deviation of $\sigma = 0.35$ (Bouwens et al., 2014b) at all redshifts. From the linear relation in Equation 2.31, this means that A_{1600} is also normally distributed. Following Smit et al. (2012), we set $A_{1600} = 0$ if $A_{1600} < 0$ and then calculate the mean, $\langle A_{1600} \rangle$. Figure 2.3 shows $\langle A_{1600} \rangle$ as a function of observed (dust-attenuated) luminosity for different redshifts. We can then obtain the relation between $\langle A_{1600} \rangle$ and the intrinsic UV luminosity at 1600Å using Equation 2.30. The intrinsic rest-frame magnitude M_{1500}^i and M_{1600}^i are calculated from SEDs using tophat bands which have a width of 100Å and centres of 1500 and 1600Å respectively.²

Dust attenuation in other UV bands can then be estimated using the reddening curve normalized by $\langle A_{1600} \rangle$. A commonly adopted reddening curve was derived by

²Calculating both M_{1500} and M_{1600} as M_{UV} is for the convenience of comparison with observations. The difference between the calculated values of M_{1500} and M_{1600} is negligible ($|M_{1500} - M_{1600}| \lesssim 0.05$).

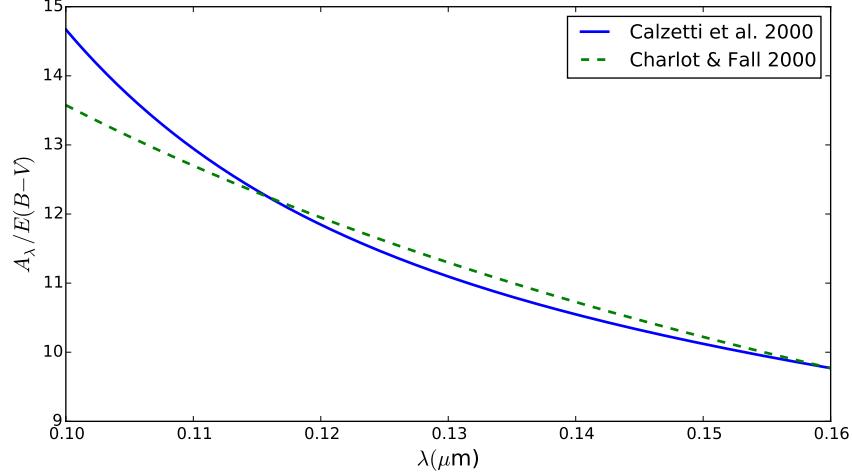


Figure 2.4: The comparison between the Calzetti et al. (2000) reddening curve and the Charlot & Fall (2000) reddening curve at $1000\text{\AA} < \lambda < 1600\text{\AA}$. The amount of dust extinction at 1600\AA is determined by the luminosity dependent continuum slopes and the Meurer et al. (1999) relation. The reddening curve determines the relative attenuation (reddening) at shorter wavelengths. We see that these two models show a very similar reddening slope.

Calzetti et al. (2000):

$$k(\lambda) = \begin{cases} 2.659 \left(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3} \right) + R_V, & 0.12\mu\text{m} \leq \lambda < 0.63\mu\text{m}, \\ 2.659 \left(-1.857 + \frac{1.040}{\lambda} \right) + R_V, & 0.63\mu\text{m} \leq \lambda < 2.20\mu\text{m}, \end{cases} \quad (2.34)$$

where the rest-frame wavelength, λ , is in units of μm , $R_V=4.05\pm0.80$ is the effective obscuration in the V band, and the coefficients are normalized to $E(B-V) = k(B) - k(V) = 1$. The Calzetti et al. (2000) reddening curve can also be approximated by a simpler reddening curve with extinction optical depth $\tau_\lambda \propto \lambda^{-0.7}$ at $0.10\mu\text{m} < \lambda < 0.16\mu\text{m}$ from Charlot & Fall (2000). The comparison between these two reddening curves are shown in Fig. 2.4. The change of magnitude due to dust attenuation is $A_\lambda = E(B-V)k(\lambda)$. To obtain the A_λ for $\lambda < 0.12\mu\text{m}$, we extrapolate the reddening curve $k(\lambda)$ to $\lambda < 0.12\mu\text{m}$.

2.5 Summary

In this chapter, we described the N -body simulations *Tiamat* (Poole et al., 2016), the semi-analytic model of galaxy formation, MERAXES (Mutch et al., 2016a), and the calculation of galaxy luminosities used in this thesis. When compared with other simu-

lations and galaxy formation models, these have a series of features that are particularly suitable for studying the formation of high- z galaxies and the process of reionization.

- The fiducial *Tiamat* simulation was run in a periodic box with side length $L = 67.8 h^{-1} \text{Mpc}$. This simulation volume is sufficient to capture the evolving structure of the reionization field. The *Tiamat/Tiny Tiamat* simulations have mass resolution of $2.64 \times 10^6 h^{-1} M_\odot / 6.9 \times 10^4 h^{-1} M_\odot$, which allows for the identification of low mass dark matter haloes ($M_{\text{vir}} \sim 10^8 M_\odot$) with the virial temperature close to the hydrogen cooling limit (Barkana & Loeb, 2001) across the redshifts relevant to reionization. *Tiamat* outputs include 100 snapshots from $z = 35$ to 5 with a constant temporal resolution of 11 Myr per snapshot. This high cadence resolves the dynamical time of galaxy discs at high redshifts, and is comparable to the lifetime of the progenitor stars of type II supernovae. The horizontal form of dark matter halo merger trees from *Tiamat* allows MERAXES to implement a self-consistent and efficient calculation of reionization structure and UVB feedback.
- Based on the high snapshot cadence of dark matter merger trees, MERAXES implements a delayed supernova feedback where a supernova may explode several snapshots after the star-formation episode. This enhances the accuracy of the supernova feedback compared to the commonly adopted instantaneous supernova feedback. Similarly, MERAXES also implements a delayed metal enrichment and stellar mass recycling scheme.
- To study the evolution of reionization structure and the UVB feedback on galaxy formation, MERAXES is the first semi-analytic galaxy formation model to couple the process of reionization (Mesinger et al., 2011) and the formation and evolution of galaxies both temporally and spatially.
- The rest-frame UV luminosities and multi-band photometric luminosities are calculated using the stellar population synthesis method. Ly α absorption is calculated using the observed Gunn-Peterson optical depth (Fan et al., 2006b) and dust extinction is estimated using an empirical luminosity-dependent extinction model (Meurer et al., 1999; Calzetti et al., 2000; Smit et al., 2012) with the new observations of UV-continuum slope of high- z galaxies (Bouwens et al., 2014b).

With these features, MERAXES provides a framework to model galaxy formation at high redshifts and make predictions for galaxy observables during the EoR. Studies of various simulation results will be shown in next chapters.

Chapter 3

UV luminosity functions

3.1 Introduction

The LF is one of the most fundamental observables for high- z galaxies. The intrinsic UV continuum of galaxies is determined by young stellar populations, and is therefore thought to be a good tracer for understanding the star-formation processes in galaxies. As a result, measurement of the UV LF constrains the buildup of stellar mass in the early Universe and the role of galaxies in reionization. Successfully describing the UV LF of galaxies at $z > 6$ is therefore essential for any models aiming to understand galaxy formation at high redshift.

In Chapter 2, we introduced the specially designed semi-analytic model of galaxy formation, MERAXES, and the calculation of the rest-frame UV and observer-frame multi-band luminosities. In this chapter, we present the calculated UV LFs, compare them with observations, and investigate the faint-end slope below current detection limits.

Galaxies at high redshift are too faint to be observed spectroscopically. However, they can be identified using multi-band photometry and selected using the Lyman-break technique (Steidel et al., 1996). Using this technique, significant progress has been made in the past two decades in characterizing the observed UV LF of galaxies towards higher redshifts (Steidel et al., 1999; Bouwens et al., 2007, 2010, 2011, 2014a; Oesch et al., 2010b; McLure et al., 2010; Finkelstein et al., 2012; Schenker et al., 2013; McLure et al., 2013; Duncan et al., 2014; Schmidt et al., 2014). The most comprehensive UV LF measurements to date at $z > 4$ were made by Bouwens et al. (2015b) and Bouwens et al. (2016b), based on the assembly of HST data sets including CANDELS, HUDF09, HUDF12, ERS and BORG/HIPPIES programmes. The large number (> 10000) of galaxies at $z \geq 4$ provide statistically reliable UV LFs for testing our semi-analytic model of galaxy formation during the reionization era.

Young galaxies are strong emitters of UV radiation. High-redshift star-forming galaxies are generally considered to be the main energy source for ionizing the IGM during the EoR. However, due to extinction by dust in the ISM, only a small fraction of UV photons can escape from galaxies (e.g. $f_{\text{esc}} \lesssim 0.2$) and ionize the IGM. Galaxies above current detection limits (e.g. $M_{\text{UV}} \sim -17$ at $z \sim 6$) are not sufficiently numerous

to maintain reionization. Rather, empirical studies and simulations show a galaxy population down to $M_{\text{UV}} = -13$ is required (e.g. Robertson et al., 2013; Duffy et al., 2014; Robertson et al., 2015; Bouwens et al., 2015a). The faint-end slope of the UV LF is therefore very important since it determines the number of reionizing photons emitted from the faint galaxies below current detection limits. On the other hand, theoretically, we expect that baryons in very low-mass dark matter haloes ($\lesssim 10^8 M_\odot$) cannot efficiently cool and form stars, implying that the LFs may have a truncation at a very faint luminosity.

Observational and numerical studies have investigated the shape of the UV LF at faint luminosities. Atek et al. (2015) recently obtained LFs down to $M_{\text{UV}} = -15.25$ at $z \sim 7$ behind lensing clusters, and found that the faint-end slope remains steep. Using the star-formation histories of Local Group dwarf galaxies obtained from a colour-magnitude diagram analysis, Weisz et al. (2014) inferred the LF at $z \sim 5$ down to $M_{\text{UV}} \sim -5$, and found no truncation. In contrast, using high-resolution cosmological hydrodynamic simulations, Wise et al. (2014) found the slope of LFs at $z > 7$ is flat at $M_{\text{UV}} > -12$. Further, O’Shea et al. (2015) found that the slope is flat for faint luminosities at $z > 12$ from calculations with a larger simulation volume. The DRAGONS simulation provides a framework within which we can consider the faint end of the LF within a self-consistent calculation of the reionization history.

This chapter is organized as follows. We first summarize our Lyman-break colour selection criteria in Section 3.2. We then show the calculated UV LFs of LBGs in Section 3.3. In Section 3.4, we calculate the fraction of total UV flux above observed luminosity limits. Finally, in Section 3.5, we present our conclusions.

3.2 Lyman-break selection

High-redshift galaxies can be selected using multi-band photometric surveys and the Lyman-break technique. To facilitate direct comparison with observed UV LFs and to study the completeness of LBG selections, we adopt the LBG colour selection criteria from Bouwens et al. (2015b) to select the model galaxies at $z \sim 5\text{--}8$:

- The colour selection criterion for $z \sim 5$ is:

$$(V_{606} - i_{775}) > 1.2 \text{ and } (z_{855} - H_{160} < 1.3) \\ \text{and } (V_{606} - i_{775} > 0.8(z_{855} - H_{160}) + 1.2). \quad (3.1)$$

- For $z \sim 6$:

$$(i_{775} - z_{850}) > 1.0 \text{ and } (Y_{105} - H_{160} < 1.3) \\ \text{and } (i_{775} - z_{850} > 0.78(Y_{105} - H_{160}) + 1.0). \quad (3.2)$$

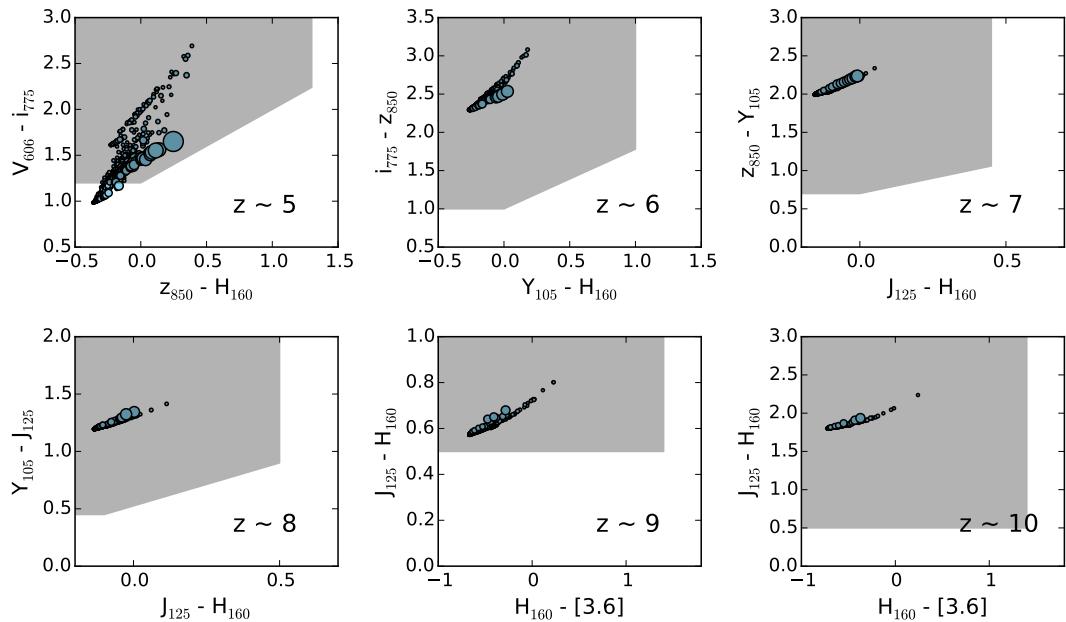


Figure 3.1: LBG selection criteria used to select star-forming galaxies at $z \sim 5\text{--}10$. The blue circles show randomly selected galaxies representing 5 per cent of the total sample with $M_{1600} < -15.75$. The areas of circles are proportional to their observed UV luminosities at rest-frame 1600Å. The grey-shaded regions are the selection regions for LBGs.

- For $z \sim 7$:

$$(z_{850} - Y_{105}) > 0.7 \text{ and } (J_{125} - H_{160} < 0.45) \\ \text{and } (z_{850} - Y_{105} > 0.8(J_{125} - H_{160}) + 0.7). \quad (3.3)$$

- For $z \sim 8$:

$$(Y_{105} - J_{125}) > 0.45 \text{ and } (J_{125} - H_{160} < 0.5) \\ \text{and } (Y_{105} - J_{125} > 0.75(Y_{125} - H_{160}) + 0.525). \quad (3.4)$$

Here V_{606} , i_{775} , z_{850} , Y_{105} , J_{125} and H_{160} represent the magnitudes in ACS and WFC3/IR filter bands $F606W$, $F775W$, $F850W$, $F105W$, $F125W$ and $F160W$ respectively. We adopt the colour criteria from Bouwens et al. (2016b) to select the galaxies at $z \sim 9$ and 10:

$$(J_{125} - H_{160} > 0.5) \text{ and } (H_{160} - [3.6] < 1.4). \quad (3.5)$$

Here [3.6] represents the magnitude in the *Spitzer*/S-CANDELS 3.6 μm filter.

We calculate the observed dust-attenuated luminosities in these bands for all model galaxies at $z \sim 5$ –10, and pass these through the colour selection criteria. Our model gives the luminosities in dropout bands no matter how faint the galaxies are. We do not want to exclude these faint galaxies, so we do not apply the non-detection criteria. We also do not consider model galaxies at redshifts other than the centre of the selection window, and so there is no contamination by interlopers from other redshifts.

Fig. 3.1 shows the LBG colour–colour selection panels. The galaxies in these panels represent a random sample of 5 per cent of the galaxies with $M_{1600} < -15.75$, which are used for our UV LF determination in Fig. 3.2. The galaxies located in the grey regions are selected as star-forming LBG galaxies. The size of the circles represents the UV luminosity of galaxies at rest-frame 1600 \AA before dust attenuation is applied.

We see that all of our model galaxies are located in the selection regions for $z \sim 6$ –10. For $z \sim 5$, a few faint galaxies fall outside of the selection region. The UV-bright galaxies have moved towards the upper right due to dust attenuation in all panels. Our study shows that if we remove the effect of dust attenuation, a significant number of the brightest galaxies will be located outside of the selection region at $z \sim 5$, where LBG selection for the UV-brightest galaxies is very sensitive to the dust attenuation model.

3.3 UV LFs

Fig. 3.2 shows our model UV LFs, $\phi(M_{\text{UV}})$, for LBGs (i.e. the galaxies which passed the LBG selection criteria) selected at redshifts $z \sim 5$ –10. The observed UV LFs are from Bouwens et al. (2015b) at $z \sim 5$, 6, 7, 8 and 10, with additional points from Bouwens

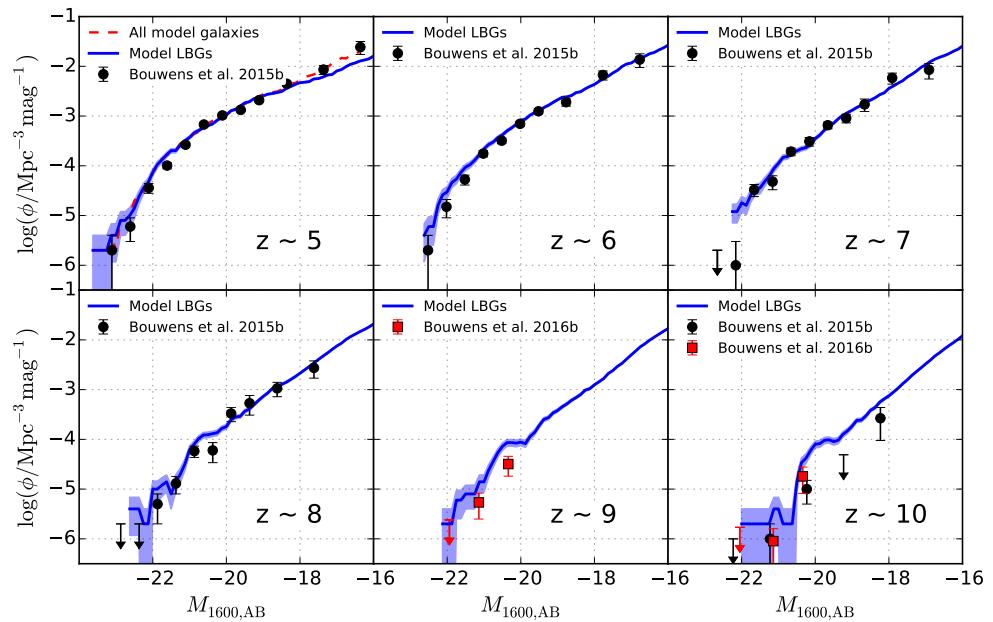


Figure 3.2: Model UV LFs at $z \sim 5\text{--}10$ from MERAKES. Blue solid lines show selected LBGs with 1σ Poisson uncertainties shown as shaded regions. Black circles and red squares are observational data from Bouwens et al. (2015b) ($z \sim 5, 6, 7, 8$ and 10) and Bouwens et al. (2016b) ($z \sim 9$ and 10) respectively. The red dashed line at $z \sim 5$ shows the UV LF for all model galaxies without the LBG selection. We see close agreement between the model and observations.

