Galaxy Formation Near the Epoch of Reionization

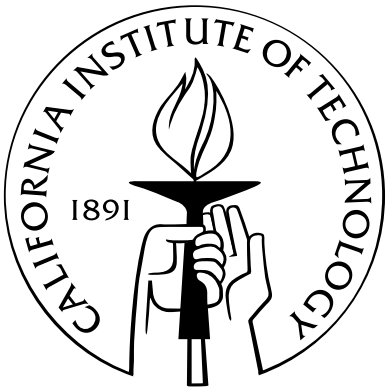
Thesis by

Michael Robert Santos

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2004

(Defended 12 September 2003)

For my family

Acknowledgements

I was very lucky during my years at Caltech to enjoy the support of so many great people. I couldn't have done my Ph.D. without them, and I wouldn't have wanted to.

First I thank my advisors, Marc Kamionkowski and Richard Ellis. No one benefited as I did from their arrivals after my first year. Both were always full advisors to me, even though I wasn't a full student for either. I am eternally grateful for their patience and impatience, and most of all for putting up with me for the better part of four years. I learned so much from them, but what really stands out is what they both taught me about effective communication.

My collaborators were important contributors to this thesis, both through their work on the papers that make up the chapters, and through all that I learned from them over the years. Avi Loeb was friendly and supportive from start to finish of my Ph.D. Volker Bromm was an encouraging and patient guide through the process of writing my first paper. Jean-Paul Kneib and Johan Richard enthusiastically made crucial contributions to the critical-line mapping project.

Before Marc and Richard, my first Caltech advisor was George Djorgovski. Though we didn't end up working together, my interest in galaxy formation started with reading I did under George's supervision and support, which I happily acknowledge.

None of the work I did at Caltech would have been possible without the able and friendly administrative support in Robinson and Bridge, especially from Shirley Hampton and Judy McClain, and computer systems administration by Chris Mach, and Cheryl, Patrick, and Anu.

I am grateful for the help of the observing assistants, instrument specialists, and all of the engineering and administrative staff at Keck Observatory, where the observations of my thesis were obtained, and to all of the scientists and engineers who contributed to the success of the observatory, the telescopes, and their instruments.

Most of the astronomy I know I learned from fellow TAPIRs during informal discussions. Marc's Monday night cosmology dinners taught me almost all of the cosmology I know, and I'm especially thankful for the contributions of Andrew Benson and Peng Oh. I also really enjoyed meeting new astronomers over strawberry lemonade (and high quality free lunch) at the Ath, so I thank the supporters and organizers of Speaker Lunch.

I would have never been at Caltech, or in astronomy (or even in science for that matter), without

the education and encouragement I received during college. At Vassar Debbie Elmegreen, my first astronomy teacher and advisor, and Fred Chromey made me even more excited about becoming an astronomer than I was about becoming a lawyer. My enthusiasm was fueled by my summer research advisors John Salzer and Caryl Gronwall, and Stephen Levine.

And of course I thank NASA, for providing some of the financial support for my graduate studies, and the NSF, for funding my next adventure at Cambridge.

Even with all of the academic support I enjoyed, I wouldn't have had much fun in grad school if it weren't for a bunch of fantastic people in my life, near and far.

I had a great quality of life during (most of) grad school, in large part due to a lot of little pleasures in life around me outside of Caltech. And so I'm thankful for: Sarkis and the staff at A&A Tire and Auto Repair, for working on my ride (picture me rollin'); the In-N-Out staff on Foothill and at the 134 and 2, for late night 3x3's and vanilla shakes; the Blockbuster on Lake, for being near the Cats; Griffith Observatory, for all the prospective visits; Point Dume; House of Pies, for their fresh strawberry pie; the Rialto, for the midnight screenings; Ralph's on Lake, for being open 24 hours; the abundance of movie theaters, for everything from Hurly Burly to Splendor to American Pie; the US women's soccer team, for the World Cup Final in the Rose Bowl; breakfast at Home; Magic Mountain, for letting us ride Viper six consecutive times without getting out, and for my pants-less Batman ride; Green Sand Beach; Burger Continental, for their welcoming weirdness; the WB network, for the glory years, especially Buffy and Dawson's; Robyn, Daisy, Lara, and the other members of the Vassar Club of Southern California steering committee, for fun times; Dr. Dre, Snoop, and Eminem, for much of the soundtrack to my Ph.D.; the waitstaff at Louise's, for their attentiveness; and the people responsible for Google and IMDb, for things you don't want to know about.

There was plenty going on at Caltech as well to keep me entertained, and I thank: the Dance Troupe, especially Sarah for her leadership, for teaching me something completely new; the other members of the Everhart Lecture Series committee, for the nights learning about Caltech research over Koo Koo Roo; Lee Armus and the rest of Ether Binge, for the chance to slide into third (and a championship); GSC Soccer and everyone who played (and especially those who ran lines) for the Cataclysmic Variables over the years, for helping astronomy represent on the field (and two championships); the PJ Young fund and the participants of the ski trips, especially the organizers, Brian and Josh, for giving me the opportunity to learn to ski, and to soar through the air...and to crash into a snow gun.

I was fortunate to keep in touch with a lot of old friends during grad school, and keeping up with their trials and triumphs always helped keep my life in perspective. They are responsible for far too

many fantastic experiences to recount here, but the least I can do is record some of the most salient ones. And so I thank: Debbie Missal, Megan, and Heather, for two great Thanksgivings at times when I really needed something to cheer me up, for which I'll always give thanks; Kevin, for the wild ride through the Scottish Highlands; Mikey, for frisbee in the streets Beverly Hills; Srinu, for our descent off Point Dume; Marissa, for sharing the most dramatic life this side of the Creek; Dave, for scrambling in Yosemite; Brian, for posing on that car in Vegas; Mark, for your perseverance in life and love and email; and Hunter, for not flinching first; Jessica, for pants; Hayden, for the thumbs up at the Peep Show; Austin, for hanging out together 10 times in four cities; Matt, for soldiering through the snowboarding trips; Steuerman, for random sports and crazy stories; Michalak, for the chance to mountainboard (and the helmet that saved my scalp); and Erich, for countless weekends that made me feel like I was in college again, and more fun than is legal in the OC.

I was especially lucky to live with two of the best roommates ever. First at 1020 Del Mar #203, I was randomly matched with Adam, quite possibly the nicest guy ever. He taught me basically everything I know about aeronautics and Islam, and anytime we hung out I always learned something new and interesting. And I got him hooked on Dawson's. Then Erin and I moved into 4444 Avocado #103, thanks to her hard work finding an affordable apartment in hip and trendy Los Feliz (and very near House of Pies). Even though we saw each other surprisingly little for roommates, living with Erin gave me a wacky and wonderful world away from work.

Grad school would have been a shadow of the rich experience it was without the friends I had there. I'm thankful for: Sam, Cindy, and Andrew, for spreading the Buffy love; anyone who ever came out for a prospective event I organized; having people to talk with at work after 2am; hanging around donuts after 4pm; Journal Beers that carried on into dinner; lunches at Chandler and Avery; birthday dinners (those who kept it real); Yama, for his Q's; anyone who brought donuts from Donut Man; FIFA in the Ram Van; watching Micol play Dance Dance Revolution; Caltech quizbowls; $6 Dodgers tickets; Speaker Lunch eating contests; and so many more little moments that meant a lot to me.

I was also lucky to have great officemates, which is important to someone as talkative as I am. Yoram was much more fun than he gives himself credit for, and I learned a lot from him. And it was a distinct pleasure to share an office and foodstuffs with Alison (actually, I did most of the sharing on both accounts). She is as much fun as she gives herself credit for; there was never a dull moment with Alison around. And just for fun: Jigglypuff.

My first officemates were my two surviving classmates, Matt and Bryan. There is no way to summarize variety of interesting things they taught me and the fun times we had. From extremely low potassium and the flaming bull, to photography, to finding the Golden Square, to diesel engines, to standing behind the lamp, to passing the qual, to xpilot, to Pie, to racing at Uncle Kal's and Speedzone, I had so many great times with them in and out of 03 Robinson. Adversity bonds people

together, and I was lucky to be able to share the struggles of first year with Matt, Bryan, and Gwen, who put the rest of us to shame in classes and then left for greener pastures. Matt had the good taste to bring Gina into the fold, and Megan has always been a part of our class. So mad props to Mix and B-Dogg for keepin' it real.

I can't fully express everything Alice's friendship has meant to me during my time in grad school. She is a kindred spirit, and has been my closest friend and confidant through the good and bad times. She is inextricably linked with my happy memories of Caltech and points spread around the globe, and I will always be thankful to have her as a friend.

And finally, my biggest thanks goes to Mom, Dad, and Scotty. The love and support of my family underlies anything I have ever accomplished.

Abstract

I present five explorations of the formation of early galaxies near the epoch of reionization: observable phenomena that demonstrate the interaction between galaxy formation and the intergalactic medium.