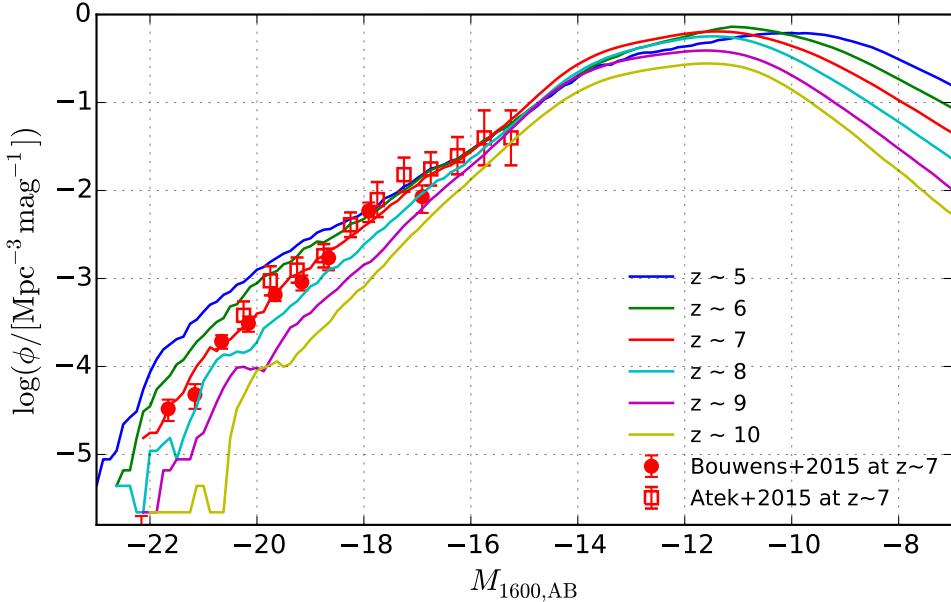


Figure 3.3: Model UV LFs from MERAXES for all model galaxies extended to low luminosities, illustrating the predicted flattening fainter than $M_{1600} \sim -14$. Red circles and red squares show the observed UV LF from (Bouwens et al., 2015b) and Atek et al. (2015) at $z \sim 7$ respectively.

et al. (2016b) at $z \sim 9$ and 10.

For the galaxies at $z \sim 5$, we also plot the UV LF for all model galaxies for comparison. We see a slight discrepancy between the LF for all model galaxies and for LBGs, due to the small fraction of faint galaxies with $M_{1600} > -19$ that do not pass the LBG selection criteria. For the galaxies at $z \sim 6\text{--}10$, all model galaxies with $M_{1600} < -15.75$ have passed the LBG selection and are identified as LBGs. Our model, which was calibrated to the stellar mass function at $z \sim 5\text{--}7$ MERAXES, produces UV LFs at $z \sim 5\text{--}10$ that are in excellent agreement with the observations.

We can use the model galaxies to study the shape and the evolution of UV LFs to much lower luminosities than observed, as shown in Fig. 3.3. To exclude the influence of LBG selection criteria, all model galaxies are used hereafter. The observed LFs from (Bouwens et al., 2015b) and Atek et al. (2015) at $z \sim 7$ are also shown. Atek et al. (2015) obtained the UV LF down to $M_{UV}=-15.25$ at $z \sim 7$. Our prediction is in good agreement with this observation.

We see that the slope of the UV LF remains steep at $M_{1600} < -14$ and becomes flat at $M_{1600} > -14$. The UV LFs have a turnover at $M_{1600} \approx -12$ and then drop towards fainter luminosities. The fact that the faint-end slope remains steep below current detection limits down to $M_{1600} = -14$ has important implications for the photon budget

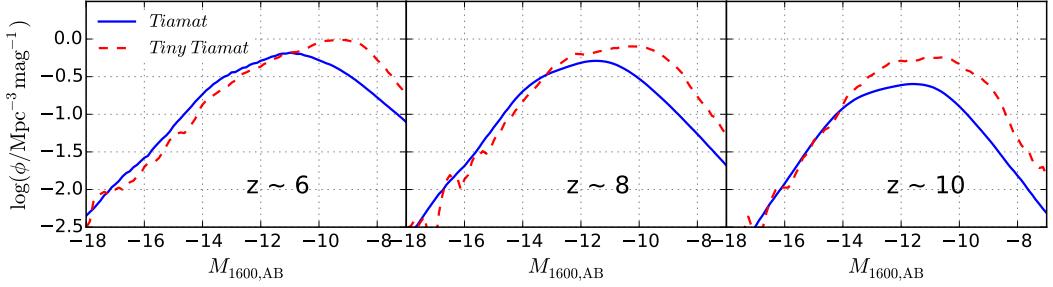


Figure 3.4: Comparison for UV LFs of all model galaxies based on *Tiamat* and *Tiny Tiamat* simulations at $z \sim 6, 8$ and 10 . The simulation box sizes are $67.8 h^{-1} \text{Mpc}$ and $10 h^{-1} \text{Mpc}$, and the particle masses are $2.64 \times 10^6 h^{-1} M_\odot$ and $6.79 \times 10^4 h^{-1} M_\odot$ for *Tiamat* and *Tiny Tiamat* respectively. We see a flattening slope at $M_{\text{UV}} \gtrsim -14$ from both simulations.

during reionization (see Section 3.4). The predicted turnover in the number density of faint galaxies can be traced to the condition that the halo mass $\sim 10^8 M_\odot$ must exceed the hydrogen cooling limit corresponding to a virial temperature of 10^4 K before stars can form. A larger value of the cooling mass or temperature will lead to a turnover at brighter UV magnitude (e.g. Muñoz & Loeb, 2011). This is also seen by comparing to the relation between the mass of dark matter haloes and UV luminosity (as discussed in Section 4.5). The flattening at $M_{1600} > -14$ of LFs is also a testable prediction of the luminosity below which it becomes likely for the halo masses to drop below the hydrogen cooling limit.

We also see that the slope of the UV LFs at $M_{1600} < -16$ steepens towards higher redshift. On the other hand, the slope at fainter magnitudes $-16 < M_{1600} < -10$ does not significantly evolve at $z > 5$. We infer that the implied continuous growth of extremely faint LFs reflects the ongoing formation of small haloes.

To investigate how the simulation volume and mass resolution affect the position of the turnover in the UV LF, we compare the predictions based on the *Tiamat* N-body simulation and the much higher mass resolution *Tiny Tiamat* N-body simulation (Poole et al., 2016) in Fig. 3.4. *Tiny Tiamat* has a high mass resolution ($6.79 \times 10^4 h^{-1} M_\odot$) which easily resolves the hydrogen cooling mass at all simulated redshifts. We see that the LFs based on *Tiamat* and *Tiny Tiamat* generally agree at $M_{1600} < -14$. However, the model UV LFs based on *Tiny Tiamat* flatten at $M_{1600} \gtrsim -12$ which is ~ 2 magnitudes fainter than those based on *Tiamat*. This difference quantifies the combined effects of the hydrogen cooling limit not being completely resolved in the *Tiamat* simulation until $z \lesssim 6$, together with merger-triggered star formation in the small haloes near the cooling limit that cannot be resolved by *Tiamat*.

The flattening slope of high- z LFs at $M_{\text{UV}} > -14$ has been previously seen in other

simulations. For example, O’Shea et al. (2015) carried out a suite of hydrodynamic simulations with an adaptive mesh refinement code (the *Renaissance Simulations*) which employed a self-consistent radiative transfer reionization scheme and included Population III stars (see also Wise et al., 2012, 2014). These simulations have a dark matter particle mass of $2.9 \times 10^4 M_\odot$, and show that the $z > 12$ LF flattens at $M_{\text{UV}} \gtrsim -14$, in good agreement with our results.

3.4 UV flux from galaxies below the detection limit

An important quantity for studies of reionization is the ionizing luminosity emitted by the overall population of galaxies. There has been extensive discussion in the literature regarding whether enough star formation has been observed to complete reionization (e.g. Fan et al., 2006a; Robertson et al., 2010, 2013, 2015; Bouwens et al., 2015a). MERAXES describes the stellar mass function of galaxies at $z \sim 5\text{--}7$, and the UV LF in the observed range from $z \sim 5\text{--}10$. The predicted UV LFs can be used to calculate the UV luminosity density originating from galaxies above a luminosity threshold L_{lim} :

$$\rho_{\text{UV}} = \int_{L_{\text{lim}}}^{\infty} L \phi(L) dL. \quad (3.6)$$

The emissivity of galaxies (number of ionizing photons emitted into the IGM per second per comoving volume) can then be estimated using

$$\epsilon = f_{\text{esc}} \xi_{\text{ion}} \rho_{\text{UV}}, \quad (3.7)$$

where ξ_{ion} is the number of ionizing photons per unit UV luminosity and f_{esc} is the fraction of ionizing photons that escape from the galaxy to ionize the IGM.

Fig. 3.5 shows the emissivity from galaxies brighter than the limit M_{lim} as a fraction of the total emissivity from all galaxies in the simulation at each redshift:

$$\frac{\epsilon_{(\leq M_{\text{lim}})}}{\epsilon_{\text{total}}} = \frac{\int_{-\infty}^{M_{\text{lim}}} \phi(M) L(M) dM}{\int_{-\infty}^{\infty} \phi(M) L(M) dM}. \quad (3.8)$$

Here we assume f_{esc} is constant for all galaxies at each redshift, although it likely scales with halo mass, possibly compensating for the relative inefficiency in star formation of the faint galaxies (e.g. Wise & Cen, 2009; Kuhlen & Faucher-Giguère, 2012; Paardekooper et al., 2015). We calculate $\epsilon/\epsilon_{\text{total}}$ for both simulations based on *Tiamat* and *Tiny Tiamat*. *Tiny Tiamat* misses the brightest galaxies due to the limited volume, and so presents conservatively low limits on the total fraction of observed flux. However, we find good agreement between estimates of faint galaxy flux levels among the simulations indicating that our model based on *Tiamat* is not missing significant luminosity.

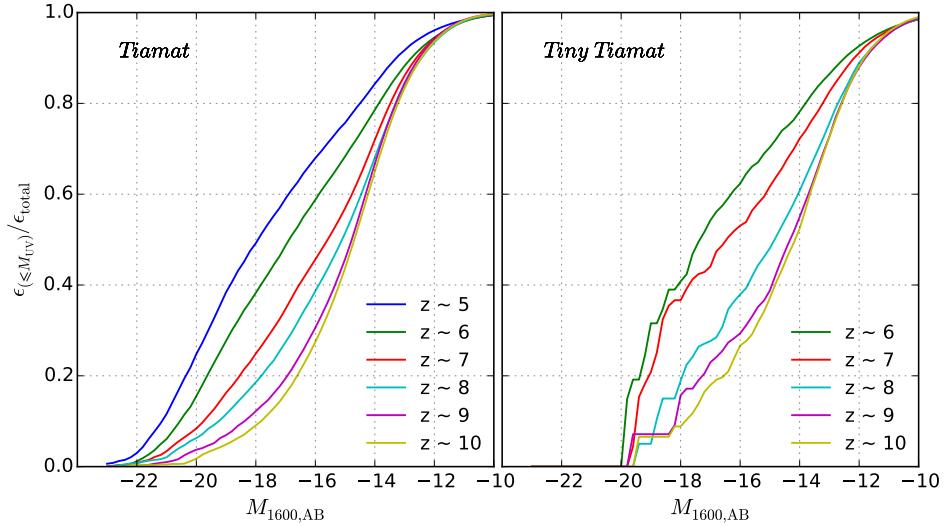


Figure 3.5: The cumulative fraction of 1600Å UV flux from the model galaxies brighter than the luminosity limit M_{1600} based on *Tiamat* (left-band panel) and *Tiny Tiamat* (right-band panel) N -body simulations. We see that more than 50 per cent of total UV flux are from galaxies fainter than $M_{\text{UV}}=-17$ at $z>6$.

Table 3.1: The fraction of UV flux at 1600Å above the luminosity limits.

z	$M_{1600} \leq -17$	$M_{1600} \leq -13$	$M_{1600} \leq -10$	
<i>Tiamat</i>	5	0.580	0.916	0.994
	6	0.478	0.884	0.995
	7	0.348	0.866	0.996
	8	0.280	0.845	0.996
	9	0.202	0.850	0.997
	10	0.167	0.840	0.997
<i>Tiny Tiamat</i>	5	—	—	—
	6	0.537	0.861	0.985
	7	0.437	0.821	0.986
	8	0.277	0.756	0.986
	9	0.230	0.729	0.988
	10	0.177	0.725	0.990

The luminosity contributions from galaxies brighter than $M_{1600} = -17, -13$ and -10 are shown in Table 3.1. Under the assumption of an escape fraction of ionizing radiation that does not depend on mass or redshift, this fraction of total luminosity equals the fraction of ionizing photons.

In Fig. 3.5, the *Tiny Tiamat*-based simulation shows a truncation of flux at $M_{\text{UV}} < -20$, indicating that simulation volume influences the flux contribution from the brightest galaxies, especially at $z < 7$. However, both simulations give similar values for fractional flux at $-18 < M_{\text{UV}} < -13$. We find that the fraction of ionizing flux from galaxies brighter than the limit $M_{\text{UV}} = -17$ evolves continuously from 17 per cent at $z \sim 10$ to 58 per cent at $z \sim 5$. This implies that bright galaxies contribute a greater fraction of UV flux at lower redshift than at high redshift. We see that faint galaxies below a detection limit of $M_{\text{UV}} = -17$ at $z \sim 6$ (10) provide more than 52 (83) per cent of the total flux, and are therefore likely to be the main source of ionizing photons for reionization. At $z > 7$, galaxies with luminosities in the range $-17 < M_{\text{UV}} < -13$ provide more than ~ 50 per cent of total UV flux. These results are in agreement with the findings of our hydrodynamic simulations (*Smaug*; see Duffy et al., 2014). Due to their inefficient formation, the faintest galaxies ($M_{\text{UV}} > -10$) contribute < 1 per cent ionizing flux at $z \sim 5-10$. Therefore, within the standard model of galaxy formation with a minimum halo mass for star formation as implemented in MERAXES, the UV flux from these faintest galaxies is negligible during the EoR, and a magnitude of $M_{\text{UV}} \sim -10$ can be considered as an appropriate integration cutoff for luminosity density calculations.

Before leaving this section, we note that although our model successfully reproduces the galaxy UV LFs, it includes a number of assumptions (e.g., IMF, dust, metallicity and binary populations) which could affect the UV luminosity of our model galaxies.

3.5 Conclusions

In this chapter we present UV LFs for model galaxies during the EoR predicted by the semi-analytic model MERAXES. Since MERAXES was calibrated using the stellar mass function at $z \sim 5-7$ (see Chapter 2), we need to mimic the LBG selection and compare the predicted UV LFs from MERAXES with observations. The predicted UV LFs are in agreement with observations, so we can use these to study the faint-end slope below current detection limits and estimate the contribution of ionizing photons from these faint galaxies. Through this procedure, we find:

- Having been calibrated to the stellar mass function at $z \sim 5-7$, our model successfully reproduces the UV LF for high- z star-forming galaxies at $z \sim 5-10$ (see Fig. 3.2). The slope of the predicted UV LF remains steep below current detection limits, and becomes flat at $M_{\text{UV}} \gtrsim -14$ before declining below $M_{\text{UV}} \sim -12$ (see Fig. 3.3). This prediction will be testable in the future based on observations of

faint galaxies with *JWST* and lensing.

- The majority (84–92 per cent) of UV flux at $z \sim 5$ –10 is produced in galaxies with $M_{\text{UV}} < -13$. At $z \sim 5$, the flux is dominated by the UV-bright galaxies ($M_{\text{UV}} \leq -17$). However, At $z \geq 7$, galaxies with $-17 \lesssim M_{\text{UV}} \lesssim -13$ are the dominant emitters of UV flux compared to their bright counterparts (see Fig. 3.5).

The above conclusions depend on modeling galaxy formation in low-mass dark matter haloes with viral temperatures close to the hydrogen cooling limit ($M_{\text{vir}} \sim 10^8 M_{\odot}$). Although the fiducial *Tiamat* simulation already has a mass resolution of $m_{\text{DM}} \sim 2.64 \times 10^6 h^{-1} M_{\odot}$, in order to resolve these low-mass haloes, we have also used semi-analytic model based on dark matter merger trees from the *Tiny Tiamat* N -body simulation which has a much higher mass resolution of $m_{\text{DM}} = 6.79 \times 10^4 h^{-1} M_{\odot}$. In *Tiny Tiamat*, a $M_{\text{vir}} \sim 10^8 M_{\odot}$ halo is represented by $\sim 10^3$ dark matter particles, so that the galaxy formation history in these small haloes can be well described. The calculations based on the *Tiny Tiamat* simulation confirm that the faint-end slope of the UV LF remains steep until at least $M_{\text{UV}} \sim -14$, and further suggests that the faint-end steepening may extend ~ 1 –2 mag fainter (see Fig. 3.4). Using cluster lensing, Bouwens et al. (2016a) have recently demonstrated that current observations permit a turn-over as bright as $M_{\text{UV}} \sim -15.3$ to -14.2 . More constraints will be made by *JWST* in the near future, which will observe faint galaxies down to $M_{\text{UV}} \sim -14$ at $z \sim 6$.

The second conclusion listed above has been also confirmed by the *Tiny Tiamat* based simulation (see 3.5). We note that a universal escape fraction for ionizing radiation has been assumed in our work. However, some simulations show that the escape fraction may decreases with increasing halo mass and cosmic time (Ferrara & Loeb, 2013). In such cases, the high- z faint galaxies below current detection limits will contribute more ionizing photons to reionization than estimated here.

The semi-analytic model used in this work does not include supermassive blackholes which can potentially generate ionizing photons and provide feedback for galaxy formation during the EoR. Qin et al. (2017) recently updated MERAXES with central supermassive blackholes included. Qin et al. (2017) showed that supermassive blackholes are unlikely to make significant contribution to ionizing photons and AGN feedback on galaxy formation can be neglected during the EoR.