I calculated the contribution of the first stars (Pop III) to the cosmic infrared background (CIRB). Pop III stars produced the observed unexplained component of the NIR CIRB if they were very massive and formed efficiently over a redshift range $25 \gtrsim z \gtrsim 7$. A small escape fraction of ionizing photons from the nebular gas immediately surrounding the Pop III stars enhances the Pop III CIRB signal.

I calculated the effect of radiative transfer through the intergalactic medium (IGM) on the observed properties of Lyman α emission lines from $z \gtrsim 6$ galaxies. The detection of a Lyman α emitting galaxy at $z = 6.5$ resulted in claims that the universe was mostly ionized there. With existing information about the source galaxy it is almost impossible to deduce the ionization state of the $z = 6.5$ IGM, especially if there were galactic-scale super-winds.

I present a method to infer the stellar population responsible for completing reionization. We simulated helium and hydrogen absorption along lines-of-sight toward high-redshift QSOs assuming either Pop II or Pop III stars dominated the UV background at $z = 5$. The final SDSS quasar sample may contain a quasar/line-of-sight combination that constrains the nature of the dominant ionizing sources only 300 Myr after the end of reionization.

I present an observational survey of low-luminosity Lyman α emitting galaxies at $4.5 < z < 6.7$. This unprecedented survey utilized strong lensing by rich clusters of galaxies at intermediate redshifts to discover six likely Lyman α emitting galaxies. The $z \sim 5$ Lyman α luminosity function may be flatter than the mass function of dark matter halos, a sign of suppression of the star-formation efficiency in low-mass halos. However, our data by themselves are consistent with the slope of the halo mass function.

I modeled the formation of metal-poor globular cluster systems (GCSs), invoking reionization to explain the bimodality observed in GCS metallicity distributions. The model reproduces the mass and spatial extent of the Milky Way metal-poor GCS, and also the increasing number of metal-poor globular clusters per unit galaxy mass for galaxies in richer environments, as observed.

Contents

Acknowledgements **iii**

Abstract **vii**

1 Introduction **1**

1.1 Introduction 1

1.2 Chapter summaries 5

2 The Contribution of the First Stars to the Cosmic Infrared Background[1] **11**

2.1 Abstract 11

2.2 Introduction 12

2.3 Star-formation rate 13

2.3.1 Overview of the star-formation model 13

2.3.2 Specific star-formation models 14

2.3.2.1 Ongoing star formation 14

2.3.2.2 Single-burst star formation 15

2.3.3 Critical halo mass for star formation 16

2.3.4 Limits on the abundance of Pop III stars 17

2.3.4.1 Black holes 17

2.3.4.2 Metal enrichment 18

2.3.5 Star-formation model results 19

2.4 Spectra 20

2.4.1 Input stellar spectrum 20

2.4.2 No escape of ionizing radiation into the IGM 21

2.4.2.1 Properties of a nebula 21

2.4.2.2 Resulting spectrum 23

2.4.3 Complete escape of ionizing radiation into the IGM 28

[1] This chapter appeared as Santos, Bromm, and Kamionkowski (2002), and is reproduced here by permission of the copyright holder, the Royal Astronomical Society.

2.4.3.1 Properties of the IGM . . . 28
2.4.3.2 Resulting spectrum . . . 30
2.5 Cosmic infrared background . . . 35
2.5.1 Model results . . . 35
2.5.2 Observational data . . . 36
2.6 Discussion . . . 37

3 Probing Reionization with Lyman α Emission Lines[2] 41
3.1 Abstract . . . 41
3.2 Introduction . . . 42
3.3 Lyman α scattering . . . 44
3.4 Lyman α line profiles . . . 46
3.4.1 Intrinsic Lyman α emission line . . . 46
3.4.2 Observed Lyman α emission line . . . 48
3.4.2.1 IGM model . . . 48
3.5 Lyman α line results . . . 52
3.5.1 Fiducial galaxy/halo model . . . 52
3.5.2 IGM model dependence . . . 54
3.5.3 Galaxy/halo parameter dependence . . . 61
3.5.4 IGM ionization dependence . . . 66
3.5.5 Redshift dependence . . . 68
3.6 Galactic winds . . . 71
3.7 Discussion . . . 73
3.7.1 Lyman α line fluxes . . . 73
3.7.2 Observed $z = 6.5$ Lyman α emitters . . . 74
3.7.3 Model assumptions . . . 79
3.8 Summary . . . 81

4 A Method to Infer the Stellar Population that Dominated the UV Background at the End of Reionization[3] 82
4.1 Abstract . . . 82
4.2 Introduction . . . 83
4.3 Absorption spectroscopy . . . 84
4.4 Ionizing background . . . 88
4.5 IGM absorption model . . . 90

[2]This chapter appeared as Santos (2004a), and is reproduced here by permission of the copyright holder, the Royal Astronomical Society.

[3]This chapter was submitted to Monthly Notices of the Royal Astronomical Society in 2003 as Santos and Loeb.

4.6 Model quasar intrinsic spectrum . . . 94
4.7 Lines-of-sight toward SDSS quasars . . . 95
4.7.1 Absorber Monte Carlo simulations . . . 96
4.7.2 Quasar absorption line spectrum simulations . . . 97
4.7.3 Quasar absorption line spectrum analysis . . . 100
4.8 Observational quasar selection techniques . . . 105
4.9 Summary . . . 109

5 The Abundance of Low-luminosity Lyman α Emitters at High Redshift[4] 111
5.1 Abstract . . . 111
5.2 Introduction . . . 112
5.3 Lyman α Surveys . . . 114
5.4 Searches for distant low-luminosity galaxies . . . 116
5.5 Observations . . . 120
5.5.1 Survey stategy . . . 120
5.5.2 Survey parameters . . . 122
5.5.3 Candidate selection and catalog . . . 124
5.5.4 Intermediate-resolution spectroscopy . . . 128
5.6 Detections . . . 128
5.6.1 $z > 4.5$ . . . 128
5.6.2 $z < 4.5$ . . . 131
5.7 Survey analysis . . . 131
5.7.1 Survey volume . . . 131
5.7.1.1 Magnification, $\mathcal{M}(\mathbf{\Omega}, z)$ . . . 133
5.7.1.2 Limiting Lyman α line flux, $f_{\rm lim}(z)$ . . . 135
5.7.1.3 Slit transmission, $T(\mathbf{\Omega})$ . . . 137
5.7.2 Volume as a function of source redshift and luminosity . . . 137
5.7.3 Number density . . . 139
5.8 Comparison with other observations and with theory . . . 142
5.8.1 Comparison with other observations . . . 142
5.8.2 Comparison with theoretical models . . . 146
5.9 Summary . . . 150

[4]This chapter appeared as Santos et al. (2004), and is reproduced here by permission of the copyright holder, the American Astronomical Society.

6 The Formation of Metal-Poor Globular Cluster Systems[5] 152

6.1 Abstract . . . 152

6.2 Introduction . . . 153

6.3 Model . . . 157

6.3.1 The importance of reionization . . . 157

6.3.2 Model parameters . . . 161

6.4 Computational methods . . . 162

6.4.1 Initial conditions . . . 162

6.4.2 N-body evolution . . . 164

6.5 Milky Way results . . . 166

6.6 Environment . . . 170

6.6.1 Schematic argument for increasing specific frequency with increasing richness of galactic environment . . . 170

6.6.2 Semi-quantitative analysis . . . 172

6.7 Summary . . . 173

[5]The work described in this chapter will be submitted in 2004 to MNRAS for publication.

List of Figures

2.1 The star-formation rate (SFR) for the ongoing and single-burst models. 19

2.2 The matter density processed through Pop III stars in units of the critical density, as a function of redshift. 20

2.3 Spectrum of a $z = 15$ Pop III star plus nebular emission, for the $f_{\rm esc} = 0$ case of section 2.4.2. 23

2.4 Properties of the ionized IGM near a Pop III star. 28

2.5 Spectrum of a $z = 15$ Pop III star plus emission from the IGM, for the $f_{\rm esc} = 1$ case of section 2.4.3. 31

2.6 The cosmic infrared background from Pop III stars. 34

3.1 Cross section for Lyman α absorption, as a function of wavelength difference from resonance in the frame of the absorbing atom. 46

3.2 IGM density as a function of radius from the fiducial galaxy. 55

3.3 Velocity of IGM gas as a function of radius from the fiducial galaxy. 56

3.4 IGM density as a function of velocity in the frame of the fiducial galaxy. 56

3.5 Schematic diagram of location and velocity of the IGM around the fiducial galaxy. . . 57

3.6 Optical depth due to IGM scattering as a function of observed wavelength. 58

3.7 Density and ionization state of the IGM as a function of radius and velocity. 59

3.8 Lyman α line profiles. 60

3.9 As in Fig. 3.7. 62

3.10 As in Fig. 3.8. 63

3.11 As in Fig. 3.7. 63

3.12 As in Fig. 3.8. 64

3.13 As in Fig. 3.8. 65

3.14 As in Fig. 3.7. 65

3.15 As in Fig. 3.8. 66

3.16 As in Fig. 3.8. 67

3.17 As in Fig. 3.8. 69

3.18 As in Fig. 3.7. 70

3.19 Lyman α line profiles. 70
3.20 As in Fig. 3.8. 72
3.21 Intrinsic and observed Lyman α line flux as a function of model parameters. 74
3.22 Intrinsic and observed Lyman α line flux as a function of model parameters, including possible effects due to galactic winds. 75
3.23 Ratio of observed to intrinsic Lyman α line flux as a function of model parameters, for both no wind and wind models. 76
3.24 Intrinsic and observed Lyman α line flux as a function of ionizing photon escape fraction. 78
3.25 Ratio of observed to intrinsic Lyman α line flux as a function of $x_{\rm HI}^{\rm IGM}$. 79