Chapter 4

Probing intrinsic galaxy and halo properties

4.1 Introduction

Intrinsic galaxy properties including SFR and stellar mass are studied through indirect observable probes. UV luminosity is one such probe for high- z galaxies which lack spectroscopic information (see Madau & Dickinson, 2014). The relationship between UV luminosity and intrinsic galaxy properties is usually calculated using a stellar population synthesis model with an assumed star-formation history (e.g. Kennicutt, 1998a; Madau et al., 1998). However, due to the shorter dynamical times and the high prevalence of mergers, galaxies at high redshifts have very different ages and may undergo different evolution and merger histories. In this chapter, we study the luminosity–SFR relation and the luminosity–stellar mass relation of galaxies at high redshifts and fainter magnitudes.

The mass and spatial distribution of dark matter haloes is also essential for understanding high- z galaxy formation. To probe dark matter haloes using observables, we need to study the connection between halo mass and galaxy luminosity. There are multiple ways to link galaxies with dark matter. One analytic approach is the halo occupation distribution (HOD; Berlind & Weinberg, 2002). This is usually represented in terms of the probability distribution $P(N|M)$, which describes the probability that a halo with virial mass of M contains N galaxies of a given type (central or satellite). The HOD relies on a number of control parameters such as the minimum halo mass hosting a galaxy and the concentration of halo density profile, and does not involve any processes of galaxy evolution. Another approach is subhalo abundance matching (SHAM; e.g. Vale & Ostriker, 2004; Conroy et al., 2006; Trujillo-Gomez et al., 2011; Mason et al., 2015), which ranks and assigns the mass of haloes to the luminosity or mass of galaxies. The halo mass is generated from the Press-Schechter mass function or dark matter N -body simulations, whilst galaxy properties are obtained from observed LFs or stellar mass functions. SHAMs assume a monotonic relationship between halo and galaxy mass/luminosity and usually neglects the statistical scatter in the galaxy–halo assignment. An alternative and potentially more powerful approach to determining the halo–observable connection is to use galaxy formation simulations. In this chapter, we

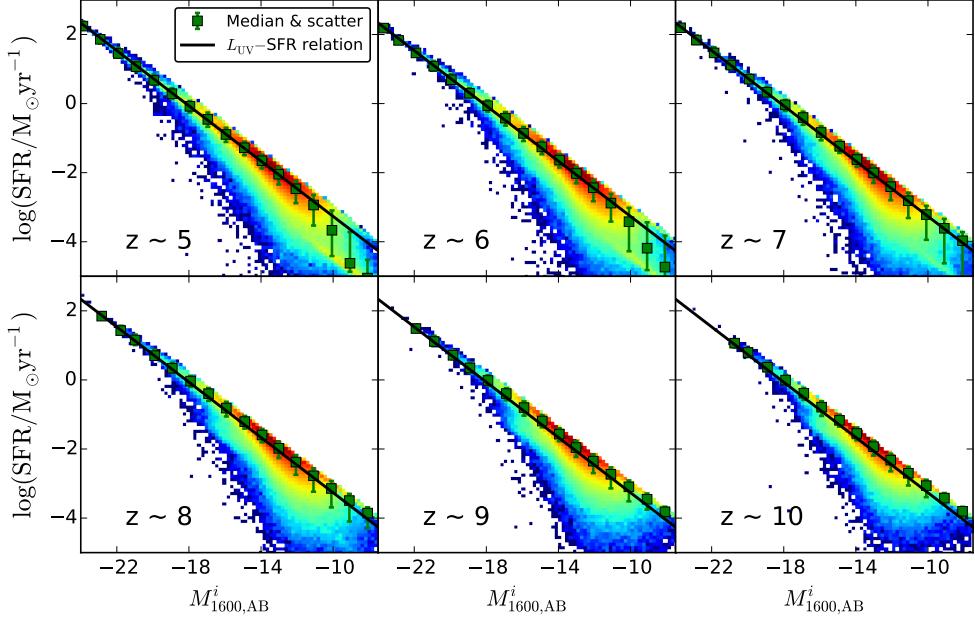


Figure 4.1: The instantaneous SFR of galaxies as a function of their (unattenuated) intrinsic UV luminosity M_{1600}^i at $z \sim 5-10$. The colour profile represents the logarithm density of the distribution. The green squares and error bars represent the median and 16th to 84th percentiles in bins which contain at least 5 model galaxies respectively. The black lines represent the linear relation between L_{UV} and SFR shown in Equation 4.1 with $\mathcal{K}_{\text{UV}} = 1.25 \times 10^{-28}$ (Kennicutt, 1998a; Madau et al., 1998). We see that the Madau-Kennicutt conversion describes the correlation between the UV luminosity and the instantaneous SFR from our model well.

use MERAXES to study the relationship between the UV luminosity of galaxies and the mass of dark matter haloes.

This chapter is organized as follows: We first study the relationship between UV luminosity and SFR in Section 4.2. The modelled SFR functions are shown in Section 4.3. We then investigate the relationship between UV luminosity and stellar mass of galaxies in Section 4.4. The connection between galaxy luminosity and the mass of its host dark matter halo is presented in Section 4.5. Finally, in Section 4.6, we present our conclusions.

4.2 UV luminosity–SFR relation

Since UV flux comes mostly from massive and short-lived stars, the intrinsic UV luminosity is proportional to SFR and independent of star-formation history over time-scales of $t \geq t_{\text{MS}}$, where t_{MS} is the main-sequence time of massive stars (Madau et al., 1998). UV luminosity is therefore thought to be a good indicator of SFR (Kennicutt, 1998a),

via the linear relation

$$\frac{\text{SFR}}{\text{M}_\odot \text{yr}^{-1}} = \mathcal{K}_{\text{UV}} \frac{L_{\text{UV}}}{\text{erg} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1}}, \quad (4.1)$$

where the L_{UV} is the intrinsic UV luminosity and \mathcal{K}_{UV} is a constant. The value of \mathcal{K}_{UV} is model dependent, but can be calibrated via a stellar synthesis model which depends on the IMF, metallicity and star-formation history. Using a Salpeter IMF in the range 0.1–125 M_\odot and an exponential burst of star formation with time-scale $\gtrsim 1$ Gyr, Madau et al. (1998) obtained $\mathcal{K}_{\text{UV}} = 1.25 \times 10^{-28}$ for L_{UV} in the wavelength range 1500–2800 Å. Assuming 100 Myr of constant star formation and a Salpeter IMF in the range 0.1–100 M_\odot , Kennicutt (1998a) calibrated the value to be $\mathcal{K}_{\text{UV}} = 1.4 \times 10^{-28}$. Wilkins et al. (2011) similarly obtained a value $\mathcal{K}_{\text{UV}} = 1.31 \times 10^{-28}$ using the STARBURST99 population synthesis model employed in this thesis.

However, at $z \gtrsim 6$ the age of the Universe is less than 1 Gyr, and many galaxies have star-formation histories shorter than 100 Myr. Therefore, we investigate how well the UV luminosity traces the instantaneous SFR using the variable star-formation histories from our model.

Fig. 4.1 shows the relation between the intrinsic UV luminosity and the instantaneous SFR for all galaxies in MERAXES at $z \sim 5$ –10. The distribution of SFRs at fixed stellar mass for the galaxies in these plots indicates the effect of variations in star-formation histories. The distribution shows a sharp upper limit, corresponding to the youngest galaxies which formed in the latest snapshot. We see that the model predicts a linear relationship between the UV luminosity and instantaneous SFR, which can be fitted well by the Madau et al. (1998) and Kennicutt (1998a) relation. The model galaxy SFRs are distributed with a scatter around the median of $\log_{10}(\text{SFR})$ varying from $\sigma \sim 0.2$ dex at $M_{1600}^i < -18$ to ~ 0.3 dex at $M_{1600}^i = -14$. The scatter is larger for UV-faint galaxies than for UV-bright galaxies. The distribution around the median L –SFR relation is not log-normal at faint luminosities in our model owing to the minimum star-formation timescale set by the finite temporal spacing of our simulation snapshots which cuts off the distribution at high SFR. We fit the L –SFR relation using Equation 4.1 for model galaxies brighter than $M_{\text{UV}} = -14$ and obtain

$$\mathcal{K}_{\text{UV}} = (1.13, 1.16, 1.17, 1.19, 1.24, 1.39) \times 10^{-28} \quad \text{at} \quad z \sim (5, 6, 7, 8, 9, 10). \quad (4.2)$$

The value of \mathcal{K}_{UV} slightly increases towards higher redshift due to the shorter galaxy-formation history and the higher fraction of young stars at higher redshift.

4.3 The SFR functions

An important quantity related to the buildup of stellar mass during reionization is the SFR function (e.g. Smit et al., 2012) which is shown in Fig. 4.2. The observed SFR

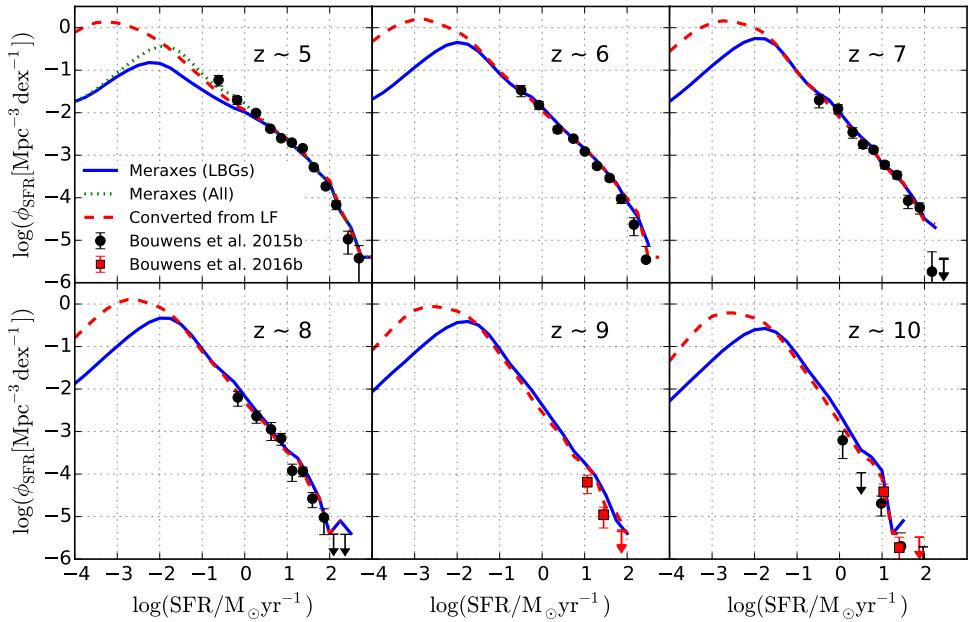


Figure 4.2: SFR functions of galaxies at $z \sim 5\text{--}10$. The blue lines show the SFR function for LBGs obtained directly from the semi-analytic model. The dotted green line is the SFR function obtained from the semi-analytic model but without LBG selection at $z \sim 5$. The red dashed lines show the model SFR converted from UV luminosity using the luminosity–SFR relation with $\mathcal{K}_{\text{UV}} = 1.25 \times 10^{-28}$ (Kennicutt, 1998a; Madau et al., 1998). Using this relation, black circles show the SFR functions converted from the observed UV LFs from Bouwens et al. (2015b) with dust correction ($z \sim 5, 6, 7, 8$ and 10). Red squares show the SFR functions converted from the observed UV LFs from Bouwens et al. (2016b) ($z \sim 9$ and 10) with dust correction applied. We see close agreement between the SFR functions calculated directly from the model and converted from the predicted UV luminosities.

function is estimated by converting the UV LF from Bouwens et al. (2015b, 2016b) using the Kennicutt (1998a) and Madau et al. (1998) relation. It is therefore important to investigate whether this assumed conversion introduces bias, and whether the scatter in the relationship effects the determination of the SFR function. We use two methods to derive the model SFR function: (i) we calculate the predicted instantaneous SFR function using the SFR calculated directly from the semi-analytic model; and (ii) to mimic observations, we convert the model UV luminosities to SFRs using Equation 4.1 with $\mathcal{K}_{\text{UV}} = 1.25 \times 10^{-28}$. The LBG selection is implemented for both of the above methods. Fig. 4.2 shows the derived model SFR functions together with the observational estimates.

There is a close agreement between the predicted model SFR functions and UV-derived model SFR functions at $\log_{10}(\text{SFR}/M_{\odot}\text{yr}^{-1}) > -2$, which in turn agree well with the observational estimates. The small difference between the predicted SFR function for LBGs and the UV derived SFR function at $z \sim 5$ is caused by the LBG colour selection criteria. However, the differences between model-predicted and UV-derived SFR functions at very low SFRs of $\log_{10}(\text{SFR}/M_{\odot}\text{yr}^{-1}) < -2$ show that the estimate of the SFR function using the Madau et al. (1998) and Kennicutt (1998a) conversion between UV luminosity and SFR will be biased by the scatter of the luminosity–SFR distribution.

4.4 UV luminosity–stellar mass relation

Galaxies which continuously form stars naturally produce a relation between luminosity and stellar mass. Fig. 4.3 shows stellar mass (M_*) as a function of observed UV luminosity (M_{1500}) for model galaxies predicted by MERAXES. We see that our model predicts UV-bright galaxies to have large stellar masses. The model galaxies have stellar masses that are distributed about the median at fixed luminosity with scatter $\sim 0.2\text{--}0.5$ dex depending on UV luminosity. For comparison we show the observed relation at $z \sim 5\text{--}7$ from Duncan et al. (2014), who measured the stellar mass–luminosity relation by fitting the observed photometric data with galaxy model SEDs. To compare with our model, the observational data with nebular emission excluded from the SED fitting is used. We convert the observed data from a Chabrier IMF to a Salpeter IMF by adding 0.24 to $\log_{10} M_*$ (Duncan et al., 2014). We linearly fit the $\log_{10} M_* - M_{\text{UV}}$ relation for our bright galaxies with $M_{1500} \leq -14$ using the relation¹

$$\log_{10} M_* = \frac{d \log_{10} M_*}{d M_{\text{UV}}} (M_{\text{UV}} + 19.5) + \log_{10} M_{*(M_{\text{UV}}=-19.5)}. \quad (4.3)$$

The best-fit slopes and intercepts are shown in Table 4.1 and Fig. 4.4.

At bright luminosities ($M_{1500} \lesssim -14$), the model galaxies are in good agreement

¹The $M_* - L_{\text{UV}}$ relation can be derived by substituting $M_{\text{UV}} = -2.5 \times \log_{10}(L_{\text{UV}}[\text{erg s}^{-1}\text{Hz}^{-1}]) + 51.6$.

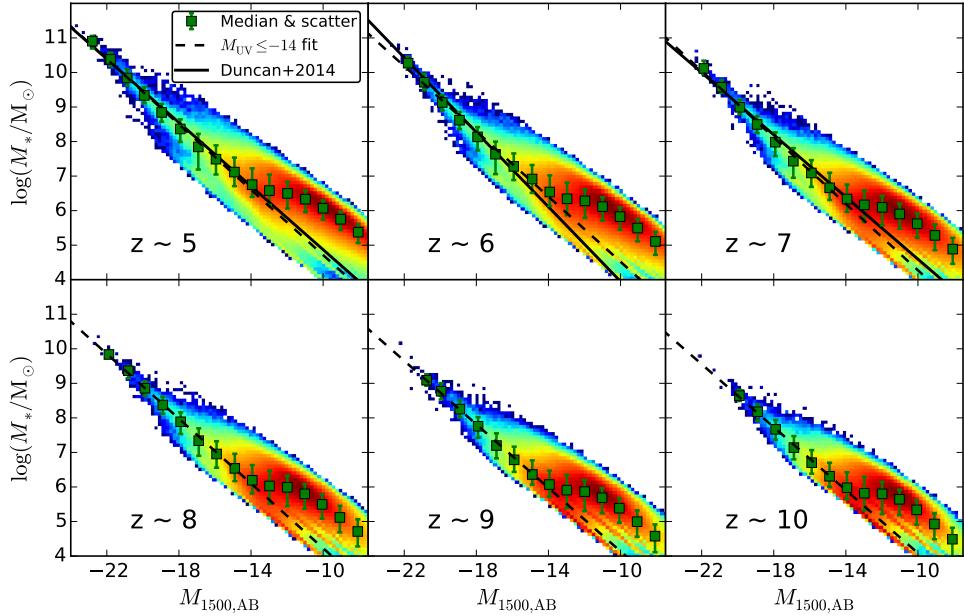


Figure 4.3: Stellar mass of galaxies as a function of UV luminosity M_{1500} . The colour profile represents the logarithm density of the distribution. The green squares and error bars represent the median and 16th to 84th percentiles in bins which contain at least 5 model galaxies. The black dashed lines are the linear fit to the medians at $M_{1600} \lesssim -14$. The black lines show the observationally fit lines from Duncan et al. (2014), where the stellar mass is converted from a Chabrier IMF to Salpeter IMF by adding 0.24 to $\log_{10} M_*$. We see a close agreement between our model and observations at $z \sim 5-7$ where the observational data is available.

Table 4.1: The best-fit slopes and intercepts of the median $\log_{10} M_* - M_{\text{UV}}$ relation (Equation 4.3) for galaxies with $M_{\text{UV}} \leq -14$.

z	$d \log_{10} M_*/d M_{\text{UV}}$	$\log_{10} M_{*(M_{\text{UV}}=-19.5)}$
5	-0.474 ± 0.013	9.21 ± 0.04
6	-0.471 ± 0.021	9.00 ± 0.05
7	-0.477 ± 0.014	8.83 ± 0.04
8	-0.470 ± 0.010	8.68 ± 0.03
9	-0.459 ± 0.012	8.50 ± 0.04
10	-0.456 ± 0.013	8.41 ± 0.05

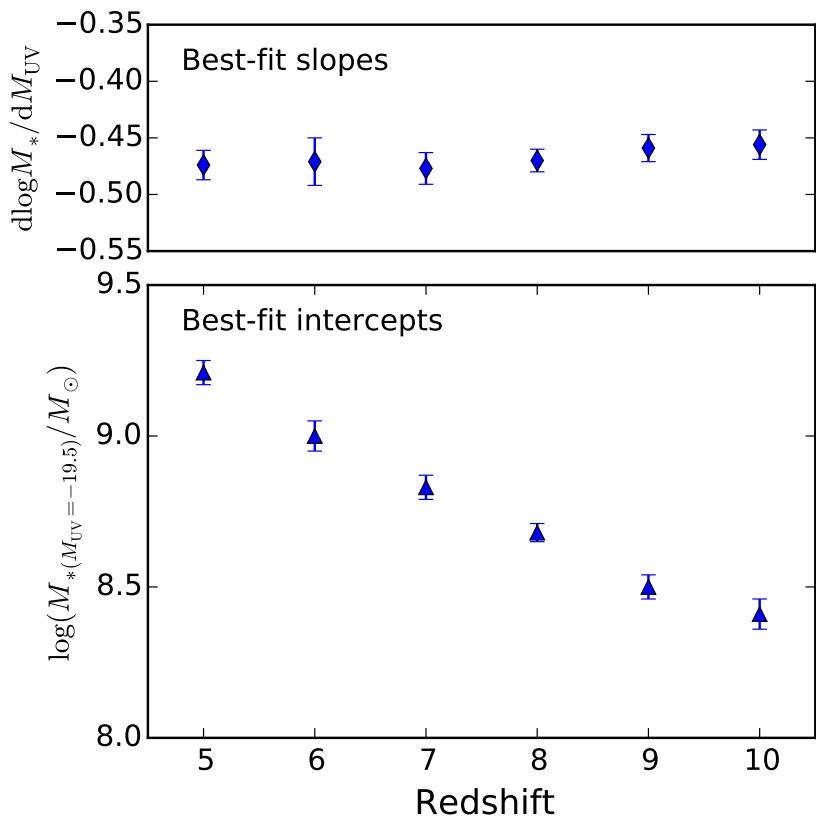


Figure 4.4: The best fitting slopes (upper panel) and the intercepts at $M_{\text{UV}}=-19.5$ (lower panel) of the $\log_{10} M_* - M_{\text{UV}}$ relation fit to the median trend for galaxies with $M_{\text{UV}} \leq -14$.