4.1 Simulated absorption spectra for $z = 5$ quasars with $\mathrm{AB}_{1450} = 20$. 86
4.2 The ratio, $\eta_{\rm thin}$, of the column density of neutral helium to the column density of neutral hydrogen, as a function of the spectral slope of the ionizing background, $\alpha_{\rm b}$. . 91
4.3 Cumulative histogram of the He I-band transmission along quasar lines-of-sight, as a function of redshift and absorber model. 98
4.4 A simulated absorption spectrum for a $z = 5$ quasar with $\mathrm{AB}_{1450} = 20$. 101
4.5 An example of the pixel-by-pixel cross-correlation of the absorption in the He I band and the H I Lyα band. 102
4.6 Contour plot of the ensemble continuum-normalized fluxes of pairs of corresponding He I-band and Lyα-band pixels. 103
4.7 Histograms of R for each of our two input values of η, 0.06 and 0.20, assuming observation of a quasar along a sightline with $T(\mathrm{He\,I}) = 0.03$. 104
4.8 Same as Fig. 4.7 but for an LOS with $T(\mathrm{He\,I}) = 0.1$. 104
4.9 Normalized cumulative histograms of R from Fig. 4.7. 105
4.10 Normalized cumulative histograms of R from Fig. 4.8. 106
4.11 Plot of $T(\mathrm{He\,I})$ versus $T(u')$. 107
4.12 Plot of $T(\mathrm{He\,I})$ versus g' of quasars. 108

5.1 The star-formation rate function based on a dark matter halo model. 119
5.2 Survey clusters, with survey area and lensing critical curves. 125
5.3 Survey clusters, with survey area and lensing critical curves. 126
5.4 Two-dimensional spectra of the three confirmed $z > 4.5$ galaxies detected in the survey. 128
5.5 Postage stamps of spectra and *HST* images. 129
5.6 Postage stamps of spectra and *HST* images. 132
5.7 Magnification due to the cluster Abell 2218 of background sources, as a function of position and redshift. 134
5.8 Cumulative histogram of the magnification factor over the entire survey area. 135

5.9 Limiting Lyman α line flux as a function of wavelength. 136
5.10 The redshift distribution of our survey volume. 138
5.11 Survey volume sensitive to sources with Lyman α line luminosity L. 139
5.12 Number density of sources brighter than luminosity L, for the three confirmed $z > 4.5$ sources. 140
5.13 Number density of sources brighter than luminosity L, for the three confirmed plus three likely $z > 4.5$ sources. 141
5.14 Number density of $z \sim 5$ sources brighter than L, for Lyman α surveys and Lyman-break galaxy surveys. 144
5.15 Number density of $z \sim 5$ sources brighter than L, for Lyman α surveys and $z \sim 5$ Lyman-break galaxy surveys. 145

6.1 Baryonic structure formation and reionization. 160
6.2 Initial positions of sites of metal-poor globular cluster formation. 165
6.3 Final positions of sites of metal-poor globular cluster formation. 168
6.4 Radial distribution of observed and simulated Milky Way globular clusters, and Milky Way dark matter halo. 169
6.5 Schematic illustration of dependence of specific frequency on galactic environment. . . 172
6.6 The specific frequency of metal-poor globular clusters. 174

List of Tables

5.1 Clusters surveyed 122
5.2 LRIS survey observations 123
5.3 Lyman α emission line candidates 127
5.4 Galaxy surveys at $z \sim 5$ 143

Chapter 1

Introduction

1.1 Introduction

The Milky Way is a robust middle-aged galaxy with a total mass of 10^{12} $M_{\odot}$. We have a very similar neighbor (the Andromeda galaxy), and several smaller companions. Our galaxy quietly and confidently goes about its business of forming stars, such as the Sun, as it has at a relatively constant rate over the last 10 billion years—most of the age of the universe—with only occasional excitement when one of our little friends drops in.

The early universe was a completely different story. After a hot but smooth beginning in the Big Bang, the universe cooled off as it expanded, and matter began to clump together under the influence of gravity. Structures resembling pancakes collapsed along one axis to form filaments; the intersections of these formed roughly spherical halos, which grew in mass as matter flowed along the filaments and poured into the halos. Less than 500 million years after the universe began, the first star in the universe was born, inside one of the halos.

The billion years after the first star was born are even more crucial to our understanding of why the universe now looks the way it does. My thesis research focused on this period, which witnessed the formation of the first clusters of many thousands of stars and the first small galaxies. These young galaxies polluted their pristine hydrogen and helium surroundings with carbon, oxygen, and other heavy elements. New elements opened up new pathways to cool gas, accelerating star formation. Pockets of intense radiation fields heated gas and blew intergalactic winds. Some small

proto-galaxies were completely destroyed by the combined impact of radiation and winds from their neighbors. These turbulent times culminated when the bubbles of intense radiation percolated throughout the universe, ionizing all of the hydrogen between the galaxies, and greatly homogenizing the radiation field. At that instant the oldest stars in the Milky Way had already formed, though another billion years passed before the Milky Way itself assembled.

My thesis presents five explorations of the formation of early galaxies near the epoch of reionization. The goal is to connect current and future observations, both of the local and high-redshift universe, to the conditions and astrophysics underlying the first epoch of galaxy formation. I will next introduce the themes and methods common to the whole thesis, then in Section 1.2 summarize each of the five projects individually.

The best current census of the contents of our universe indicates that about 80% of the mass is non-baryonic dark matter (Spergel et al., 2003). Dark matter is believed to interact almost exclusively through gravity with itself and baryonic matter. As a consequence, it is relatively simple to calculate the density evolution of dark matter from some initial conditions to a later cosmic time, both analytically (e.g., Padmanabhan, 1993; Peebles, 1993, for pedagogical reviews) and numerically (see Bertschinger, 1998, for a recent review). Moreover, the initial conditions of the dark matter are easy to calculate under the paradigm that inflation nearly completely smoothed out the early universe, leaving behind only tiny, scale-free (or nearly so), fluctuations in the matter distribution. On scales of interest the density perturbations are well approximated by linear perturbation theory until $z \sim 100$ or later (Zeldovich, 1970). Overdense perturbations are unstable to collapse, as dark matter is nearly pressureless, and the end result of the most overdense perturbations is collapse to a nearly spherical halo of dark matter (see Cooray & Sheth, 2002, for a review). Details of the evolution of dark matter are still an active research topic (e.g., Power et al., 2003), but the qualitative features and quantitative calculation techniques are generally agreed upon.

The baryonic universe is a more complex place, though, because forces beyond gravity are relevant. However, in many situations of interest, gravitational forces from a dark matter halo dominate the other forces felt by the associated baryons on large scales. Thus, as a useful first approximation

to the fundamentally baryonic process of galaxy formation, it has proved useful to associate galaxies with dark matter halos (see White, 1999, for a review), partially justified by the dark matter content inferred in local galaxies and clusters of galaxies. In most of the volume of the universe, though, thermal pressure of the baryonic gas is the most important force. Consequently, the distribution of material in the inter-galactic medium (IGM) is a more challenging calculation than the distribution of galaxies. Though there are many outstanding issues in galaxy formation, much current research is devoted to relating galaxy formation to the IGM.

The reionization of intergalactic hydrogen is an important example of the coupling between galaxy formation and the IGM. Reionization was dominated by photoionization of intergalactic gas, most likely from sources associated with early galaxies (see Barkana & Loeb, 2001; Loeb & Barkana, 2001, for reviews). So an understanding of early galaxy formation is key to predicting the reionization epoch and process. The interplay between galaxies and the IGM persisted even after the universe was finally and nearly completely ionized at $z \sim 6$ (Fan et al., 2002, and references therein). The sudden temperature rise associated with photoionization of the IGM hydrogen increased the Jeans mass until it was comparable to the characteristic mass scale of dark matter structure. This was a relatively unique event, because throughout most of the history of galaxy formation, both before and longer after reionization, the Jeans mass falls at considerably smaller scales than the characteristic scale of dark matter halos. But as reionization was completing, IGM gas resisted collapse on all but the most massive scales, possibly seriously modifying galaxy formation at that epoch in comparison to earlier and later times. The first results from the *WMAP* satellite suggest that reionization may be a longer and more complicated process than previously believed (Kogut et al., 2003). The study of the reionization epoch will be an important challenge to theoretical and observational astronomy in the coming decade.