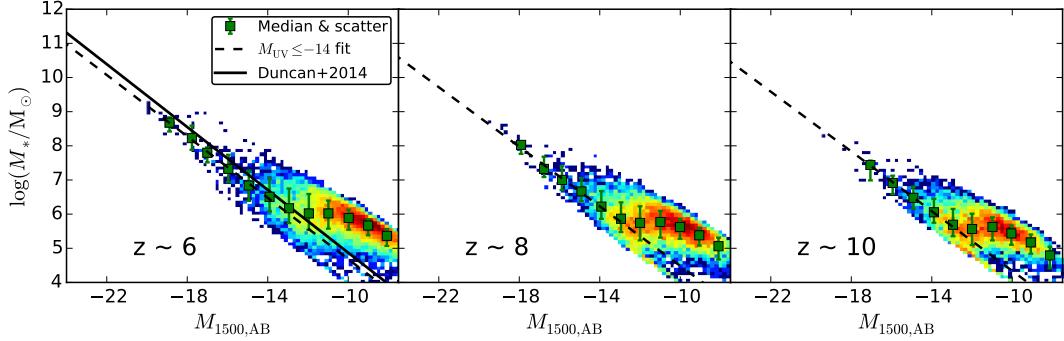


Figure 4.5: Same as Fig. 4.3 but for model galaxies at $z \sim 6, 8$ and 10 based on the *Tiny Tiamat* N-body simulation, which has a smaller volume but higher mass resolution than the *Tiamat* simulation.

with the observed mass–luminosity relation. The best-fit slope (with median ~ -0.47) for model galaxies at $z \sim 5\text{--}7$ does not significantly change with redshift, and is close to the observed slopes of $d \log_{10} M_*/dM_{\text{UV}} \sim -0.45$ to -0.54 from Duncan et al. (2014) at $z \sim 5\text{--}7$. Similar slopes are also found by Lee et al. (2012) and Shibuya et al. (2015).

We see that the best-fit intercept $\log_{10} M_{*(M_{\text{UV}}=-19.5)}$ evolves linearly with redshift. Assuming a constant slope $d \log_{10} M_*/dM_{\text{UV}} = -0.47$ for the bright galaxies at $z \sim 5\text{--}10$, we find the evolution of the luminosity–stellar mass relation can be estimated using the relation

$$\log_{10} \left(\frac{M_*}{10^8 M_\odot} \right) = -0.47(M_{1500} + 19.5) - 0.15(z - 7) + 0.86. \quad (4.4)$$

For fainter galaxies ($-14 < M_{\text{UV}} < -11$) the slope of the luminosity–stellar mass relation changes significantly to $d \log_{10} M_*/dM_{\text{UV}} \sim -0.1$. This flattening of the $\log_{10} M_*$ – M_{UV} relation depends on the mass resolution of our dark matter N-body simulations.

Since the dark matter haloes are well resolved in the model based on *Tiny Tiamat*, we also present the luminosity–stellar mass relation based on this high-resolution simulation in order to investigate whether the flattening at low luminosities seen in Fig. 4.3 is due to resolution effects. Fig. 4.5 shows the $\log_{10} M_*$ – M_{UV} relation for the model based on *Tiny Tiamat*. We linearly fit the relation for the bright galaxies with $M_{1500} \lesssim -14$. The best-fit slopes and intercepts are shown in Table 4.2. We find a similar luminosity–stellar mass relation based on *Tiny Tiamat* with best-fit slopes slightly larger than the slopes from the model based on *Tiamat*. We see that the luminosity–stellar mass relation based on *Tiny Tiamat* becomes flat at $M_{1500} \gtrsim -13$ which is ~ 1 magnitude fainter than the model based on *Tiamat*. This difference is due to the additional star formation in the merger triggered starburst phase in low-mass haloes of $M_{\text{vir}} \sim 10^8 M_\odot$.

Table 4.2: The best fitting slopes and the intercepts of the $\log_{10} M_* - M_{\text{UV}}$ relation for galaxies with $M_{\text{UV}} \leq -14$ based on the *Tiny Tiamat* N-body simulation.

z	$d \log_{10} M_*/dM_{\text{UV}}$	$\log_{10}(M_{*(M_{\text{UV}}=-19.5)})$
6	-0.447 ± 0.016	8.94 ± 0.05
7	-0.424 ± 0.015	8.71 ± 0.05
8	-0.435 ± 0.029	8.63 ± 0.11
9	-0.405 ± 0.017	8.42 ± 0.06
10	-0.437 ± 0.009	8.49 ± 0.04

Table 4.3: The best fitting slopes and the intercepts of the median $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation (Equation 4.5) for galaxies with $M_{\text{UV}} \leq -14$.

z	$d \log_{10} M_{\text{vir}}/dM_{\text{UV}}$	$\log_{10} M_{\text{vir},(M_{\text{UV}}=-19.5)}$
5	-0.321 ± 0.006	11.05 ± 0.02
6	-0.326 ± 0.006	10.88 ± 0.02
7	-0.345 ± 0.005	10.75 ± 0.02
8	-0.347 ± 0.007	10.63 ± 0.02
9	-0.360 ± 0.008	10.52 ± 0.03
10	-0.366 ± 0.009	10.45 ± 0.03

4.5 UV luminosity–halo mass relation

In this section, we discuss the relation between the masses of dark matter haloes (FoF groups) and the UV luminosity of hosted galaxies. If a halo contains more than one galaxy, the luminosity is obtained by summing up all galaxies. Fig. 4.6 shows the halo mass–luminosity relation for all model galaxies at $z \sim 5$ –10. We see that the UV-bright galaxies tend to be located in massive dark matter haloes. We fit the relation between $\log_{10} M_{\text{vir}}$ and UV magnitude for galaxies brighter than −14 mag at each redshift using a linear relation, as shown by the dashed lines in Fig. 4.6:

$$\log_{10} M_{\text{vir}} = \frac{d \log_{10} M_{\text{vir}}}{dM_{\text{UV}}} (M_{\text{UV}} + 19.5) + \log_{10} M_{\text{vir},(M_{\text{UV}}=-19.5)} \quad (4.5)$$

The values of the best-fit slopes and intercepts are shown in Table 4.3 and Fig. 4.7.

We find that the slope of the lines (with median ~ -0.35) slightly steepens towards higher redshift, and the best-fit intercept $\log_{10} M_{\text{vir},(M_{\text{UV}}=-19.5)}$ evolves linearly with redshift. Assuming a constant slope $d \log_{10} M_{\text{vir}}/dM_{\text{UV}} = -0.35$ for the UV-bright galaxies at $z \sim 5$ –10, we find the evolution of luminosity–halo mass relation can be estimated using

$$\log_{10} \left(\frac{M_{\text{vir}}}{10^{10} M_{\odot}} \right) = -0.35(M_{1500} + 19.5) - 0.13(z - 7) + 0.79. \quad (4.6)$$

The scatter of the distribution at high luminosities is much smaller than at lower luminosities, varying from ~ 0.2 dex at $M_{1500} = -20$ to ~ 0.5 dex at $M_{1500} = -16$. Galaxies

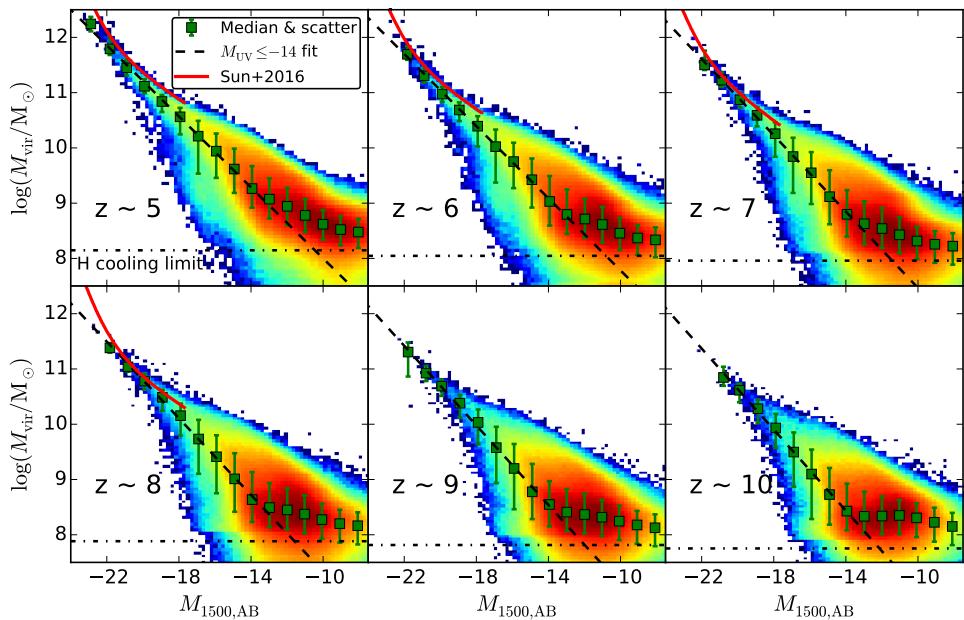


Figure 4.6: Dark matter halo (FoF group) mass as a function of UV luminosity. The colour profile represents the logarithm density of the distribution. The green squares and errorbars show the median and 16th to 84th percentiles in bins which contain at least 5 galaxies. The black dashed lines are the linear fit to the medians at $M_{1500} \lesssim -14$. The dash-dotted horizontal lines show the hydrogen cooling limit at each redshift. The red solid lines show the relation obtained using halo mass abundance matching technique from Sun & Furlanetto (2016) at $z \sim 5-8$.

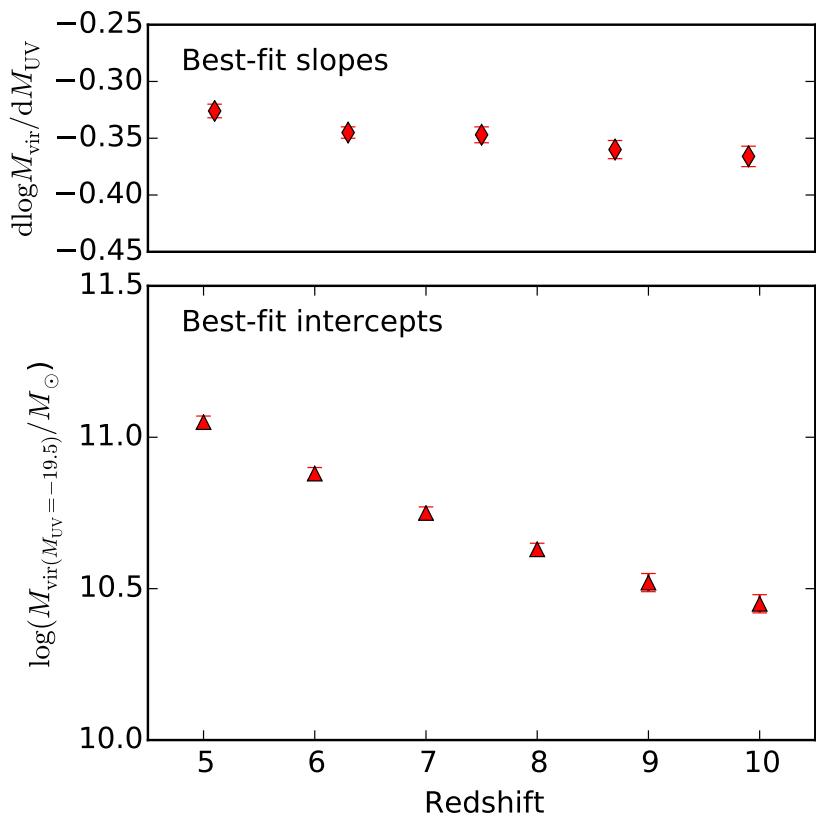


Figure 4.7: The best fitting slopes (upper panel) and the intercepts at $M_{\text{UV}}=-19.5$ (lower panel) of the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation fit to the median trend for galaxies with $M_{\text{UV}} \leq -14$.

with the same UV luminosity tend to be hosted by dark matter haloes with lower masses at higher redshift. This is partly because the stellar populations of galaxies are generally younger at higher redshift. At $z \sim 6$, galaxies with $M_{1500} = -20$ reside in haloes with a mass of $\sim 10^{11.0 \pm 0.1} M_\odot$. A detailed analysis of clustering with MERAXES will be presented in Park et al. (in preparation). However, this value is in agreement with the clustering analysis in Barone-Nugent et al. (2014). Fig. 4.6 shows that at the lowest luminosities, halo masses remain constant at $\sim 10^8 M_\odot$. This is because the model prevents star formation in haloes below the hydrogen-cooling limit, and explains why the LF turns over at low luminosities. We plot the cooling limit for dark matter haloes using the relation between T_{vir} and M_{vir} at high redshifts provided by Barkana & Loeb (2001) and assuming $T_{\text{cool}} = 10^4$ K.

We note that the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ distribution is well described by a linear relation for $-22 < M_{\text{UV}} < -14$ and host halo masses of $M_{\text{vir}} < 10^{12} M_\odot$. We do not have a significant sample of $M_{\text{vir}} > 10^{12} M_\odot$ haloes, due to the simulation volume and so do not have reliable predictions in that mass range. Moreover, for haloes more massive than $10^{12} M_\odot$, AGN feedback, which is not included in the present simulations, may play a significant role in suppressing star formation in galaxies and lead to a steep slope of the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation.

The $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation can also be estimated from observations using the halo abundance matching technique (HAM Vale & Ostriker, 2004; Mason et al., 2015; Sun & Furlanetto, 2016; Mashian et al., 2016). We plot the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation from Sun & Furlanetto (2016) at $z \sim 5-8$ in Fig. 4.6 for comparison with our results. We see that the relations from our model agree with HAM results within statistical uncertainty at $M_{\text{UV}} > -22$. We do see a small difference at $M_{\text{UV}} > -19$ between these two models, which may have arisen from two systematic errors in the HAM methodology: (i) assuming a monotonic relation with an incorrect estimate of scatter between halo mass and galaxy luminosity; (ii) neglecting multiple halo occupation, which leads to the subhalo luminosities paired with central haloes.

We also present the halo mass–luminosity relation for all model galaxies based on *Tiny Tiamat* at $z \sim 5-10$ which is shown in Fig. 4.8. We fit the relation for galaxies brighter than $M_{\text{UV}} \sim -14$ at each redshift using a linear relation, as shown by the dashed lines in Fig. 4.8. The values of best-fit slopes and intercepts are shown in Table 4.4. The figure clearly shows that the model prevents star formation in haloes below the hydrogen-cooling limit. Therefore, the conclusions regarding relations between luminosity and stellar or halo mass are not an artifact of simulation resolution.

4.6 Conclusions

In this chapter, we have studied the relationship between the UV luminosity and intrinsic galaxy properties including SFR and stellar mass, and the relationship between

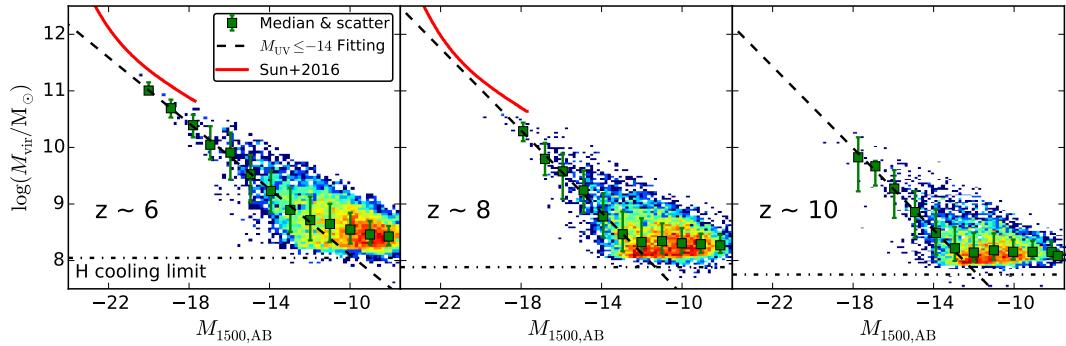


Figure 4.8: Same as Fig. 4.6 but for model galaxies at $z \sim 6, 8$ and 10 based on the *Tiny Tiamat* N-body simulation, which has a smaller volume but higher mass resolution than the *Tiamat* simulation.

Table 4.4: The best fitting slopes and the intercepts of the $\log_{10} M_{\text{vir}} - M_{\text{UV}}$ relation for galaxies with $M_{\text{UV}} \leq -14$ based on the *Tiny Tiamat* N-body simulation.

z	$d \log_{10} M_{\text{vir}} / d M_{\text{UV}}$	$\log_{10}(M_{\text{vir}}, (M_{\text{UV}} = -19.5))$
6	-0.290 ± 0.009	10.86 ± 0.03
7	-0.325 ± 0.018	10.84 ± 0.07
8	-0.365 ± 0.019	10.86 ± 0.07
9	-0.319 ± 0.024	10.46 ± 0.09
10	-0.379 ± 0.006	10.61 ± 0.02

galaxy UV luminosity and the mass of host dark matter halo. We find that the fitted luminosity–SFR, luminosity–stellar mass, and luminosity–halo mass relations are in good agreement with empirical or observed results, and we find that these relations remain valid below current detection limits until at least $M_{\text{UV}} \sim -14$. Other key findings are as follows:

- Model galaxies with $M_{\text{UV}} \lesssim -14$ are distributed around the luminosity–SFR relation from Kennicutt (1998a) and Madau et al. (1998) with a scatter of 0.1–0.3 dex (see Fig. 4.1). However, we find that the conversion between the high- z UV luminosity and SFR functions will be significantly biased by unaccounted-for scatter in the luminosity–SFR distribution for low-SFR galaxies ($\text{SFR} < 0.01 M_{\odot} \text{yr}^{-1}$; see Fig. 4.2).
- Model galaxies with $M_{\text{UV}} \lesssim -14$ have stellar mass–luminosity relations that are consistent with the observed relations from Duncan et al. (2014) at $z \sim 5$ –7. The $\log_{10} M_{*} - M_{\text{UV}}$ relation has a slope of ~ -0.47 which corresponds to $M_{*} \propto L_{\text{UV}}^{1.2}$ (see Fig. 4.3).
- For dark matter haloes with galaxy formation, there is a linear relation between $\log_{10} M_{\text{vir}}$ and M_{UV} with a slope of ~ -0.35 which corresponds to $M_{\text{vir}} \propto L_{\text{UV}}^{0.88}$. The scatter in this relation decreases with increasing luminosity. Galaxies with luminosities of $M_{\text{UV}} = -20$ at $z \sim 6$ are hosted in dark matter haloes of mass $M_{\text{vir}} \sim 10^{11.0 \pm 0.1} M_{\odot}$. This mass decreases linearly towards higher redshift until $z \sim 9$ (see Fig. 4.6).