My thesis addresses baryonic properties of the universe exclusively, yet generally begins with the basic assumption that baryonic structure traces dark matter structure. Beyond that, though, I have sought observable phenomena that demonstrate the interface of galaxy formation with the IGM. Two of my projects focus on the generation of meta-galactic radiation fields by galaxies. I

considered the integrated radiation field produced by the first generation of stars in the universe, and a future observational test of the spectrum responsible for the completion of reionization. I also studied an important current (and future) observational test of the progress of reionization. My final two projects are devoted to observed manifestations of the impact of reionization on the galaxy formation process.

One theme that underlies my approach to almost all of my thesis work is the relationship between the dynamically evolved universe, such as it is today and as it was around the redshift of reionization, and the density field imprinted on the nascent universe. The association of virialized halos with "peaks" in the initial density field, an analytic prediction confirmed, at least grossly, by numerical simulation (see, e.g., Peebles 1993 for a review), is a very powerful tool, especially now that the properties of the initial density field are tightly constrained by a combination of microwave background and large-scale structure data (Spergel et al., 2003).

The dynamical evolution of a density peak is described analytically usually under the simplifying assumption that the density distribution has a "spherical top-hat" shape, meaning a constant density within some sphere, and the cosmic mean density outside (first described in Partridge & Peebles 1967). For a given background cosmology, this assumption implies a unique relationship between the amplitude of the density peak and the time for the mass within the peak to collapse. Press & Schechter (1974) showed that the relationship between peak height and collapse time could be used to predict the number density of collapsed objects as a function of collapsed mass. This Press-Schechter mass function has been roughly confirmed, though not exactly (see, e.g., Sheth & Tormen 2002), by numerical simulations of structure formation.

Press & Schechter (1974) evaluated the initial density field statistically in order to determine the mass function of collapsed objects in the whole universe, as a function of time. Subsequent work (Bond et al., 1991; Lacey & Cole, 1993) extended the idea of the Press-Schechter analysis to the density distribution around a single peak, in the spirit of spherical top-hats embedded within spherical top-hats (but with different amplitudes). The "excursions" of spherically averaged density around a point in space as a function of the averaging radius could then be used, again in conjunction

with the simple relationship between density and collapse time, to determine statistically the past and future mass evolution of a halo, given that halo's mass at some time.

My thesis began with my study of these techniques, as well as me learning to implement them numerically. Though I don't employ any traditional merger-tree history calculations in this thesis, that approach to structure formation informs all of my thinking about the galaxy formation process. In particular, the "single-burst" star-formation rate model introduced in Chapter 2 is a direct outgrowth of thinking about the first time any branch of the merger tree crosses some physically interesting mass threshold. Another example, in Chapter 3, is the extension of the application of the spherical collapse model to the non-linear, but not yet virialized, stage of mass assembly of a halo. And as shown numerically in the final chapter, as anticipated earlier by analytic work, the peak-patch split formalism may ultimately be a necessary tool in understanding the properties of globular cluster systems.

In the next section I summarize each project in more detail.

1.2 Chapter summaries

The following five chapters are five separate journal articles. For each project I present the motivation, provide a summary of the techniques and major conclusions.

Chapter 2 is an analytic calculation of the contribution of the first stars to the cosmic infrared background (CIRB) (Santos et al., 2002). Two groups detected an unexplained component of the all-sky infrared background at 2.2 μm; that is, after subtracting the contribution from foregrounds and known extra-galactic sources, there was a statistically significant signal remaining. The first generation of star formation in the universe produced stars with primordial composition (called Pop III stars). Existing theoretical models suggested such stars would be very efficient at emitting far-ultraviolet radiation over million-year timescales. Since Pop III stars are predicted to form at $z \sim 10$–20, they were natural candidates for a population of sources that might have produced a radiation background observed in the near-infrared (NIR) today, but that would also have escaped detection as individual sources by even the deepest imaging surveys to date.

We computed the abundance of Pop III star-forming halos, the Pop III star-formation rate, and a model for the reprocessing of ionizing photons into the CIRB. We located a limited region of our parameter space within which Pop III stars would produce the unexplained component of the NIR CIRB: Pop III stars must be very massive, and form efficiently over a redshift range $25 \gtrsim z \lesssim 7$. Moreover, we found that the escape fraction of ionizing photons from the nebular gas immediately surrounding the Pop III stars (i.e., ISM, not IGM, gas) should be small to enhance the Pop III CIRB signal. We incidentally showed that if the escape fraction was near unity, Pop III stars could reionize the universe themselves at a redshift as high as $z \sim 30$, though at the time of publication is was believed that reionization didn't occur before $z \simeq 10$. A lower escape fraction, favored to produce the CIRB, delays reionization. Since our publication of this work, a similar study with similar conclusions was published (Salvaterra & Ferrara, 2003), and two papers have appeared suggesting that new observations of the angular correlation function of NIR CIRB fluctuations favor explanation by Pop III stars (Magliocchetti et al., 2003; Cooray et al., 2004).

This project explored one of the earliest (in cosmic time) connections between galaxy formation and the IGM, the production of ionizing photons and heavy metals by the Pop III stars. More directly, we showed the first stars in the universe may make a significant contribution to the NIR meta-galactic radiation field observed locally.

This work was published as Santos et al. (2002). The idea for this paper was suggested by Volker Bromm, who also supplied the Pop III stellar spectrum, helped with some of the numerical computations, and wrote some of the text of the paper. Marc Kamionkowski supervised the work and contributed to the discussion and presentation of the paper.

Chapter 3 is a calculation of the effect of the radiative transfer through the IGM on the observed properties of Lyman α emission lines from $z \gtrsim 6$ galaxies (Santos, 2004a). Following the detection of a Lyman α emitting galaxy at $z = 6.5$, there were claims in the literature that the very observation of any such Lyman α emission line indicated that the IGM was highly ionized out to $z = 6.5$, at least along that sight-line. The rationale was that if the IGM were neutral at $z = 6.5$, any Lyman α photons emitted by a galaxy would be scattered by the IGM into a low surface brightness

halo, inconsistent with the observations. However, there was also a claim that the observation was consistent with a fully neutral IGM at $z = 6.5$. I attempted to resolve this discrepancy by using a more realistic IGM model than previously considered in this context to explore a wide range of parameters that influence the observed Lyman α emission line from a high-redshift galaxy.

I modeled the dynamics of IGM around a high-redshift galaxy using a model I developed based on a suggestion by K. L. Adelberger, and submitted independently by Barkana (2004). This model extended the excursion-set formalism and spherical top-hat model used to calculate halo formation histories to describe the density and velocity profiles of the IGM around a galaxy. The ionization state of this IGM was described by an overall ionizing background or ionization fraction and a contribution from the ionizing flux of the galaxy itself (a proximity effect). The properties of the emitted Lyman α line were calculated self-consistently with the ionizing radiation. I concluded that the many free parameters associated with interpreting an observed Lyman α line without any other information about the source galaxy make it almost impossible to draw conclusions about the ionization state of the IGM. In particular, galactic-scale super-winds, if present, play a crucial and previously unrecognized role in the observability of Lyman α emission from $z \gtrsim 6$ galaxies.

This project critically analyzed Lyman α lines as a probe of the reionization history. Given recent success detecting Lyman α lines at up to $z = 6.5$, they could play a key role in constraining the reionization history at even higher redshifts. I described the complementary observations required to reconstruct the IGM ionization state from current and future Lyman α emission line detections.

This chapter was published as Santos (2004a).

Chapter 4 describes a method to infer the stellar population responsible for completing the reionization process (Santos & Loeb, 2003). Deep x-ray surveys suggest that stars, rather than accreting black holes, emitted the ionizing photons that completed the reionization process at $z \sim 6$. However, particularly in light of the *WMAP* results, it is still an open question whether the ionizing background during reionization was dominated by massive, metal-free Pop III stars, or garden-variety Pop II stars. One approach to this question is to characterize individual sources at $z \sim 6$, assessing their ionizing flux. We proposed an alternative, to infer the spectrum of the ionizing background

using the relative ionization states of hydrogen and helium. Since Pop III stars are predicted to have much harder spectra than Pop II stars, we expected the IGM ionization state to be sensitive to the difference between these scenarios.

We created Monte Carlo simulations of lines-of-sight toward high-redshift QSOs of the sort discovered recently by the Sloan Digital Sky Survey (SDSS). We fixed our model of the hydrogen absorption systems to match current observations, then predicted the absorption due to He I as a function of ionizing background spectrum. (He II absorption is a more sensitive test for a hard ionizing spectrum, but measuring this will likely never be feasible toward $z \sim 6$ quasars.) We showed that the final SDSS quasar sample is reasonably likely to contain a quasar/line-of-sight combination useful for constraining the shape of the ionizing spectrum, and thus the nature of the dominant sources, at $z \sim 5$, only 300 Myr after the end of the reionization epoch.