In this chapter, we provided the redshift dependent values for \mathcal{K}_{uv} in the Madau-Kennecutt relation (see Equation 4.2) for high- z galaxies. We also presented the fitted luminosity–stellar mass relation and luminosity–halo mass relation. These can be used to probe the formation of galaxies as well as dark matter structures from observations.

The predicted linear luminosity–stellar/halo mass relations which break at faint luminosities should be due to the hydrogen cooling limit in low mass haloes. We also study these relations based on *Tiny Tiamat* to investigate the effect of simulation resolution. We find that there are ~ 1 –2 mag differences in the predicted luminosities within the smallest haloes, owing to the additional star formation in the merger triggered starburst phase found in the higher resolution simulation.

Chapter 5

The sizes of galaxies

5.1 Introduction

The evolution of galaxy size during the EoR provides an additional probe for understanding galaxy formation in the early Universe. In the hierarchical structure formation scenario (White & Rees, 1978), dark matter haloes form first, then baryonic gas cools and falls into their potential wells to form galaxies. Within this scheme, Fall & Efstathiou (1980) studied the formation of galaxy discs. In this model, the spin of a rotationally supported galaxy disc originates from the conservation of angular momentum during the collapse of cooling gas. Further analytic modelling by Mo et al. (1998) provided a relation between the disc scale length of a galaxy, R_d , and the virial radius of its dark matter halo, R_{vir} for infinitesimally thin discs with exponential surface density profiles. The disc size can be written as

$$R_d = \frac{\lambda}{\sqrt{2}} \left(\frac{j_d}{m_d} \right) R_{\text{vir}}, \quad (5.1)$$

where m_d and j_d are the fractions of mass and angular momentum in the disc relative to the halo and λ is the spin parameter of the halo, which is a dimensionless measure of the angular momentum of the system.

The virial radius of a dark matter halo scales with redshift and virial velocity, V_{vir} , or virial mass, M_{vir} , as

$$R_{\text{vir}} = \left(\frac{GM_{\text{vir}}}{100H^2(z)} \right)^{1/3} = \frac{V_{\text{vir}}}{10H(z)}, \quad (5.2)$$

where $H(z)$ is the Hubble parameter, and $H(z) \propto (1+z)^{3/2}$ at high redshifts (Carroll et al., 1992). Therefore, from Equation 5.1, the proportionality of R_d with R_{vir} predicts that the sizes of discs scale with redshift as $(1+z)^{-3/2}$ at fixed circular velocity, or $(1+z)^{-1}$ at fixed halo mass.

Observations of Lyman break galaxies (LBGs) show that galaxies are more compact at higher redshift, and that average sizes evolve with redshift as $(1+z)^{-m}$ with $m \sim 1-1.5$ (e.g. Ferguson et al., 2004; Bouwens et al., 2004; Oesch et al., 2010a; Grazian et al., 2012; Ono et al., 2013; Kawamata et al., 2015; Holwerda et al., 2015; Shibuya et al.,

Table 5.1: Observed evolution of galaxy sizes, $R_e \propto (1+z)^m$ from literature, where $L_{z=3}^*$ corresponds to UV magnitude $M_{\text{UV}} = -21.0$.

z	m	Sources
$L = (0.3-1)L_{z=3}^*$		
2–6	1.05 ± 0.21	Bouwens et al. (2004)
2–8	1.12 ± 0.17	Oesch et al. (2010a)
2–12	1.30 ± 0.13	Ono et al. (2013)
2.5–12	1.24 ± 0.10	Kawamata et al. (2015)
0.5–10	1.10 ± 0.06	Shibuya et al. (2015)
5–10	1.32 ± 0.43	Holwerda et al. (2015)
$L = (0.12-0.3)L_{z=3}^*$		
2–8	1.32 ± 0.52	Oesch et al. (2010a)
2–12	1.30 ± 0.13	Ono et al. (2013)
0.5–10	1.22 ± 0.05	Shibuya et al. (2015)
5–10	0.76 ± 0.12	Holwerda et al. (2015)

2015).

Semi-analytic models have had considerable success studying the formation and evolution of galaxies in the past two decades (e.g. White & Frenk, 1991; Kauffmann et al., 1993; Cole et al., 2000; Croton et al., 2006; Bower et al., 2006; Lacey et al., 2011, 2015). Galaxy sizes are important for semi-analytic models since the cold gas is assumed to settle in discs where star formation occurs at a rate depending on the surface density (e.g. Croton et al., 2006). Reproducing the evolution of galaxy sizes in the early and dense Universe is therefore important for semi-analytic models of reionization. On the other hand, feedback mechanisms are already known to play an important role in suppressing star formation in galaxies.

Using the observed size evolution and the LF of galaxies, Wyithe & Loeb (2011) presented a simple model to constrain the feedback mechanism using galaxy sizes:

$$R_e \propto L^{\frac{1}{3(1+a)}} (1+z)^{-m}. \quad (5.3)$$

Here L is the galaxy luminosity, and a and m are free parameters which can be constrained using both the slope of the galaxy LF and galaxy size evolution. Feedback arising from energy release and momentum outflow could affect the luminosity at fixed disc sizes. Based on the observed relation between size, luminosity and redshift, Wyithe & Loeb (2011) ruled out the no-supernova feedback model with high confidence, and suggested a supernova feedback model through the transfer of momentum. Here we improve on this analysis using a more realistic semi-analytic model. Investigation of galaxy sizes using semi-analytic models have previously been made using galaxies in both the local and high redshift Universe (e.g. Cole et al., 2000; González et al., 2009;

Shankar et al., 2010; Xie et al., 2015; Stevens et al., 2016; Tonini et al., 2016). The purpose-designed semi-analytic model MERAXES provides a tool to study galaxy sizes during the EoR.

In this chapter, we run simulations to investigate the size–luminosity relation, the size–stellar mass relation and the redshift evolution of galaxy sizes at $5 < z < 10$. We aim to use the evolution of galaxy sizes to probe the physics of galaxy formation during the EoR. In particular, we study how sensitive galaxy sizes are to feedback, especially from supernovae feedback during the EoR.

This chapter is organized as follows. In Section 5.2 we briefly the discussion of the modelling of galaxy sizes in MERAXES. In Section 5.3 we study the relation between sizes and UV luminosities of galaxies. In Section 5.4 we discuss the probability of resolving galaxies using *HST*, *JWST* and *GMT*. In Section 5.5 we study the size–stellar mass relation of model galaxies. In Section 5.6 we present the redshift evolution of galaxy sizes and compare this with observations. In Section 5.7 we discuss the interpretation of our model sizes in the context of recent high- z observations. In Section 5.8, we present our conclusions.

5.2 Modeling disk sizes

In our semi-analytic model, we adopt the disc scale radius from Mo et al. (1998) as shown in Equation 5.1, and the standard assumption $j_d/m_d = 1$ (Fall & Efstathiou, 1980), for which the specific angular momentum of the material forming the disc is the same as that of the host halo. The spin parameter, λ , is calculated from the N -body simulation using Equation 1.60 (see Angel et al., 2016, for a discussion of spin parameters for haloes in *Tiamat*).

Equation 5.1 was obtained assuming a simple model in which dark matter haloes have singular spherical isothermal density profiles and the gravitational effects of baryonic discs are neglected. It is therefore important to note that inclusion of gravity from the disc may alter the size and rotation curve of galaxies and modify the dark matter concentration in the inner region of the halo. However, Mo et al. (1998) showed that a more realistic model with NFW halo profiles (Navarro et al., 1997) and self-gravitating discs results in only minor modifications to Equation 5.1.

Simulations also show that inclusion of the self-gravity of discs will lead to instabilities of gas and stars, which drives disc material towards the centre of galaxies and results in instability-driven star bursts and bulge growth in galaxies. This will have an impact on the distribution of disc sizes (e.g. Cole et al., 2000; Bower et al., 2006; Stevens et al., 2016; Tonini et al., 2016). Another significant assumption in the model is $j_d/m_d = 1$, since lots of angular momentum in the gas component lost during galaxy assembly would lead to a smaller disc. On the other hand, strong feedback mechanisms which release the energy and angular momentum to the interstellar medium will

suppress the formation of small discs.

To quantify these effects in semi-analytic models, Guo et al. (2016) compared galaxy sizes from semi-analytic models L-GALAXIES and GALFORM at $z < 2$. GALFORM includes the self-gravity of discs while L-GALAXIES ignores it. Guo et al. (2016) showed that self-gravity does not significantly affect the sizes of galaxies with $M_* < 10^{9.5} M_\odot$. However, for galaxies with $M_* > 10^{9.5} M_\odot$, self-gravity of discs in GALFORM reduces galaxy sizes and results in a decreasing size–mass relation. In this work, which considers the small galaxies that drive reionization, we do not have a large number of galaxies with $M_* > 10^{9.5} M_\odot$ at $z > 6$. Thus, we utilize the simple model of Mo et al. (1998) in this study, as has been common in semi-analytic models (e.g. Croton et al., 2006; De Lucia & Blaizot, 2007).

From Equations 5.1 & 1.59, we see that the disc sizes of galaxies are determined by the properties of dark matter haloes. We assume that star formation and feedback processes do not directly modify the disc sizes. On the other hand, the size of the disc does play a fundamental role in the build up of stellar mass. As discussed in Section 2.3.2, The cold gas, which cools from the hot gas reservoir of the host FoF group, is assumed to fall onto the galaxy hosted by the central halo. MERAXES assumes that the cold gas settles in a rotationally supported disc with an exponential surface density profile. Star formation is assumed to occur if the total amount of cold gas in the disc exceeds a critical value, above which gas cannot maintain stability and will start forming stars. The critical mass of the disc is calculated by integrating the critical surface density to the size of the disc through Equation 2.7.

Through this star-formation process, disc size affects a number of galaxy properties, including UV luminosities. The size–luminosity relation therefore becomes an important predictor from galaxy formation models. We note that the star forming process is rather complicated. It is not only determined by the galaxy sizes but also by other effects including cooling, mergers and feedback. To study the role of supernova feedback in the build up of the size-luminosity (stellar mass) relation, we also run a simulation with the supernova feedback switched off. This no-supernova model cannot reproduce the stellar mass function in detail, but is recalibrated to provide the observed stellar mass density at $z = 5$ (Mutch et al., 2016a).

In this chapter, to compare with observations we present the sizes of model galaxies using the physical effective radius (i.e. half-light radius), R_e , within which half of the galaxy’s luminosity originates. Here R_e is estimated using $R_e = 1.678R_d$, where the constant originates from the assumed exponential surface density profile and constant mass-to-light ratio.

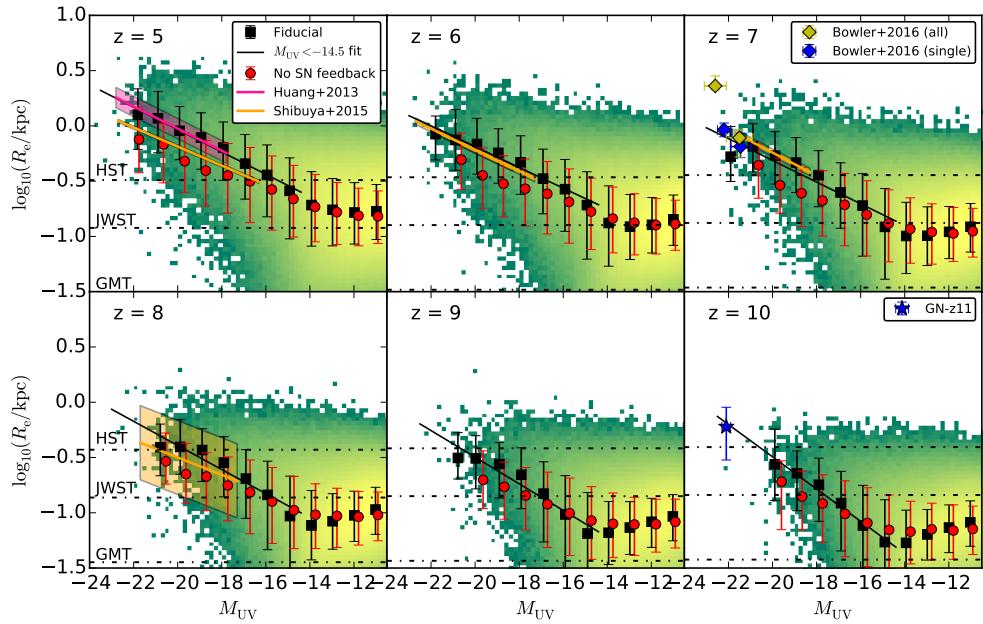


Figure 5.1: Effective radius of galaxies as a function of UV luminosity at $z \sim 5\text{--}10$. The colour profile shows the logarithm density of the distribution. The black squares and error bars represent the median and 16th to 84th percentiles of the R_e distribution in bins which contain at least 10 galaxies. The black solid lines are the linear best-fits for galaxies with $M_{1600} < -14.5$, and are extended to brighter luminosities. The pink and orange lines and associated shaded regions show the observed relations from Huang et al. (2013) and Shibuya et al. (2015). The blue and yellow diamonds show the observations at $z \sim 7$ from Bowler et al. (2016). The blue star shows luminous galaxy GN-z11 found by Oesch et al. (2016). For model comparison, the red circles and error bars show the median and distribution of size–luminosity from the model with supernova feedback turned off. The dash-dotted lines represent the minimum measurable effective radii of *HST*, *JWST* and *GMT*.

5.3 Size–luminosity relation

We first investigate the relationship between the physical size and UV luminosity of model galaxies. Fig. 5.1 shows the relation between the effective radius and UV magnitude M_{UV} for model galaxies at $z \sim 5\text{--}10$, where the UV magnitude M_{UV} is the dust-extincted luminosity at the rest-frame 1600 Å. We see that at $M_{\text{UV}} \lesssim -14$, galaxies with brighter UV luminosity tend to have larger sizes.

We note that the effective radius does not significantly change with luminosity for the galaxies with luminosities $M_{\text{UV}} > -14$. This is because galaxies fainter than $M_{\text{UV}} \sim -14$ are located in the dark matter haloes of the minimum gas cooling mass. This is similar to the turnover at $M_{\text{UV}} \sim -14$ in the relation between UV luminosity and the mass of dark haloes found in Chapter 3. We see that at fixed luminosity, the size of galaxies grows from $z \sim 10\text{--}5$. We discuss the redshift evolution of galaxy sizes further in Section 5.6.

For comparison with our simulations we show the observed $R_{\text{e}}\text{--}M_{\text{UV}}$ relations from Huang et al. (2013) at $z \sim 5$ and Shibuya et al. (2015) at $z \sim 5\text{--}8$, where the latter is calculated by us using the sizes data from Shibuya et al. (2015). Our results are in close agreement with the observations.

Recently, Oesch et al. (2016) found an unexpectedly luminous galaxy (GN-z11) at $z \sim 11$, which has $M_{\text{UV}} = -22.1 \pm 0.2$ and $R_{\text{e}} = 0.6 \pm 0.3$ kpc. In Mutch et al. (2016b) we demonstrated that the properties of GN-z11 are in good agreement with the results of our model in terms of stellar mass, SFR and UV luminosities. We show the measured size of GN-z11 in Fig. 5.1 and we find that it is in agreement with our fitted size–luminosity relation at $z \sim 10$.

The relation between the galaxy size and luminosity is commonly fitted by

$$R_{\text{e}} = R_0 \left(\frac{L_{\text{UV}}}{L_0} \right)^{\beta}, \quad (5.4)$$

where R_0 is the effective radius at L_0 , and β is the slope. We set $L_0 = L_{z=3}^*$ which corresponds to $M_0 = -21$ (Steidel et al., 1999). This equation can be rewritten as

$$\log_{10} R_{\text{e}} = -0.4 \times \beta(M_{\text{UV}} + 21) + \log_{10} R_0. \quad (5.5)$$

We linearly fit the $\log_{10} R_{\text{e}}\text{--}M_{\text{UV}}$ relation for galaxies brighter than $M_{\text{UV}} = -14.5$ at each redshift. The best-fitting values for R_0 and β at $z \sim 5\text{--}10$ are shown in Table 5.2.

We see that the slope of the size-luminosity relation, β , does not significantly change at $z \sim 5\text{--}9$ and has a median value of $\beta \sim 0.25$ for galaxies with UV luminosity brighter than $M_{\text{UV}} \sim -14$. This value agrees with observational studies for both local and high- z galaxies. For example, de Jong & Lacey (2000) found $\beta = 0.253 \pm 0.020$ for local spiral

Table 5.2: The best-fitting parameters R_0 and β (Equation 5.4) for the model galaxies with UV magnitudes $M_{\text{UV}} < -14$ at $z \sim 5\text{--}10$.

z	R_0/kpc	β
5	1.17 ± 0.05	0.25 ± 0.02
6	0.80 ± 0.05	0.23 ± 0.02
7	0.61 ± 0.07	0.25 ± 0.04
8	0.53 ± 0.07	0.28 ± 0.04
9	0.42 ± 0.06	0.30 ± 0.04
10	0.45 ± 0.04	0.36 ± 0.03

galaxies. Shen et al. (2003) derived a slope of $\beta \sim 0.26$ for the late-type galaxies from SDSS. Courteau et al. (2007) obtained $\beta = 0.321 \pm 0.010$ from local field and cluster spiral galaxies. Grazian et al. (2012) found $\beta = 0.3\text{--}0.5$ for LBGs at $z \sim 7$, while Holwerda et al. (2015) derived $\beta = 0.24 \pm 0.06$ using the Grazian et al. (2012) data. In addition, Huang et al. (2013) found $\beta = 0.22$ and 0.25 for the galaxies in GOODS and HUDF fields at $z \sim 4$ and $z \sim 5$, respectively. Finally, Shibuya et al. (2015) investigated the galaxy effective radius from a large *HST* sample and obtained $\beta = 0.27 \pm 0.01$ at $z \sim 0\text{--}8$. They also showed that β does not significantly evolve over this redshift range.

Due to limitations in sample volumes and selection biases, observed values of β often have large uncertainties and vary between studies. For example, observations are generally biased towards galaxies with high surface brightness and are not sensitive to measured properties of fainter, more spatially extended galaxies. Because a model does not suffer from these selection effects and can have a large sample of both bright and faint galaxies, we are able to investigate the true scatter of the size–luminosity relation.

The size-luminosity relation fitted to the model predictions is also consistent with the analytic prediction (Equation 5.3) of Wyithe & Loeb (2011). In that work they considered a supernova feedback model where supernova-driven winds conserve momentum in the interaction with the galactic gas. The model results in a luminosity scaling of $a = 1/3$ which corresponds to $R_e \propto L^{0.25}$. While the model without supernova feedback yields $a = 0$ which corresponds to $R_e \propto L^{0.33}$.

To study the role of supernova feedback on the build up of galaxy sizes, we show the size-luminosity relation for the no-supernova feedback model in Fig. 5.1 (red circles). The size-luminosity relation for the no-supernova feedback model is also flat at $M_{\text{UV}} > -14$. This is because the minimum size is set by the mass scale of efficient cooling in both models. There is no clear difference between the fiducial and no-supernova feedback model at $M_{\text{UV}} > -17$, where the accumulated effect from supernova feedback on star-formation histories is not significant enough to be observed. However, at $M_{\text{UV}} < -17$, the median size of galaxies from the no-supernova feedback model is notably smaller than the fiducial model. In other words, for the same size galaxy, the no-supernova

feedback model results in a much brighter luminosity. We note that removing supernova feedback allows more stars to form, and so the model has been recalibrated to produce the correct stellar mass density at $z = 5$. The luminosity difference is $\sim 2\text{--}3$ mag at $z = 5\text{--}7$, which is larger than the ~ 1 mag difference at $z = 8\text{--}10$. This is also due to the correct galaxy mass only being achieved at $z = 5$ in this recalibrated model.

The different size-luminosity relations from these two models arise because the supernova feedback in the fiducial model suppresses star-formation resulting in a more gradual star-formation history. In contrast, galaxies without supernova feedback have much burstier star-formation histories and contain more young stellar populations which are UV bright. These effects are more significant at lower redshift due to the longer star-formation histories. We also ran a simulation with both supernova and reionization feedback mechanisms switched off. However, we found the result to be almost identical to the no-supernova feedback model, with only a tiny difference at lower redshifts ($z \sim 5\text{--}6$).

5.4 Resolving galaxies with *HST*, *JWST* and *GMT*

The spatial resolution of a telescope with effective diameter D_{tel} is

$$\Delta l = \Delta\theta d_A = \frac{1.22\lambda}{D_{\text{tel}}} d_A, \quad (5.6)$$

where $\Delta\theta$ is the angular resolution determined by the Rayleigh criterion, $\lambda = 1600(1+z)$ Å is the observed wavelength of UV photons and d_A is the angular diameter distance. In Equation 5.6, the observed wavelength is scaled by a factor of $(1+z)$ at fixed intrinsic wavelength, the angular diameter distance decreases at a similar rate at $z \gtrsim 1$. Thus the spatial resolution does not rapidly change with redshift. Galaxy sizes are usually measured through light profile fitting (e.g. Peng et al., 2002). As a result, one can trace the galaxy outskirt light, and obtain an effective radius below the spatial resolution of the telescope. The minimum observable size of a disc depends on many galaxy properties such as the light profile and the image depth. The comparison between the observed R_e and the spatial resolution limits of the *Hubble Space Telescope (HST)* indicates that values of R_e can be measured which are smaller than the resolution limit of the telescope by roughly a factor of ~ 2 (e.g. Ono et al., 2013; Shibuya et al., 2015).

In Fig. 5.1 we show the minimum observable disc size R_{\min} of *HST*, *JWST*, and the *Giant Magellan Telescope (GMT)*, where we adopt the relation $R_{\min} \approx \Delta l/2$ as discussed above. We see that *HST* ($D_{\text{tel}} = 2.4$ m) can resolve the R_{\min} of observed galaxies at $z \sim 5\text{--}7$, and the structures of typical $z > 8$ galaxies can not be resolved. The larger diameter *JWST* ($D_{\text{tel}} = 6.5$ m) will resolve the R_{\min} for galaxies brighter than $M_{\text{UV}} = (-14, -16, -18)$ at $z = (6, 8, 10)$. However, with an exposure time $t_{\text{exp}} = 10^6$ s, *JWST* will observe galaxies to $M_{\text{UV}} = (-15.0, -15.8, -16.3)$ with signal-to-noise ratio

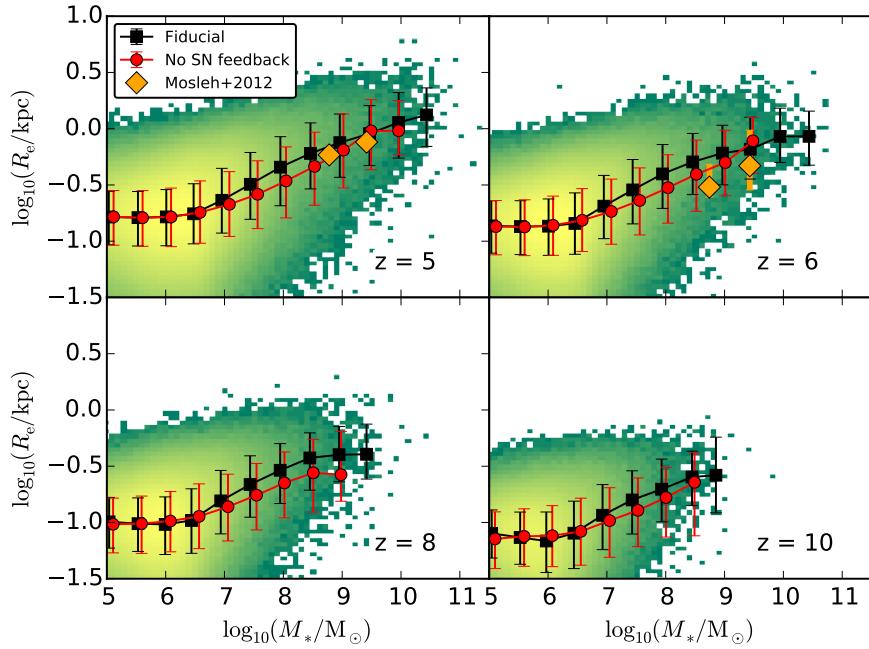


Figure 5.2: Size-mass relation of model galaxies at $z = 5, 6, 8, 10$. The colour profile shows the logarithm density of the distribution. The black squares and red circles show the median relation in bins which contain at least ten galaxies. The error bars represent the median and 16th to 84th percentiles of the intrinsic scatter. The orange diamonds show the observations from Mosleh et al. (2012)

$S/N = 10$ at these redshifts, hence a significant fraction of $z > 8$ galaxies will be still unresolved. Due to the large mirror size, *GMT* ($D_{\text{tel}} = 25 \text{ m}$) will have the ability to resolve all galaxies in haloes above the atomic cooling limit.

5.5 Mass-size relation

Fig. 5.2 shows the relation between the effective radius and stellar mass of galaxies at $z \sim 5, 6, 8$ and 10 for both fiducial and no-supernova feedback models. Observed data from Mosleh et al. (2012) are also shown. The model size-mass relation is in good agreement with these observations. We see that for galaxies with stellar masses above $10^{6.5} M_{\odot}$, more massive galaxies tend to have larger sizes. The galaxies from the fiducial model have larger sizes than the galaxies from no-supernova feedback model at fixed stellar mass. However, the difference in the size-mass relation between the fiducial and no-supernova feedback model is much smaller than in the size-luminosity relation. This is expected because we have tuned both models to produce the galaxy stellar mass density. However, star-formation histories including supernovae lead to less variable UV luminosities resulting in larger difference seen in Fig. 5.1. For galaxies

with $M_* < 10^{6.5} M_\odot$, our two models show similar galaxy sizes due to the inefficient star formation in the minimum cooling mass, as was the case in the size–luminosity relation in Fig. 5.1.

5.6 Redshift evolution of sizes

The redshift evolution of galaxy sizes provides another important measurement in addition to the luminosity dependence (e.g. Ferguson et al., 2004; Bouwens et al., 2004; Oesch et al., 2010a; Grazian et al., 2012; Ono et al., 2013; Kawamata et al., 2015; Holwerda et al., 2015; Shibuya et al., 2015). Fig. 5.3 shows the redshift evolution of the effective radius predicted by our model. To compare with observations of size evolution, galaxies were selected using their luminosity in ranges of $(0.3\text{--}1)L_{z=3}^*$ and $(0.12\text{--}0.3)L_{z=3}^*$. These luminosity ranges correspond to UV magnitudes from -21.0 to -19.7 and from -19.7 to -18.7 , respectively. Both fiducial and no-supernova feedback models are shown in the figure. For comparison the observed galaxy sizes from Bouwens et al. (2004), Oesch et al. (2010a), Ono et al. (2013), Kawamata et al. (2015), Holwerda et al. (2015) and Shibuya et al. (2015) are also shown.

We see that the evolution of galaxy sizes from our fiducial model is in good agreement with observations. However, the galaxy sizes in the no-supernova feedback model are underestimated at each redshift. For example, sizes at fixed luminosity in the no-supernova feedback model are ~ 60 (70) percent of those in the fiducial model at $z \sim 5$ (10). This corresponds to surface brightness densities which are ~ 3 (2) times larger than the fiducial model prediction. These are distinguishable differences. To investigate the influence of parameter calibration in the no-supernova model, we have also run an uncalibrated no-supernova feedback simulation and find a qualitatively similar result. Therefore, we conclude that the galaxy size evolution provides an additional observable for determining the importance of supernova feedback in early galaxy formation.

We fit the model size evolution at $z \sim 5\text{--}10$ using $R_e \propto (1+z)^{-m}$ and find $m = 2.00 \pm 0.07$ with $R_e(z=7) = 0.61 \pm 0.01$ kpc for galaxies with luminosity in the range $(0.3\text{--}1)L_{z=3}^*$ and $m = 2.02 \pm 0.04$ with $R_e(z=7) = 0.50 \pm 0.01$ kpc for galaxies with luminosity in the range $(0.12\text{--}0.3)L_{z=3}^*$. The fitted relations are shown as grey solid lines in Fig. 5.3. We also show $\Delta\chi^2 = 1$ confidence intervals using the observations from Bouwens et al. (2004), Oesch et al. (2010a) and Ono et al. (2013), as well as combined observations from all data shown in Fig. 5.4. Here we only include the observational data at $z > 5$ and do not include more precise measurements at $z < 5$ which could dominate the fit.

We see that the fitted m from our model is comparable to observations. For example, $m = 1.64 \pm 0.30$ and $m = 1.82 \pm 0.51$ are derived using the combined observations shown in Fig. 5.4 with luminosities in the ranges $(0.3\text{--}1)L_{z=3}^*$ and $(0.12\text{--}0.3)L_{z=3}^*$ respectively. We note that the fits from our model as well as $z > 5$ observations give larger values for

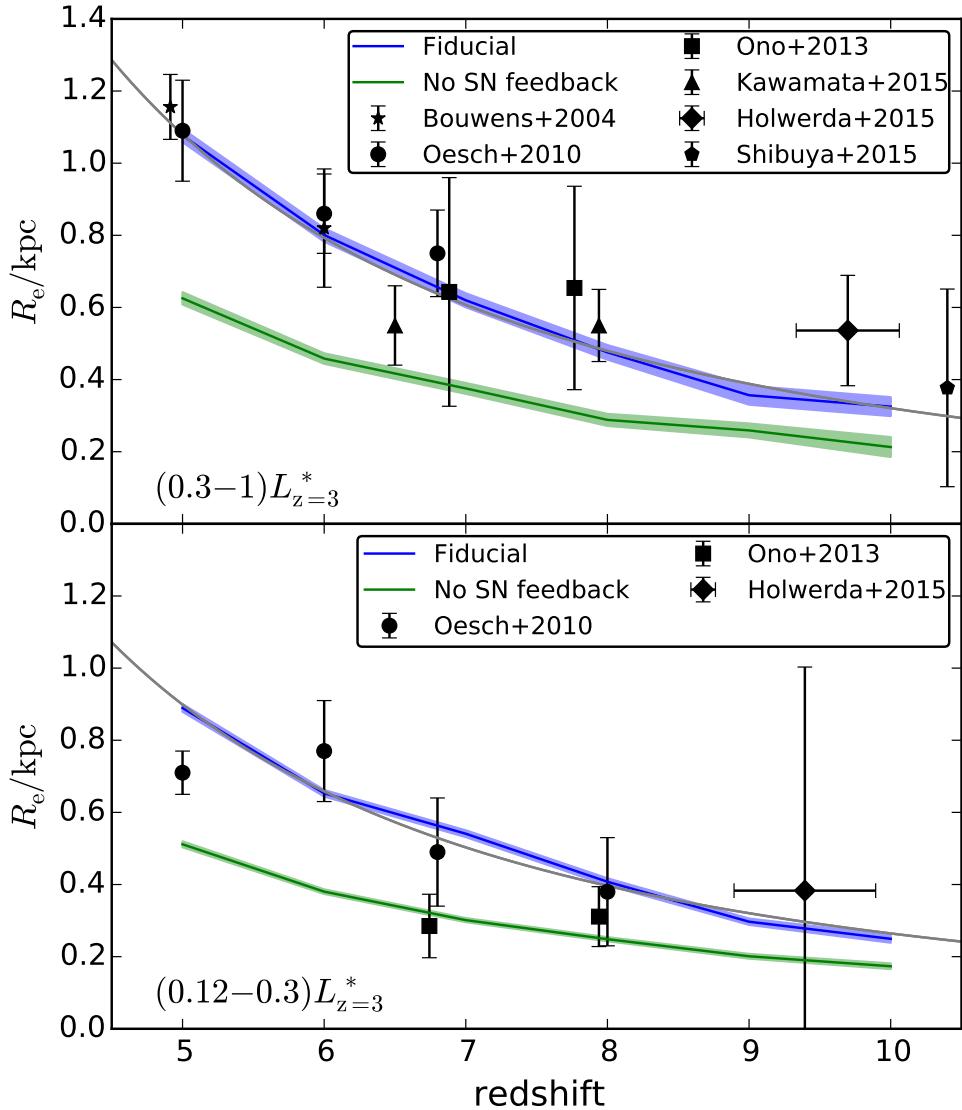


Figure 5.3: The redshift evolution of the mean effective radius for galaxies in the luminosity range $(0.3-1)L_{z=3}^*$ (upper panel) and $(0.12-0.3)L_{z=3}^*$ (lower panel). The blue line shows the mean effective radius from the fiducial model and the green line shows the mean effective radius from the model without supernova feedback. The shaded regions show the associated 1σ uncertainties of the means. The grey solid lines show the power law fit to our model. For comparison, we show the observed mean sizes from Bouwens et al. (2004), Oesch et al. (2010a), Ono et al. (2013), Kawamata et al. (2015), Holwerda et al. (2015) and Shibuya et al. (2015). We see that our fiducial model agrees with observations, while the no-supernova model significantly underestimates the galaxy sizes.

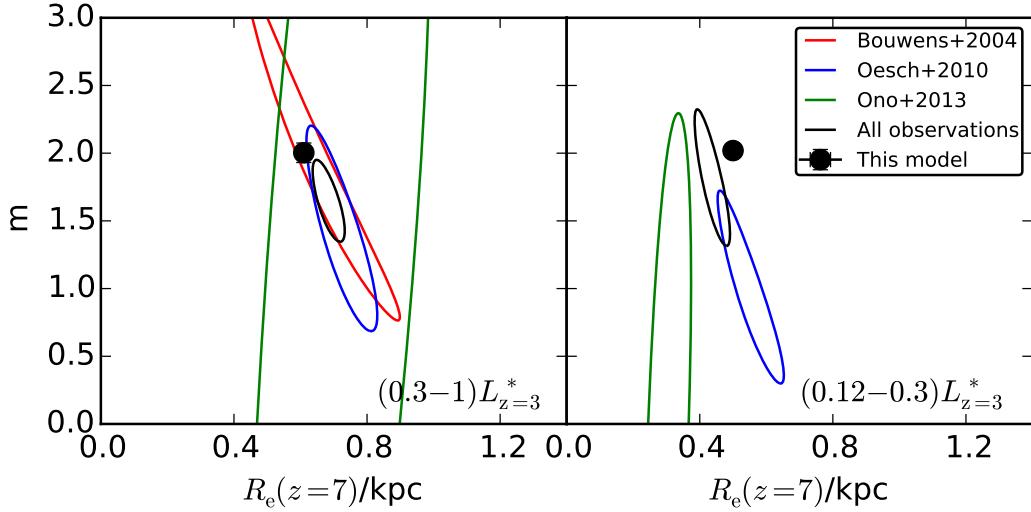


Figure 5.4: Confidence ellipses with $\Delta\chi^2 = 1$, which projects 1σ uncertainties on m and R_e axes. The red, blue and green contours are $z \gtrsim 5$ only observations from Bouwens et al. (2004), Oesch et al. (2010a) and Ono et al. (2013) respectively. The black contours are from all observations shown in Fig. 5.3. Our best-fitting values are shown as black filled circles.

m compared to observations that includes $z < 5$ data as shown in Table 5.1. This may suggest that galaxy sizes undergo faster evolution at $z > 5$ compared to the evolution at lower redshift.

The normalization $R_e(z=7)$ for model galaxies with luminosity in the range $(0.12\text{--}0.3)L_{z=3}^*$ is slightly larger than the combined observations. However, these $z > 5$ observations are also inconsistent with each other due to the large uncertainties from the small sample. We find that $R_e(z=7)$ is in agreement with combined observations with 3σ uncertainty.

5.7 Measuring galaxy sizes

Before concluding this chapter, we discuss the applicability of R_d as a measure of galaxy size. In observations, morphologies of LBGs are often found to be irregular and clumpy, sometimes showing multiple components (e.g. Giavalisco et al., 1996; Ravindranath et al., 2006; Curtis-Lake et al., 2016; Shibuya et al., 2016; Guo et al., 2015; Bowler et al., 2016). This could be due to two different formation mechanisms: (i) galaxy interactions, such as mergers (e.g. Lotz et al., 2006; Overzier et al., 2008); (ii) distributed and clumpy star-formation regions within the same collapsing cloud due to instabilities (e.g. Goldader et al., 2002; Law et al., 2007; Dekel et al., 2009; Oesch et al., 2010a; Jiang et al., 2013; Behrendt et al., 2016).