This project proposed a relatively direct test of the nature of the sources responsible for the ionizing background radiation at the end of reionization.

This chapter was submitted to the Monthly Notices of the Royal Astronomical Society in 2003 as Santos & Loeb (2003). Avi Loeb suggested pursuing He I absorption as a project, and contributed to the discussion and presentation of the paper.

Chapter 5 presents the results and interpretation of an observational survey of low-luminosity Lyman α emitting galaxies at $4.5 < z < 6.7$ (Santos et al., 2004). This study utilized strong lensing by clusters of galaxies at intermediate redshifts to search for Lyman α emitting galaxies at unprecedented depth, complementary to many other surveys for higher Lyman α luminosity systems.

Our survey primarily used the LRIS spectrograph on the Keck I telescope to take spectra blindly of the sky near strong-lensing clusters. We chose clusters for which accurate mass models have been constructed, allowing us to focus our survey area on the sky regions with highest magnifications for sources at our redshifts of interest: half of our survey area is magnified by more than a factor of 10 (reaching similar depths over the same area without lensing would take 10 times more telescope time than our technique). We used *HST* imaging and additional spectroscopy to separate low-redshift interlopers from our high-redshift Lyman α emission line sample. Our survey volume contains 3

confirmed and 3 likely emission lines identified as Lyman α at $z > 4.5$, up to $z = 6.25$. We reconstruct our complicated survey volume and present a measurement of the cumulative number density of Lyman α emitters down to an order of magnitude fainter than previous surveys achieved. At the Lyman α luminosities where our data overlap other surveys, there is marginal consistency. Analysis of all available data suggested that at $z \sim 5$ the Lyman α luminosity function may be flatter than the mass function of dark-matter halos, a sign of suppression of the star-formation efficiency in low-mass halos. However, our data by themselves are consistent with the slope of the halo mass function, and we present alternate interpretations.

This project approached the direct detection of suppression of star-formation in low-mass $z \sim 5$ halos, a prediction of reionization theory. Our observations have not conclusively demonstrated either a star-formation suppression or a lack thereof, but we have demonstrated that this measurement is possible as more data become available in the near future.

This chapter was published as Santos et al. (2004). This project was initiated by Richard Ellis, Jean-Paul Kneib, and Konrad Kuijken, in 2000; Richard Ellis served as PI on all of the observations made for this project. Data were reduced by Konrad Kuijken, Johan Richard, and me. Data were analyzed by all members of the collaboration, but identification of candidate Lyman α emission lines was performed primarily by Richard Ellis, Konrad Kuijken, and Johan Richard. I calculated the survey selection function and performed the analysis of our number density data, as well as the observational and theoretical comparisons. I wrote most of the paper, with assistance from Richard Ellis and Johan Richard, and contributions to the discussion and presentation from Jean-Paul Kneib and Konrad Kuijken.

Chapter 6 presents a new model for the formation of metal-poor globular cluster systems (GCSs) (Santos, 2004b). The metallicity distribution of the Milky Way GCS is bimodal, with peaks at [Fe/H]$\simeq -1.6$ and -0.6, and a relative lack of globular clusters at [Fe/H]$\simeq -1$. This metallicity gap is universally observed in the metallicity distributions of giant-galaxy GCSs, and thus is an important part of a successful model of GCS formation. Many other observations challenge theoretical models of GCS formation, such as the mass of the Milky Way GCS, the concentration of Milky Way globular

clusters within 40 kpc of the galactic center, and the observed variation of the number of globular clusters per unit galaxy mass (specific frequency) as a function of galactic environment. In contrast to some existing globular cluster formation models, I sought to explain these observations in the context of the modern cosmological framework of galaxy formation.

I first argue that reionization played a crucial role suppressing the formation of metal-poor globular clusters, resulting in the bimodality observed in GCS metallicity distributions. The assumption that metal-poor globular clusters formed before the end of reionization is powerful because it allowed me to use structure formation theory to compute the abundance and properties of metal-poor globular cluster formation sites, which I did using constrained realizations of the initial density perturbations in a region of space that evolved into a Milky Way-like galaxy at $z = 0$. I show that our model reproduced the mass of the Milky Way metal-poor GCS. I use N-body simulations to follow the collisionless evolution of metal-poor globular cluster formation sites, to demonstrate for the first time in detail that a pre-galactic globular cluster formation scenario naturally predicts a Milky metal-poor GCS spatial extent exactly matching observations. Additionally, I show that the correlation between metal-poor specific frequency and galaxy environment is also naturally predicted, and arises due to the correlation between galactic environment and collapse history.

This project introduced reionization as a key component of globular cluster system formation. The impact of reionization on globular cluster formation may explain the universal bimodality of globular cluster system metallicity distributions. Moreover, I showed that this model passes some important observational tests.

The work in this chapter will be submitted as a single-author paper to the Monthly Notices of the Royal Astronomical Society in 2004.

Chapter 2

The Contribution of the First Stars to the Cosmic Infrared Background[1]

2.1 Abstract

We calculate the contribution to the cosmic infrared background from very massive metal-free stars at high redshift. We explore two plausible star-formation models and two limiting cases for the reprocessing of the ionizing stellar emission. We find that Population III stars may contribute significantly to the cosmic near-infrared background if the following conditions are met: (i) The first stars were massive, with $M \gtrsim 100$ M$_\odot$. (ii) Molecular hydrogen can cool baryons in low-mass haloes. (iii) Pop III star formation is ongoing, and not shut off through negative feedback effects. (iv) Virialized haloes form stars at about 40 per cent efficiency up to the redshift of reionization, $z \sim 7$. (v) The escape fraction of the ionizing radiation into the intergalactic medium is small. (vi) Nearly all of the stars end up in massive black holes without contributing to the metal enrichment of the Universe.

[1]This chapter appeared as Santos, Bromm, and Kamionkowski (2002), and is reproduced here by permission of the copyright holder, the Royal Astronomical Society.

2.2 Introduction

It has long been realized that observations of the cosmic infrared background (CIRB) can place important constraints on the energy production associated with the formation of cosmological structure (see Hauser & Dwek 2001 and references therein). The cosmic background is the locally measured radiation density from all extragalactic sources. Starlight dominates the CIRB in the near-IR, whereas the mid- and far-IR backgrounds result primarily from dust emission (e.g., Dwek et al., 1998).

Observational efforts to measure the near-IR CIRB are hampered by strong and uncertain foreground contamination from zodiacal dust. Cambrésy et al. (2001), using data from the Diffuse Infrared Background Experiment (DIRBE) on the *COBE* satellite, found that the integrated light from all galaxies (estimated from deep surveys) is not sufficient to account for the measured CIRB at 1.25 μm (J band) or 2.2 μm (K band). Wright & Johnson (2001) analyzed the same DIRBE data as Cambrésy et al. (2001), but subtracted a different zodiacal light model; they find a J-band CIRB that is consistent with the contribution from galaxies, but their K-band CIRB value is larger than the background inferred from galaxy counts. Totani et al. (2001) model the contribution from galaxies missed by deep galaxy surveys and find that it is unlikely to be greater than 30 per cent of the measured light.

Bond et al. (1986) suggested that the first generation of stars in the Universe, called Population III (Pop III) stars because they are assumed to have near-zero metallicity, may contribute significantly to the cosmic background. The cosmological impact of Pop III stellar radiation has been addressed before (e.g., Carr et al., 1984; Bond et al., 1986; Couchman & Rees, 1986; Haiman & Loeb, 1997; Ciardi et al., 2001). Recent theoretical advances have improved our understanding of the physics of Pop III star formation (see Barkana & Loeb 2001 and references therein); in particular, three-dimensional numerical simulations suggest that the primordial initial mass function (IMF) may have been concentrated at stellar masses $\gtrsim 100$ M$_\odot$ (Bromm et al., 1999, 2002; Abel et al., 2000). Stars with masses $\gtrsim 100$ M$_\odot$, which we will call 'very massive' (Carr et al., 1984), have spectra and luminosities that asymptotically approach the blackbody form and the Eddington limit, respectively (e.g., Bromm et al., 2001b).

In this paper, we investigate whether a significant fraction of the near-IR CIRB may come from very massive Pop III stars. Since these stars are very luminous per unit stellar mass over a short lifetime, they may contribute substantially to the CIRB. We use two different models for the formation of dark matter (DM) haloes to calculate the rate at which baryons are processed through haloes with virial temperatures high enough to allow baryonic cooling (§2.3). In addition, we consider two possibilities for the re-processing of the stellar emission by the gas in the halo and by the intergalactic medium (§2.4). In §2.5 we show our model results and the observational data. Section 2.6 contains a discussion of our results and compares them with the data.

We adopt the following values for the cosmological parameters: $h = 0.7$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_{\rm B}h^2 = 0.019$, $\sigma_8 = 0.9$. Here h is the dimensionless Hubble constant, $H_0 = 100h$ km s^{-1}Mpc^{-1}; Ω_m, Ω_Λ and $\Omega_{\rm B}$ are the total matter, cosmological constant and baryon density in units of the critical density and σ_8 gives the normalization of the power spectrum on the $8h^{-1}$Mpc scale.

2.3 Star-formation rate

2.3.1 Overview of the star-formation model

We assume that the star-formation rate (SFR) of Pop III stars is related to the cooling of baryons in collapsed dark-matter haloes (henceforth just 'haloes'). The cooling rate of baryons is a very strong function of baryon temperature near certain temperature thresholds: above these thresholds a new cooling mechanism can act within the gas. For a given choice of cooling mechanism, we would like to calculate the rate at which baryons are heated above the associated temperature threshold, and thus may collapse to high density and form stars.

If baryons in a halo are shock-heated to the virial temperature of their halo, the temperature of those baryons is related to the mass of that halo according to $T \propto M^{2/3}$. So as a proxy for the temperature history of the baryons, we may calculate the mass history of haloes, assuming a constant ratio of baryonic matter to dark matter in all haloes, and that baryons below the cooling threshold are always heated to the virial temperature of their halo. In this picture haloes above a critical

mass may form stars, because those haloes have heated their gas above the critical temperature, and haloes below the critical mass do not form any stars. Additionally, we parametrize the fraction of eligible baryons that actually do form stars with a constant star formation efficiency, η.

The mass-assembly history of dark-matter haloes is computed using the extended Press-Schechter formalism (Lacey & Cole, 1993), where the $\sigma(M)$–M relation is evaluated with the power spectrum of Eisenstein & Hu (1999). Here $\sigma(M)$ is the standard deviation of the linear density field smoothed on scales containing a mean mass M. Our choices of $M_{\rm crit}(z)$ are described in §2.3.3.

We consider two models for the SFR. In both models, star formation is triggered in a halo as it accumulates enough mass to put it above the threshold mass $M_{\rm crit}(z)$, but the models differ in how they treat additional material that merges into such haloes. In the 'ongoing' model, all gas in haloes with $M \geq M_{\rm crit}(z)$ is eligible for star formation. Conversely, the 'single-burst' model only allows star formation in haloes that have not previously formed stars and have no progenitor halo that has formed stars, i.e., both merging haloes are crossing the critical threshold for the first time.

2.3.2 Specific star-formation models

2.3.2.1 Ongoing star formation

In the ongoing model, the SFR in a single halo is proportional to the growth (in mass) of that halo, after it has become more massive than $M_{\rm crit}(z)$. Star formation does not occur in a quiescently evolving (i.e., non-accreting) halo; the star formation is not *continuous* in the usual sense. But star formation is not inhibited after the first generation of Pop III stars is born in a halo, which assumes that radiation and mechanical outflows from the life and death of the star(s) have no impact (zero feedback) on future star formation.

The SFR in the ongoing model is proportional to the time derivative of the total mass contained in haloes that have mass $M \geq M_{\rm crit}(z)$. The SFR per comoving volume, $\psi_{\rm on}$, is

$$\psi_{\rm on}(z, M_{\rm crit}(z)) = \eta \frac{\Omega_{\rm B}}{\Omega_m} \frac{\mathrm{d}}{\mathrm{d}t} \int_{M_{\rm crit}(z)}^{\infty} \mathrm{d}M \; M \frac{\mathrm{d}n_{\rm PS}}{\mathrm{d}M}(M, z). \tag{2.1}$$

Here $n_{\rm PS} \equiv n_{\rm PS}(M, z)$ is the comoving number density of haloes of mass M at redshift z, given by Press & Schechter (1974). The integral expresses the collapsed mass per comoving volume contained in haloes above the critical mass. The time derivative converts this to a mass rate, and the prefactors convert from total mass to stellar mass.

2.3.2.2 Single-burst star formation

In the single-burst model, a halo forms stars when the accretion of matter pushes its mass above the critical mass, $M_{\rm crit}(z)$; that is the only time that the halo will form stars, regardless of additional mergers and mass accretion. This is a very extreme model which assumes Pop III stars make a long-term change in their surroundings which permanently prohibits future Pop III stars from forming there. This extreme model could be realistic if Pop III stars heavily enrich the interstellar medium with metals. Then, though stars will likely eventually form out of that material, Pop III stars will not. The SFR may lie somewhere between the prediction of this model, and the prediction of the ongoing model discussed above (in which all baryons in haloes with $M > M_{\rm crit}(z)$ are eligible to form stars).

The SFR in the single-burst model can be calculated with the extended Press-Schechter formalism: A halo grows in mass through discrete mergers and gradual accretion, which is treated as a series of small mergers in extended Press-Schechter theory. As an example, take a halo with mass M_1 at some redshift z that merges with mass ΔM for a total mass of $M_2 = M_1 + \Delta M$ at redshift $z + {\rm d}z$, where ${\rm d}z$ is small and negative, i.e., a short time later. If $M_1 < M_{\rm crit}(z)$ and $M_2 \geq M_{\rm crit}(z)$, then, provided $\Delta M < M_{\rm crit}(z)$, the halo undergoes a burst of star formation. We parametrize the mass of stars formed, M_*, by

$$M_* = \eta \frac{\Omega_{\rm B}}{\Omega_m} M_2. \tag{2.2}$$

The SFR per comoving volume, $\psi_{\rm burst}$, is

$$\psi_{\rm burst}(z, M_{\rm crit}(z)) = \frac{1}{2} \frac{\eta\ \Omega_{\rm B}}{\Omega_m}$$

$$\times \int_0^{M_{\rm crit}(z)} {\rm d}M_1 \frac{{\rm d}n_1}{{\rm d}M_1} \int_{M_{\rm crit}(z)}^{M_{\rm crit}(z)+M_1} {\rm d}M_2 \; M_2 \frac{{\rm d}^2 P}{{\rm d}M_2 \; {\rm d}t}, \tag{2.3}$$

where $n_1 \equiv n_{\rm PS}(M_1, z)$ and $P \equiv P(M_1, M_2, z)$ is the probability that a halo with mass M_1 merges to a new mass $M_2 > M_1$ at redshift z (Lacey & Cole, 1993). The prefactor of $1/2$ corrects for double counting in the integrals (for fixed M_1 and ΔM, the integral as written counts both $M_1 + \Delta M = M_2$ and $\Delta M + M_1 = M_2$). [2] A similar approach to model the history of star formation at high redshifts has been discussed by Barkana & Loeb (2000).

2.3.3 Critical halo mass for star formation

There is a relationship between the mass of a halo and its virial temperature. Using the assumptions we made in §2.3.1, we can substitute a critical temperature for baryonic cooling, $T_{\rm crit}$, into that relation to find the corresponding critical mass:

$$M_{\rm crit}(z) = 0.94 \times 10^8 \; {\rm M}_\odot \left(\frac{h}{0.7}\right)^{-1} \left(\frac{\Omega_m}{0.3}\right)^{-1/2}$$

$$\times \left(\frac{1+z}{10}\right)^{-3/2} \left(\frac{\mu}{0.6}\right)^{-3/2} \left(\frac{T_{\rm crit}}{10^4 \; {\rm K}}\right)^{3/2}, \tag{2.4}$$

where μ is the mean molecular weight ($\mu = 0.6$ for ionized gas [$T_{\rm crit} \gtrsim 10^4$ K] and $\mu = 1.2$ for neutral gas [$T_{\rm crit} \lesssim 10^4$ K]) (Barkana & Loeb, 2001).

We seek to model only two cooling processes: radiative cooling from molecular hydrogen and radiative cooling from atomic hydrogen. For molecular cooling, we choose $T_{\rm crit} = 400$ K (e.g., Tegmark et al., 1997; Abel et al., 1998). For atomic cooling, we choose $T_{\rm crit} = 10^4$ K (Barkana & Loeb, 2001).

[2] There is some ambiguity in the SFR arising from the extended Press-Schechter merger rates. We defer discussion of that to future work, and here evaluate the SFR as written.

2.3.4 Limits on the abundance of Pop III stars

A realistic model of Pop III star formation must conform to some observational constraints. The matter density processed through Pop III stars as a fraction of the critical density is

$$\Omega_{\rm III}(z) = \frac{1}{\rho_{\rm c}} \int_z^\infty \psi(z') \left|\frac{{\rm d}t'}{{\rm d}z'}\right| {\rm d}z', \tag{2.5}$$

with

$$\left|\frac{{\rm d}t}{{\rm d}z}\right| = \left[H_0(1+z)E(z)\right]^{-1},$$

and, in a flat universe,

$$E(z) = [\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}.$$

Here, ψ is the SFR and $\rho_{\rm c}$ is the density required to close the Universe. A strong upper limit on $\Omega_{\rm III}$ is that it must not exceed $\Omega_{\rm B}$. Below we discuss two constraints on the possible products of Pop III stars, black holes and heavy elements (see also Schneider et al. 2002b).