Morphological studies at very high redshift are more challenging. Shibuya et al.

(2016) investigated the evolution of clumpy galaxies with large *HST* samples and found that the clumpy fraction increases from $z \sim 0$ to $z \sim 1$ but subsequently decreases from $z \sim 1\text{--}3$ to $z \sim 8$. On the other hand, high-resolution cosmological simulations show that galaxies at $z \gtrsim 6$ are dominated by disc morphologies (e.g. Pawlik et al., 2011; Romano-Díaz et al., 2011; Feng et al., 2015). For example, using the large-volume BLUETIDE simulation, Feng et al. (2015) found that at $z = 8\text{--}10$, up to 70 per cent of the galaxy population more massive than $10^{10} M_\odot$ are disc galaxies. Detailed measurement of more compact and clumpy galaxies are limited by the angular resolution of instruments, and the origin of observed clumpy morphologies at high redshift is still under debate.

Bowler et al. (2016) recently published size measurements for a sample of extremely luminous galaxies at $z \sim 7$. Bowler et al. (2016) divided the sample into two groups (single and multi-component) according to their morphologies. The size measurements are shown as the yellow (all galaxies) and blue (single component) diamonds in Fig. 5.1. We see that the size-luminosity relation for the single morphology galaxies is in good agreement with our model while including clumpy morphology galaxies leads to larger sizes. This may suggest that the multi-component galaxies are merging systems (Bowler et al., 2016). However, we are not able to rule out the clumpy-formation scenario due to the simplification of our semi-analytic model. Also, limited by the volume and mass resolution of our *N*-body simulation, the bright multi-component galaxies which undergo mergers will not be resolved by our model.

5.8 Conclusions

We have used the semi-analytic model MERAXES to study the dependence of galaxy size on UV luminosity, stellar mass and redshift at $z \sim 5\text{--}10$. We also studied the effect of supernova feedback on the evolution of galaxy sizes. We show that the rotationally supported disc model generally adopted in semi-analytic models can be used to study the sizes of high- z galaxies. Our primary findings as follows:

- The effective radius scales with UV luminosity as $R_e \propto L^{0.25}$ for galaxies with luminosity $M_{\text{UV}} \lesssim -14$. Galaxies with the same disc size in the no-supernova feedback model have brighter UV magnitudes than in the fiducial model.
- Our fiducial model with strong supernova feedback successfully reproduces the redshift evolution of average galaxy sizes at $z > 5$, which is slightly steeper than $z < 5$ observations. The model with no-supernova feedback produces a significantly smaller radius at fixed luminosity than the fiducial model.
- The recently identified luminous galaxy GN-z11 at $z \sim 11$ (Oesch et al., 2016) lies on our model-fitted size–luminosity relation. The fitted relation is also in agreement with the size measurements of very luminous galaxies containing sin-

gle components and with individual components of luminous multi-component systems at $z \sim 7$ (Bowler et al., 2016).

- A significant fraction of $z > 8$ galaxies will not be resolved by *JWST*. However, *GMT* will have the ability to resolve all galaxies in haloes above the atomic cooling limit.

We conclude that galaxy sizes provide an important additional constraint on galaxy formation physics during reionization, and that current observations of galaxy size and evolution reinforce the importance of supernova feedback. These findings are in agreement with results based on the stellar mass function and luminosity function.

Chapter 6

Summary and future work

6.1 Summary

Modelling galaxy observables at high redshifts is essential for understanding galaxy formation in the early Universe and interpreting high- z observations. The UV luminosities of undetected galaxies at $z > 6$ also play an important role in the ionizing photon budget for reionization. As the title indicates, the aim of thesis is to study observable properties of high- z galaxies using semi-analytic galaxy formation models in order to investigate both of these questions.

As discussed in **Chapter 2**, the dark matter N -body simulation *Tiamat* (Poole et al., 2016; Angel et al., 2016) and the semi-analytic galaxy formation model MERAXES (Mutch et al., 2016a) used in the thesis have been specially designed for modelling galaxy formation during the EoR. *Tiamat* (*Tiny Tiamat*) has a high mass resolution of $2.64 \times 10^6 h^{-1} M_\odot$ ($6.9 \times 10^4 h^{-1} M_\odot$) allowing us to model galaxies within low-mass haloes with virial temperatures close to and below the hydrogen cooling limit ($M_{\text{vir}} \sim 10^8 M_\odot$). The *Tiamat* simulation has a high snapshot cadence of 11 Myr, which resolves the dynamical time of galaxy discs at $z > 6$ and the main-sequence life time of massive stars. Using these high mass and temporal resolutions, MERAXES implements a time resolved supernova feedback scheme, and also includes a temporally and spatially coupled UV background feedback which plays an important role for high- z galaxy formation. In Chapter 2, we have also presented the calculation of SEDs and multi-band photometric luminosities for model galaxies which I have carried out. This is accomplished by integrating the star-formation and assembly histories of galaxies with template SEDs from stellar population synthesis models (Leitherer et al., 1999; Vázquez & Leitherer, 2005; Leitherer et al., 2010). Ly α absorption and dust extinction are also included. The calculated luminosities and colours can then be directly compared with high- z observations.

Since MERAXES was calibrated using the stellar mass function at $z \sim 5\text{--}7$, in **Chapter 3**, we first compared the predicted UV LFs with recent observations (Bouwens et al., 2015b, 2016b). We found that our model successfully reproduces the UV LF for high- z star-forming galaxies at $z \sim 5\text{--}10$. We then discussed how the faint-end slope of the galaxy UV luminosity function behaves below current detection limits, and how much

contribution undetected faint galaxies make to reionization. By studying the UV LFs below current detection limits, we found that the slope of the UV LF remains steep until at least $M_{\text{UV}} \sim -14$ ($M_{\text{UV}} \sim -12$ based on *Tiny Tiamat*) at $z \sim 5\text{--}10$. This will be testable in the future based on observations of faint and lensed galaxies with *JWST*. Assuming a constant escape fraction of ionizing photons, we found that the UV flux is dominated by the UV-bright galaxies ($M_{\text{UV}} \leq -17$) at $z \sim 5$. However, at $z \geq 7$, galaxies with $-17 \lesssim M_{\text{UV}} \lesssim -13$ are the dominant contributors (> 50 per cent) of the total UV flux. Considering a higher escape fraction in faint and high- z galaxies, these undetected sources will emit a significant amount of ionizing photons and may solve the problem of the ionizing photon budget for reionization.

As a direct observable, UV luminosities are commonly used as a probe of intrinsic galaxy properties. In **Chapter 4**, we therefore studied the luminosity dependence of SFR and stellar mass in high- z galaxies. We also studied the relationship between luminosity and halo mass in order to build the connection between the observable of galaxies and the mass of dark matter haloes. We found that model galaxies are distributed around the luminosity–SFR relation from Kennicutt (1998a) and Madau et al. (1998) with a scatter of 0.1–0.3 dex, and that the conversion between the high- z UV LFs and SFR functions will not be significantly biased by unaccounted-for scatter in the luminosity–SFR distribution except for low-SFR galaxies ($\text{SFR} < 0.01 M_{\odot} \text{yr}^{-1}$). We also found that both the luminosity–stellar mass relation and luminosity–halo mass relation can be well fitted using a singular power law as $M_{*} \propto L_{\text{UV}}^{1.2}$ and $M_{\text{vir}} \propto L_{\text{UV}}^{0.88}$, respectively. These relations are valid until as faint as $M_{\text{UV}} \sim -14$ ($M_{\text{UV}} \sim -12$ based on *Tiny Tiamat*), and may break at fainter luminosities due to the hydrogen cooling limit in low-mass haloes.

In addition to the luminosity, galaxy size is another important observable. In MERAXES, the size of a galaxy, calculated using the virial radius and spin parameter of its host halo (Mo et al., 1998) regulates the star formation history through setting the critical mass in the disc. We studied the evolution of galaxy sizes and the size–luminosity relation in **Chapter 5**. We showed that MERAXES successfully reproduces the size–luminosity at $z \sim 5\text{--}10$ and that the recently identified luminous galaxy GN-z11 (Oesch et al., 2016) lies on the fitted size–luminosity relation. These results indicate that the rotationally supported disc model (Mo et al., 1998), widely adopted in semi-analytic models, can be used to model the sizes of galaxies at high redshifts. We studied the redshift evolution at $z > 5$ and found that our model suggests a rapid size evolution compared to $z < 5$ observations. We also investigated the effect of supernova feedback on the evolution of galaxy sizes, and found that a model without supernova feedback produces a significantly smaller radius at fixed luminosity than our fiducial model. This confirms that observations of galaxy sizes will provide an independent constraint on

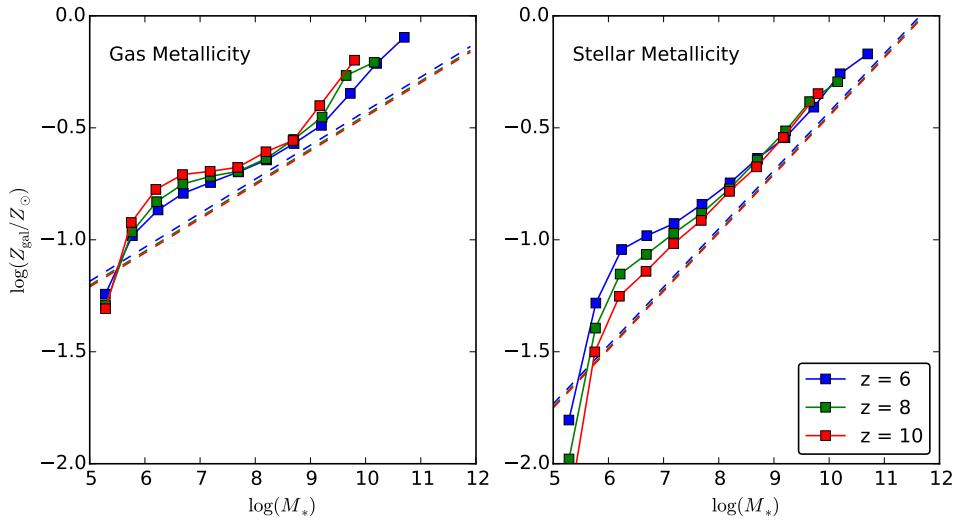


Figure 6.1: The median galaxy mass–metallicity relation (solid squares) from MERAXES at $z \sim 6, 8$ and 10 . The left panel shows the gas-phase metallicity (in the ISM) whilst the right panel shows the stellar metallicity (in stars). Dashed lines are simulation results extrapolated from Ma et al. (2016) which is fitted at $z \sim 0-6$.

galaxy formation including the importance of feedback during the EoR.

Using these findings, our understanding about galaxy observables including UV luminosities and disk sizes can be extrapolated to fainter luminosities below current detection limits and to the early Universe at $z > 10$. Next generation telescopes including *JWST* will have the ability to test our predictions, and our predicted luminosity-dependent relations can be used to study and interpret new observations.

Additionally, mock galaxy catalogs calculated in this work have been used for subsequent studies. For example, Mutch et al. (2016b) studied the origins and fate of the recently discovered $z = 11.1$ luminous galaxy GN-z11 (Oesch et al., 2016). We found that GN-z11 like galaxies in MERAXES possess relatively smooth but extremely rapid mass growth histories, and that they will no longer be the most massive or brightest galaxies by $z = 5$. In another example, Park et al. (in preparation) study galaxy clustering at $z \sim 6-8$ and find that the predicted galaxy bias at fixed luminosity increases with redshift, in agreement with observations.

6.2 Future work

Tiamat and MERAXES provide a framework to study galaxy formation and reionization in the early Universe. The study of galaxy observables will not end with this thesis. We therefore wish to conclude by describing additional possible research directions:

- *The metallicity evolution of high- z galaxies.* Metallicity provides another tool to

study the galaxy formation and evolution. Correlations between stellar mass and metallicity, or luminosity and metallicity are fundamental relations of galaxies. Metallicities of galaxies can be divided into gas-phase metallicities and the stellar metallicities. These reflect the chemical enrichment state across the star-formation history. The metallicity in the ISM is also a parameter for determining the gas cooling rate and the metallicity with stars are born has an impact on the stellar evolution. As the metallicity of galaxies is related to dust abundance and the escape fraction of ionizing photons, it is also important for the study of reionization. In the local Universe, strong monotonic relations between stellar mass/luminosity and gas-phase (Tremonti et al., 2004; Lee et al., 2006), or stellar metallicity (Gallazzi et al., 2005; Kirby et al., 2013) have been found for both star-forming and dwarf galaxies. Similar relations have also been found at redshifts up to $z \sim 3$ (Erb et al., 2006; Maiolino et al., 2008; Mannucci et al., 2009). Metallicity is measured from the emission or absorption lines in a galaxy spectrum, but these spectral lines are extremely weak for distant galaxies. For high- z galaxies, metallicity can be studied using hydrodynamical simulations (e.g. Ma et al., 2016) and semi-analytic models (e.g. De Lucia et al., 2004; Croton et al., 2006; Somerville et al., 2008). In MERAXES, metallicity is the result of the combination of star formation, mass recycling, and gas ejection, accretion, and stripping. We can study the stellar mass–metallicity and the luminosity–metallicity relation for galaxies at $z > 6$ using MERAXES, and the results can be used to compare with relevant high- z observations when they become available. As a preliminary result, Fig. 6.1 shows the relationship between metallicity and stellar mass from MERAXES. We can see that MERAXES reproduces the stellar mass–metallicity relation for high- z galaxies, in agreement with the result from hydrodynamical simulations (Ma et al., 2016).

- *Environment dependence of high- z galaxy properties.* Environment can be characterized via the local mass overdensity with respect to the average density of the Universe (see Chapter 1). On large scales, galaxies and their host haloes are distributed in the cosmic web surrounding large voids (see Fig. 1.4), so that galaxies can reside in various environments. In observations, some galaxy properties show dependencies on the local environment. A well known example is the morphology–density relation, where early-type galaxies are common in high-density regions, while late-type galaxies dominate in low-density regions (Dressler, 1980). Observations also show that galaxies in denser environments are redder, less gas-rich, more massive, more concentrated and have lower specific SFRs compared to galaxies in low-density regions (Kauffmann et al., 2004; Baldry et al., 2006; Brough et al., 2013). These environmental dependences are also found at

higher redshifts (Peng et al., 2010; Strazzullo et al., 2013). However, many of these properties are strongly correlated, so that we need to study which are truly environment dependent.

In hierarchical galaxy formation models, galaxy properties are controlled by the host dark matter halo (e.g. via mass, size and spin), and there is no direct dependence of galaxy formation on the density field surrounding the halo. However, galaxy properties can be modified through the gravitational interactions with nearby galaxies including stripping and strangulation (e.g. van den Bosch et al., 2008; Peng et al., 2015). We can divide the environment into subhalo scales (within the FoF group) and superhalo scales, and study them differently. For example, on superhalo scales, MERAXES includes UVB feedback, where the baryonic infall in a halo can be modified by the local UVB radiation. Mutch et al. (2016a) found that UVB feedback has only a small effect on the stellar mass function at high redshifts. In the future, we can systematically study the environmental dependence of galaxy observables including luminosity, colour, size and other intrinsic properties such as SFR, age, etc.

Bibliography

- Angel P. W., Poole G. B., Ludlow A. D., Duffy A. R., Geil P. M., Mutch S. J., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 459, 2106
- Arrigoni M., Trager S. C., Somerville R. S., Gibson B. K., 2010, MNRAS, 402, 173
- Atek H. et al., 2015, ApJ, 814, 69
- Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
- Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
- Barnes J., Hut P., 1986, Nature, 324, 446
- Barone-Nugent R. L. et al., 2014, ApJ, 793, 17
- Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101
- Behrendt M., Burkert A., Schartmann M., 2016, ApJ, 819, L2
- Benson A. J., Bower R., 2010, MNRAS, 405, 1573
- Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002a, MNRAS, 333, 177
- Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002b, MNRAS, 333, 156
- Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
- Bertschinger E., 1985, ApJS, 58, 39
- Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
- Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
- Boesgaard A. M., Steigman G., 1985, ARA&A, 23, 319
- Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325
- Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
- Bouwens R. J. et al., 2014a, ApJ, 795, 126
- Bouwens R. J., Illingworth G. D., Blakeslee J. P., Broadhurst T. J., Franx M., 2004, ApJ, 611, L1
- Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928
- Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015a, ApJ, 811, 140
- Bouwens R. J. et al., 2012, ApJ, 754, 83
- Bouwens R. J. et al., 2011, ApJ, 737, 90

- Bouwens R. J. et al., 2014b, *ApJ*, 793, 115
- Bouwens R. J. et al., 2010, *ApJ*, 709, L133
- Bouwens R. J. et al., 2015b, *ApJ*, 803, 34
- Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2016a, ArXiv e-prints
- Bouwens R. J. et al., 2016b, *ApJ*, 830, 67
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
- Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2016, ArXiv e-prints
- Brough S. et al., 2013, *MNRAS*, 435, 2903
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, *MNRAS*, 321, 559
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Carroll S. M., Press W. H., Turner E. L., 1992, *ARA&A*, 30, 499
- Cen R., 2003, *ApJ*, 591, 12
- Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, *ApJ*, 437, L9
- Chabrier G., 2003, *PASP*, 115, 763
- Chabrier G., 2005, in *Astrophysics and Space Science Library*, Vol. 327, *The Initial Mass Function 50 Years Later*, Corbelli E., Palla F., Zinnecker H., eds., p. 41
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Coe D. et al., 2013, *ApJ*, 762, 32
- Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, *MNRAS*, 271, 781
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
- Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201
- Courteau S., Dutton A. A., van den Bosch F. C., MacArthur L. A., Dekel A., McIntosh D. H., Dale D. A., 2007, *ApJ*, 671, 203
- Cox T. J., Primack J., Jonsson P., Somerville R. S., 2004, *ApJ*, 607, L87
- Crocce M., Pueblas S., Scoccimarro R., 2006, *MNRAS*, 373, 369
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- Curtis-Lake E. et al., 2016, *MNRAS*, 457, 440
- Daddi E. et al., 2009, *ApJ*, 694, 1517
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- de Blok W. J. G., Walter F., Brinks E., Trachternach C., Oh S.-H., Kennicutt, Jr. R. C., 2008, *AJ*, 136, 2648
- de Jong R. S., Lacey C., 2000, *ApJ*, 545, 781