2.3.4.1 Black holes

A nonrotating star with mass above ~ 260 M$_\odot$ is predicted to evolve directly to a black hole without any metal ejection (Fryer et al., 2001). The fate of rotating stars in that mass range is unknown, but may be similar. If so, and if the Pop III IMF were concentrated at masses above ~ 260 M$_\odot$, most of the baryons once in Pop III stars would now be in black holes. Such black holes would populate the halo of our galaxy, but they would be very difficult to detect unless they have a companion; e.g., the MACHO microlensing study places no constraint on the fraction of the Milky Way halo contained in $\gtrsim 100$ M$_\odot$ black holes (Alcock et al., 2001). Detailed theoretical studies may eventually reveal the fraction of such black holes that currently have a companion object, and searches for those binary systems, or future microlensing studies, could lead to constraints on $\Omega_{\rm BH}$ (e.g., Agol & Kamionkowski, 2002). But currently even if $\Omega_{\rm BH} = \Omega_{\rm III}$ there is no additional constraint on $\Omega_{\rm III}$ from limits on black holes.

2.3.4.2 Metal enrichment

A nonrotating star with mass between ~ 140 $M_\odot$ and ~ 260 $M_\odot$ is expected to end its evolution in a pair-instability supernova, which completely disrupts the star and leaves no remnant (Heger & Woosley, 2002). In this case all of the nucleosynthetic products of the star are ejected, and potentially pollute the intergalactic medium (IGM) (cf. §2.3.4.1, where all were swallowed into a black hole). Thus metal abundances in the Universe, especially at high redshift, are limits on the number of Pop III stars which resulted in supernovae, given assumptions about the mixing of the metals.

If a fixed fraction $\epsilon_{\rm pi}$ ('pi' for 'pair instability') of the mass converted into Pop III stars is expelled in pair-instability supernovae, the nucleosynthetic yield calculation from Heger & Woosley (2002) may be used to determine the contribution from Pop III stars to the metallicity of the IGM. For

$$[i] \equiv \log_{10}\left(\frac{f_i^U}{f_i^\odot}\right), \tag{2.6}$$

where i is a species and f_i^U and $f_i^\odot$ are the mass fractions of species i in the Universe and Sun, respectively. Similarly, if we denote the ejected mass fraction of species i by f_i^E, then

$$f_i^U = \epsilon_{\rm pi}\frac{\Omega_{\rm III}}{\Omega_B} f_i^E \tag{2.7}$$

and

$$[i] = \log_{10}\left(\epsilon_{\rm pi}\frac{\Omega_{\rm III}}{\Omega_{\rm B}} P_i\right), \tag{2.8}$$

where $P_i \equiv f_i^E/f_i^\odot$ is the production factor of species i. Heger & Woosley (2002, table 4) calculated P_i for many isotopes and stellar masses. We here take ~ 200 $M_\odot$ as a fiducial mass for a pair-instability supernova progenitor.

A few values of interest from their table are $P_i(i = {}^{12}{\rm C}, {}^{16}{\rm O}, {}^{24}{\rm Mg}, {}^{28}{\rm Si}, {}^{56}{\rm Fe}) = (13.2, 45.8, 85.7, 353, 49.8)$. In general, the highest values of P_i are for the alpha elements.

Observations of the abundance of species i in the high-redshift universe are sometimes quoted

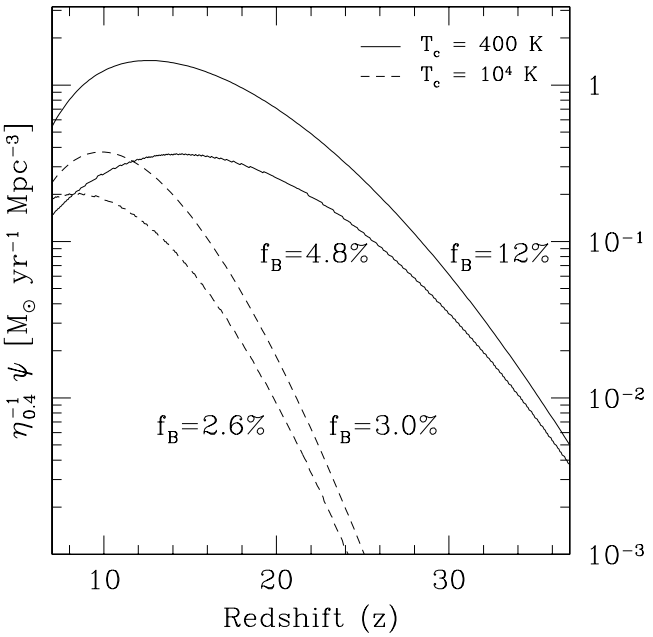


Figure 2.1 The star-formation rate (SFR) for the ongoing and single-burst models, for critical temperatures of $T_c = 400$ (*solid lines*) and 10^4 K (*dashed lines*). For each T_c the upper curve is the ongoing SFR and the lower curve is the single-burst SFR. All curves are for $\eta = 0.4$; each curve is labelled by the value of $f_{\rm B} \equiv \Omega_{\rm III}(z = 7)/\Omega_{\rm B}$ it implies.

in slightly different notation from [i]. Assuming $\epsilon_{\rm pi}\Omega_{\rm III} \ll \Omega_{\rm B}$, $[i/{\rm H}] \simeq [i]$. Using tabulated values for f_i^E (Heger & Woosley, 2002, table 3), it is straightforward to calculate $\Omega_i = f_i^U \Omega_{\rm B}$. With assumptions about the mixing of the ejecta from Pop III stars, measurements of metals in the IGM at high redshifts (or the metallicities of stars formed at high redshifts) can place a limit on a combination of the IMF of Pop III stars (which we simplify into the parameter $\epsilon_{\rm pi}$) and $\Omega_{\rm III}$. For example, if $\epsilon_{\rm pi} = 1$ and $f_{\rm B} \equiv \Omega_{\rm III}/\Omega_{\rm B} \simeq 3 \times 10^{-3}$, then Pop III stars would, in the mean, enrich the Universe to the solar abundance of silicon already at a redshift $\gtrsim 7$. In contrast, if $\epsilon_{\rm pi} = 0$, then metallicity measurements place no constraints on the abundance of Pop III stars.

2.3.5 Star-formation model results

To compare our models, we present the SFR as a function of redshift in Fig. 2.1. We assume that Pop III star formation shuts off at $z_{\rm end} \simeq 7$; this may be due to photoevaporation of low-mass haloes due to reionization (Barkana & Loeb, 1999) or some other process. Figure 2.1 shows the SFRs normalized to $\eta = 0.4$. The curves are labelled with the corresponding value of $f_{\rm B}$.

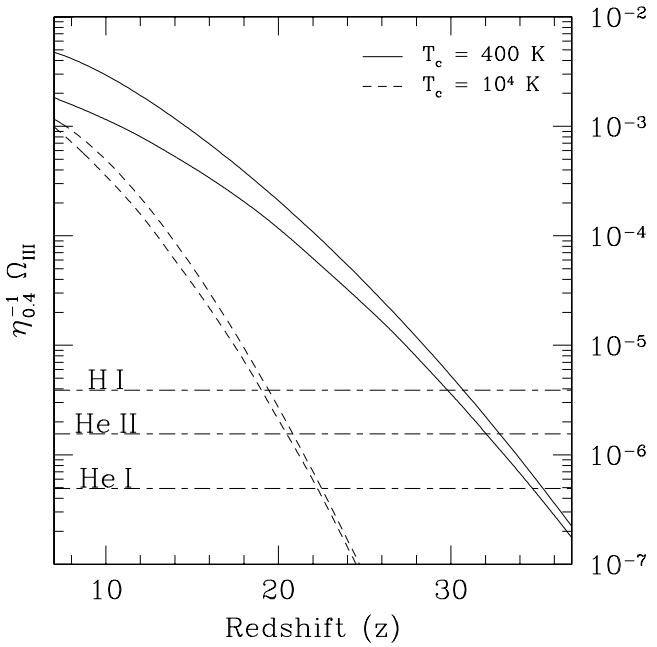


Figure 2.2 The matter density processed through Pop III stars in units of the critical density, as a function of redshift. The four curves correspond to the four different SFRs of Fig. 2.1. The horizontal dashed lines labelled H I, He I and He II correspond to the $\Omega_{\rm III}$ required for Pop III stars to produce 10 ionizing photons per particle of each species. These lines are only relevant for the $f_{\rm esc} = 1$ case; see section 2.4.

Figure 2.2 shows $\Omega_{\rm III}(z)$ for the SFRs of Figure 2.1. The horizontal lines are explained in §2.6.

2.4 Spectra

2.4.1 Input stellar spectrum

Metal-free very massive stars have spectra similar to a $\sim 10^5$ K blackbody spectrum (Bromm et al., 2001a). Most of the energy is then radiated in photons with energies > 13.6 eV. There are also many photons produced capable of ionizing He II (but see Schaerer, 2002). The specific luminosity per solar mass of very massive stars is almost independent of stellar mass (Bromm et al., 2001a). We will therefore take the spectrum of a 1000 $M_\odot$ star as our fiducial input Pop III stellar spectrum.