- De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
- De Lucia G., Kauffmann G., White S. D. M., 2004, MNRAS, 349, 1101
- De Lucia G., Tornatore L., Frenk C. S., Helmi A., Navarro J. F., White S. D. M., 2014, MNRAS, 445, 970
- Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785
- Dressler A., 1980, ApJ, 236, 351
- Duffy A. R., Wyithe J. S. B., Mutch S. J., Poole G. B., 2014, MNRAS, 443, 3435
- Duncan K. et al., 2014, MNRAS, 444, 2960
- Efstathiou G., 1992, MNRAS, 256, 43P
- Efstathiou G., Jones B. J. T., 1979, MNRAS, 186, 133
- Ellis R. S. et al., 2013, ApJ, 763, L7
- Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813
- Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
- Fan X., Carilli C. L., Keating B., 2006a, ARA&A, 44, 415
- Fan X. et al., 2004, AJ, 128, 515
- Fan X. et al., 2006b, AJ, 132, 117
- Feng Y., Di Matteo T., Croft R., Tenneti A., Bird S., Battaglia N., Wilkins S., 2015, ApJ, 808, L17
- Ferguson H. C. et al., 2004, ApJ, 600, L107
- Ferrara A., Loeb A., 2013, MNRAS, 431, 2826
- Field G. B., 1958, Proceedings of the IRE, 46, 240
- Finkelstein S. L. et al., 2012, ApJ, 756, 164
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
- Furlanetto S. R., 2016, in Astrophysics and Space Science Library, Vol. 423, Understanding the Epoch of Cosmic Reionization: Challenges and Progress, Mesinger A., ed., p. 247
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
- Genel S. et al., 2014, MNRAS, 445, 175
- Giavalisco M., Steidel C. C., Macchetto F. D., 1996, ApJ, 470, 189
- Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581
- Goldader J. D., Meurer G., Heckman T. M., Seibert M., Sanders D. B., Calzetti D., Steidel C. C., 2002, ApJ, 568, 651
- González J. E., Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2009, MNRAS, 397, 1254
- González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJ, 735, L34
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, MNRAS, 439, 264

- Grazian A. et al., 2012, *A&A*, 547, A51
- Grazian A. et al., 2015, *A&A*, 575, A96
- Gunn J. E., 1977, *ApJ*, 218, 592
- Gunn J. E., Peterson B. A., 1965, *ApJ*, 142, 1633
- Guo Q. et al., 2016, *MNRAS*, 461, 3457
- Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Guo Y. et al., 2015, *ApJ*, 800, 39
- Guth A. H., 1981, *Phys. Rev. D*, 23, 347
- Haardt F., Madau P., 1996, *ApJ*, 461, 20
- Heinis S. et al., 2014, *MNRAS*, 437, 1268
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, *MNRAS*, 431, 3373
- Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J., 1996, *ApJ*, 457, L51
- Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles
- Holwerda B. W., Bouwens R., Oesch P., Smit R., Illingworth G., Labbe I., 2015, *ApJ*, 808, 6
- Huang K.-H., Ferguson H. C., Ravindranath S., Su J., 2013, *ApJ*, 765, 68
- Hubble E., 1929a, *Proceedings of the National Academy of Science*, 15, 168
- Hubble E. P., 1929b, *ApJ*, 69
- Jeans J. H., 1902, *Philosophical Transactions of the Royal Society of London Series A*, 199, 1
- Jiang L. et al., 2013, *ApJ*, 773, 153
- Kauffmann G., 1996, *MNRAS*, 281, 475
- Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, *MNRAS*, 303, 188
- Kauffmann G., White S. D. M., Guiderdoni B., 1993, *MNRAS*, 264, 201
- Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, *MNRAS*, 353, 713
- Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., 2015, *ApJ*, 804, 103
- Kennicutt, Jr. R. C., 1998a, *ARA&A*, 36, 189
- Kennicutt, Jr. R. C., 1998b, *ApJ*, 498, 541
- Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, *ApJ*, 779, 102
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kroupa P., 2002, *Science*, 295, 82
- Kuhlen M., Faucher-Giguère C.-A., 2012, *MNRAS*, 423, 862

- Lacey C., Cole S., 1993, MNRAS, 262, 627
- Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, MNRAS, 412, 1828
- Lacey C. G. et al., 2015, ArXiv e-prints
- Law D. R., Steidel C. C., Erb D. K., Pettini M., Reddy N. A., Shapley A. E., Adelberger K. L., Simenc D. J., 2007, ApJ, 656, 1
- Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1
- Lee H., Skillman E. D., Cannon J. M., Jackson D. C., Gehrz R. D., Polomski E. F., Woodward C. E., 2006, ApJ, 647, 970
- Lee K.-S. et al., 2012, ApJ, 752, 66
- Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, ApJS, 212, 14
- Leitherer C., Heckman T. M., 1995, ApJS, 96, 9
- Leitherer C., Ortiz Otálvaro P. A., Bresolin F., Kudritzki R.-P., Lo Faro B., Pauldrach A. W. A., Pettini M., Rix S. A., 2010, ApJS, 189, 309
- Leitherer C. et al., 1999, ApJS, 123, 3
- Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
- Lineweaver C. H., Tenorio L., Smoot G. F., Keegstra P., Banday A. J., Lubin P., 1996, ApJ, 470, 38
- Lotz J. M., Madau P., Giavalisco M., Primack J., Ferguson H. C., 2006, ApJ, 636, 592
- Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
- Ma X., Kasen D., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Kereš D., Murray N., 2015, MNRAS, 453, 960
- Madau P., Dickinson M., 2014, ARA&A, 52, 415
- Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648
- Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
- Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, A&A, 496, 57
- Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35
- Maiolino R. et al., 2008, A&A, 488, 463
- Mannucci F. et al., 2009, MNRAS, 398, 1915
- Martin C. L., 1999, ApJ, 513, 156
- Mashian N., Oesch P. A., Loeb A., 2016, MNRAS, 455, 2101
- Mason C. A., Trenti M., Treu T., 2015, ApJ, 813, 21
- McLure R. J. et al., 2013, MNRAS, 432, 2696
- McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, MNRAS, 403, 960

- McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043
- McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, ApJ, 743, 82
- Mellema G., Koopmans L., Shukla H., Datta K. K., Mesinger A., Majumdar S., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 10
- Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
- Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64
- Mihos J. C., Hernquist L., 1994, ApJ, 431, L9
- Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
- Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution
- Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
- Mosleh M. et al., 2012, ApJ, 756, L12
- Muñoz J. A., Loeb A., 2011, ApJ, 729, 99
- Mukhanov V., 2005, Physical Foundations of Cosmology. p. 442
- Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
- Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016a, MNRAS, 462, 250
- Mutch S. J. et al., 2016b, MNRAS, 463, 3556
- Mutch S. J., Poole G. B., Croton D. J., 2013, MNRAS, 428, 2001
- Nagashima M., Lacey C. G., Baugh C. M., Frenk C. S., Cole S., 2005a, MNRAS, 358, 1247
- Nagashima M., Lacey C. G., Okamoto T., Baugh C. M., Frenk C. S., Cole S., 2005b, MNRAS, 363, L31
- Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
- Norman M. L., Reynolds D. R., So G. C., Harkness R. P., Wise J. H., 2015, ApJS, 216, 16
- Ocvirk P. et al., 2016, MNRAS, 463, 1462
- Oesch P. A. et al., 2010a, ApJ, 709, L21
- Oesch P. A. et al., 2010b, ApJ, 709, L16
- Oesch P. A. et al., 2013, ApJ, 773, 75
- Oesch P. A. et al., 2014, ApJ, 786, 108
- Oesch P. A. et al., 2016, ApJ, 819, 129
- Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
- Ono Y. et al., 2013, ApJ, 777, 155
- O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12
- Overzier R. A. et al., 2008, ApJ, 677, 37
- Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544
- Pannella M. et al., 2009, ApJ, 698, L116

- Pawlak A. H., Milosavljević M., Bromm V., 2011, *ApJ*, 731, 54
- Pawlak A. H., Schaye J., Dalla Vecchia C., 2015, *MNRAS*, 451, 1586
- Peebles P. J. E., 1968, *ApJ*, 153, 1
- Peebles P. J. E., 1982, *ApJ*, 263, L1
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Peng Y., Maiolino R., Cochrane R., 2015, *Nature*, 521, 192
- Peng Y.-j. et al., 2010, *ApJ*, 721, 193
- Persic M., Salucci P., Stel F., 1996, *MNRAS*, 281, 27
- Planck Collaboration et al., 2014, *A&A*, 571, A23
- Planck Collaboration et al., 2016, *A&A*, 594, A13
- Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, *MNRAS*, 459, 3025
- Poole G. B. et al., 2015, *MNRAS*, 449, 1454
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 505
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Pritchard J. R., Loeb A., 2012, *Reports on Progress in Physics*, 75, 086901
- Qin Y. et al., 2017, ArXiv e-prints
- Rauch M., 1998, *ARA&A*, 36, 267
- Ravindranath S. et al., 2006, *ApJ*, 652, 963
- Reddy N. A., Steidel C. C., 2009, *ApJ*, 692, 778
- Reddy N. A., Steidel C. C., Fadda D., Yan L., Pettini M., Shapley A. E., Erb D. K., Adelberger K. L., 2006, *ApJ*, 644, 792
- Rees M. J., 1984, *ARA&A*, 22, 471
- Richards G. T. et al., 2006, *AJ*, 131, 2766
- Robertson B. E., Ellis R. S., Dunlop J. S., McLure R. J., Stark D. P., 2010, *Nature*, 468, 49
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, *ApJ*, 802, L19
- Robertson B. E. et al., 2013, *ApJ*, 768, 71
- Romano-Díaz E., Choi J.-H., Shlosman I., Trenti M., 2011, *ApJ*, 738, L19
- Rubin V. C., Ford, Jr. W. K., Thonnard N., 1980, *ApJ*, 238, 471
- Rubin V. C., Ford, Jr. W. K., Thonnard N., Burstein D., 1982, *ApJ*, 261, 439
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schechter P., 1976, *ApJ*, 203, 297
- Schenker M. A. et al., 2013, *ApJ*, 768, 196
- Schiminovich D. et al., 2005, *ApJ*, 619, L47
- Schmidt K. B. et al., 2014, *ApJ*, 786, 57
- Schmidt M., 1959, *ApJ*, 129, 243

- Shankar F., Marulli F., Bernardi M., Boylan-Kolchin M., Dai X., Khochfar S., 2010, MNRAS, 405, 948
- Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
- Shibuya T., Ouchi M., Harikane Y., 2015, ApJS, 219, 15
- Shibuya T., Ouchi M., Kubo M., Harikane Y., 2016, ApJ, 821, 72
- Silk J., 1968, ApJ, 151, 459
- Smit R., Bouwens R. J., Franx M., Illingworth G. D., Labb   I., Oesch P. A., van Dokkum P. G., 2012, ApJ, 756, 14
- So G. C., Norman M. L., Reynolds D. R., Wise J. H., 2014, ApJ, 789, 149
- Sobacchi E., Mesinger A., 2013a, MNRAS, 432, 3340
- Sobacchi E., Mesinger A., 2013b, MNRAS, 432, L51
- Sokasian A., Yoshida N., Abel T., Hernquist L., Springel V., 2004, MNRAS, 350, 47
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
- Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
- Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
- Song M. et al., 2016, ApJ, 825, 5
- Spergel D. N. et al., 2007, ApJS, 170, 377
- Spergel D. N. et al., 2003, ApJS, 148, 175
- Springel V., 2005, MNRAS, 364, 1105
- Springel V. et al., 2005, Nature, 435, 629
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001a, MNRAS, 328, 726
- Springel V., Yoshida N., White S. D. M., 2001b, New A, 6, 79
- Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
- Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17
- Stevens A. R. H., Croton D. J., Mutch S. J., 2016, MNRAS, 461, 859
- Strazzullo V. et al., 2013, ApJ, 772, 118
- Sun G., Furlanetto S. R., 2016, MNRAS, 460, 417
- Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
- Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608
- Tonini C., Mutch S. J., Croton D. J., Wyithe J. S. B., 2016, MNRAS, 459, 4109
- Tremonti C. A. et al., 2004, ApJ, 613, 898
- Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
- Uhlig M., Pfrommer C., Sharma M., Nath B. B., Enslin T. A., Springel V., 2012, MNRAS, 423, 2374
- Vale A., Ostriker J. P., 2004, MNRAS, 353, 189

- van den Bergh S., 2000, The Galaxies of the Local Group. Cambridge
- van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79
- Vázquez G. A., Leitherer C., 2005, ApJ, 621, 695
- Vogelsberger M. et al., 2014, Nature, 509, 177
- Wagoner R. V., 1973, ApJ, 179, 343
- Wagoner R. V., Fowler W. A., Hoyle F., 1967, ApJ, 148, 3
- Walker T. P., Steigman G., Kang H.-S., Schramm D. M., Olive K. A., 1991, ApJ, 376, 51
- Weinberg S., 2008, Cosmology. Oxford University Press
- Weisz D. R., Johnson B. D., Conroy C., 2014, ApJ, 794, L3
- White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
- White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
- Wilkins S. M., Bouwens R. J., Oesch P. A., Labb   I., Sargent M., Caruana J., Wardlow J., Clay S., 2016, MNRAS, 455, 659
- Wilkins S. M., Bunker A. J., Lorenzoni S., Caruana J., 2011, MNRAS, 411, 23
- Wise J. H., Cen R., 2009, ApJ, 693, 984
- Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
- Wouthuysen S. A., 1952, AJ, 57, 31
- Wyithe J. S. B., Loeb A., 2003, ApJ, 586, 693
- Wyithe J. S. B., Loeb A., 2007, MNRAS, 375, 1034
- Wyithe J. S. B., Loeb A., 2008, MNRAS, 383, 606
- Wyithe J. S. B., Loeb A., 2011, MNRAS, 413, L38
- Wyithe S., Geil P., Kim H., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 15
- Xie L., Guo Q., Cooper A. P., Frenk C. S., Li R., Gao L., 2015, MNRAS, 447, 636
- Zeldovich Y. B., Kurt V. G., Syunyaev R. A., 1968, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 55, 278
- Zheng W. et al., 2012, Nature, 489, 406

Appendix A

UV LFs with different metallicities

In this thesis we assumed a constant metallicity of $Z = 0.001$ for stellar population synthesis, which is appropriate during the EoR. To check the effect of assumed metallicity on calculated UV LFs, we present the calculations with different constant metallicities of $Z = 0.001, 0.004$, and 0.008 for calculating SEDs from stellar population synthesis. As shown in Fig. A.1. We find that there is no significant difference between these calculations.

We also calculate UV LFs for galaxies with evolving metallicity which is modelled from MERAXES using metal enrichment scheme (See Chapter 2). For a galaxy, stellar populations are identified using both ages and metallicities, which are obtained by tracing all star-forming, merging and metal enrichment histories. The comparison between the UV LFs with evolving metallicity and with constant metallicity is shown in Fig. A.2. We find that there is no significant difference between the results from these calculations at the relevant redshifts.

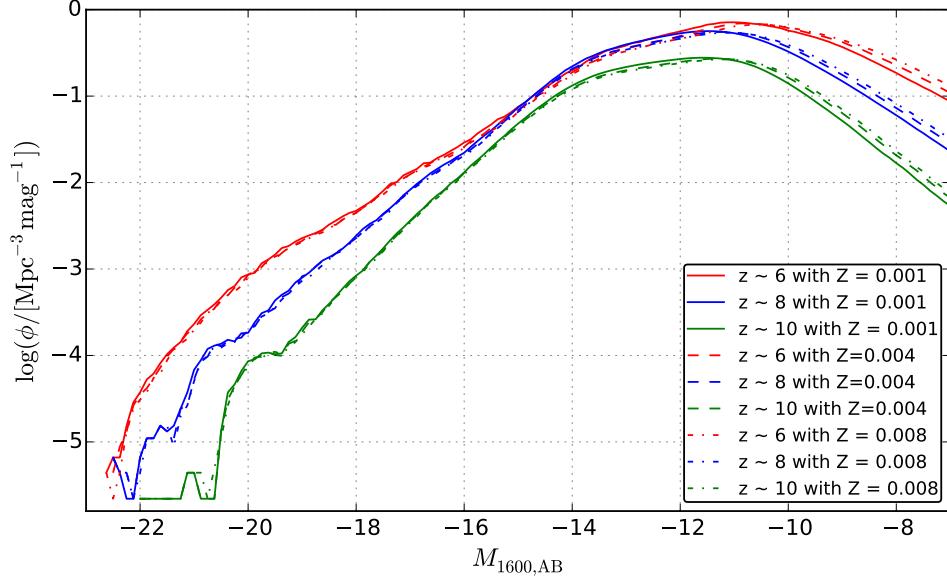


Figure A.1: The calculated UV LFs with constant metallicity of $Z = 0.001, 0.004$ and 0.008 at $z \sim 6, 8$ and 10 . We see that there is no significant difference between the UV LFs with different metallicities.

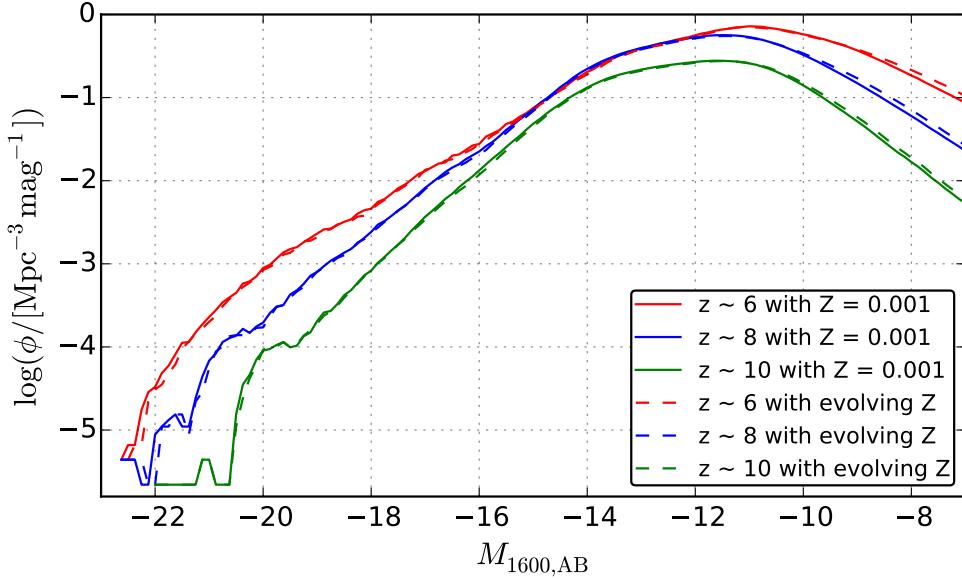


Figure A.2: The calculated UV LFs using evolving metallicity compared with the calculated UV LFs using constant metallicity at $z \sim 6, 8$ and 10 . We find that there is no significant difference between these two calculations at the relevant redshifts.

University Library



MINERVA
ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Liu, Chuanwu

Title:

Galaxy observables during the epoch of reionization from semi-analytic models

Date:

2017

Persistent Link:

<http://hdl.handle.net/11343/191356>

File Description:

Galaxy observables during the epoch of reionization from semi-analytic models

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.