When a Pop III star becomes luminous, we expect the host halo to contain gas not incorporated into stars; we call this gas the 'nebula.' We refer to the gas outside the collapsed halo as the IGM. The physical environment of Pop III star formation will likely undergo a transition from dense gas

in the nebula to a diffuse IGM, but we will model all gas as either part of the nebula or part of the IGM. *We also assume that both the nebula and IGM are entirely free of dust.*

Since the nebula and IGM are neutral in the absence of stellar radiation, both have the potential to play an important role in reprocessing ionizing photons from a star. We examine two limiting cases for the importance of each of these phases: In the first case, Pop III stars are enshrouded in dense nebulae, and all of the reprocessing of ionizing radiation takes place in the halo (the IGM still plays a role in scattering Lyα photons). In this case, the escape fraction of ionizing photons from the nebula, f_{esc}, is zero. In the second case, the nebula plays no role and $f_{esc} = 1$; all reprocessing occurs in the IGM. In the rest of this section, we discuss the reprocessed spectrum of Pop III stellar radiation for each of these cases.

2.4.2 No escape of ionizing radiation into the IGM

2.4.2.1 Properties of a nebula

Numerical simulations suggest that when a Pop III star forms, the nebula consists of a higher density phase, with $n_{\rm H} \simeq 10^4$ cm^{-3}, and a lower density phase (e.g., Bromm et al., 1999). We make the simplifying assumption that half of the nebula's mass is contained in a homogeneous phase with density $n_{\rm H} = 10^4$ cm^{-3} that completely covers the star(s), and ignore the lower density gas. We take the mass fractions of hydrogen and helium as $X = 0.75$ and $Y = 0.25$, respectively.

Ionizing radiation from the star(s) creates an H II region in the dense nebula. Because of the hardness of our input spectrum, in the inner part of the H II region helium is doubly ionized (the He III region). In the outer part it is singly ionized (the He II region; the spectrum is hard enough that there is no H II/He I region). The majority of photons above the He II ionization threshold ionize He II rather than H I. We iteratively solve for the sizes and temperatures of these regions, using the thermodynamic equations from Cen (1992). The He III region, comprising 0.4 of the volume of the H II region, has a temperature of 3.6×10^4 K; the He II region is cooler, at 2.7×10^4 K. In both regions the primary cooling mechanism is H I recombination, but free-free emission and cooling via collisional excitation of H I are also important. The total nebular emission is not very sensitive to

the relative sizes of the He II and He III regions.

Given these properties of an ionized region, we can determine $f_{\rm esc}$.[3] The volume and mass of the He III region are

$$V_{\rm HeIII} = \frac{Q_{\rm HeII}}{n_{\rm e} n_{\rm HeIII} \alpha_{\rm B}({\rm He\,II}, T = 3.6 \times 10^4\ {\rm K})}, \tag{2.9}$$

$$M_{\rm HeIII} = \mu\, m_{\rm H}\, (n_{\rm H} + n_{\rm He})\, V_{\rm HeIII} \simeq 0.7 \left(\frac{M_*}{{\rm M}_\odot}\right), \tag{2.10}$$

where $Q_{\rm HeII}$ is the stellar emission rate of photons energetic enough to ionize He II, $\alpha_{\rm B}({\rm He\,II}, T)$ is the Case B recombination coefficient for He II, here M_* is the mass of the ionizing Pop III star(s) in the nebula and $\mu = 1.2$ is the mean molecular weight. Because in the He III region recombinations to He II provide enough photons to keep the hydrogen ionized, the volume and mass of the He II region are

$$V_{\rm HeII} = \frac{Q_{\rm HI} - Q_{\rm HeII}}{n_{\rm e} n_{\rm HII} \alpha_{\rm B}({\rm H\,I}, T = 2.7 \times 10^4\ {\rm K})}, \tag{2.11}$$

$$M_{\rm HeII} = \mu\, m_{\rm H}\, (n_{\rm H} + n_{\rm He})\, V_{\rm HeII} \simeq 1.0 \left(\frac{M_*}{{\rm M}_\odot}\right), \tag{2.12}$$

where $Q_{\rm HI}$ is the rate of H I-ionizing photons, and $\alpha_{\rm B}({\rm H\,I}, T)$ is the hydrogen recombination coefficient.

The total mass of ionized gas is then 1.7 times the mass of the ionizing star(s). In the single-burst star-formation model, the nebula mass in $n_{\rm H} = 10^4$ cm^{-3} gas, $M_{\rm neb}$, is

$$M_{\rm neb} = \frac{1}{2}\, \frac{1-\eta}{\eta}\, M_*. \tag{2.13}$$

For $\eta = 0.4$, $M_{\rm neb} = 0.75\, M_*$, which implies $f_{\rm esc} = 0$ would be difficult to achieve for our model nebula. In the ongoing model, though, $M_{\rm neb} > (1-\eta) M_*/(2\eta)$ in general, and especially at lower redshift. This is due to the short lifetime of Pop III stars: star formation in a halo subsequent to the first episode usually takes place after previous generations of stars in the halo have stopped radiating. Thus $f_{\rm esc} = 0$ may be possible in these haloes.

[3]The dynamical evolution of the gas on the timescale of the lifetime of a Pop III star may be important, as the pressure of the nebular gas may exceed the gravitational pressure, but here we treat the nebula as static.

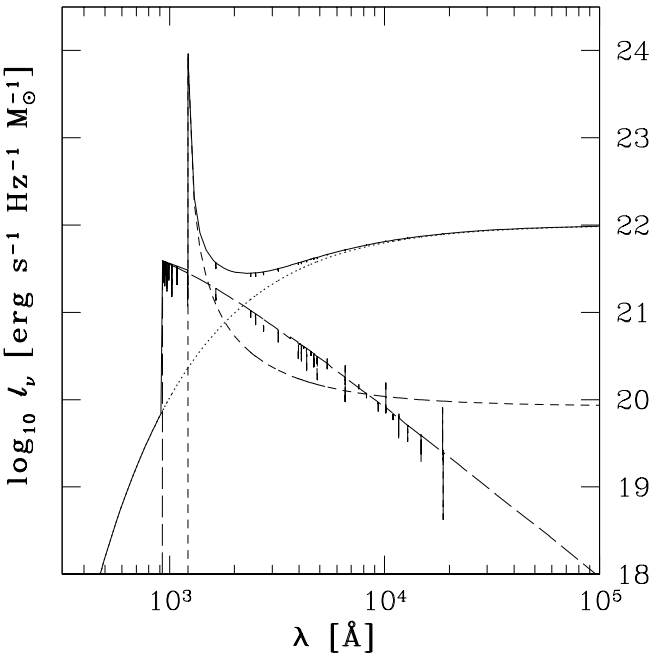


Figure 2.3 Spectrum of a $z = 15$ Pop III star plus nebular emission, for the $f_{\rm esc} = 0$ case of section 2.4.2. The long-dashed line is the spectrum of the star, cut off for absorption shortward of the Lyman limit. The dotted line is the spectrum of free-free emission from the nebula. The short-dashed line is emission from Lyα recombination in the nebula at $z = 15$, corrected for scattering in the IGM. The solid line is the sum of the spectra. All spectra are in the rest frame of the star.

2.4.2.2 Resulting spectrum

Recombination emission. Approximately 1/2 of the energy radiated by a Pop III star in a nebula is ultimately reradiated by recombinations. The mean energy of a free electron just before it recombines with a proton may be estimated from

$$\langle E \rangle = \frac{\beta_{\rm B}({\rm H\,I}, T)\, kT}{\alpha_{\rm B}({\rm H\,I}, T)}, \tag{2.14}$$

where $\beta_{\rm B}({\rm H\,I}, T)$ is the recombination emission coefficient. Using $T = 3{\times}10^4$ K, we find $\langle E \rangle = 1.4$ eV (Seaton, 1959; Cen, 1992). The average energy radiated per recombination, then, is 15.0 eV. Under the on-the-spot approximation, all Lyman series photons degrade into $n = 2 \rightarrow 1$ transitions, where n is the energy level of an excited hydrogen atom. Since 10.2 eV is released per $n = 2 \rightarrow 1$ transition, about 2/3 of the recombination energy is released in Lyα or two-photon emission, and 1/3 of the recombination energy (and about 1/6 of the total emitted energy) is emitted in other recombination

lines and free-bound continuous emission. We ignore those processes here; see Schaerer (2002) for a more complete computation of nebular emission lines and free-bound spectra.

The relative importance of Lyα emission compared to two-photon emission is determined by the effective recombination coefficients to the $2p$ and $2s$ states, respectively, and the collisional excitation rate from the $2s$ state to the $2p$ state. At $T = 3 \times 10^4$ K, 0.75 Lyα photons are emitted for every H I recombination (Osterbrock, 1989; Storey & Hummer, 1995); we ignore the contribution of two-photon emission (about 12 per cent of the total emission) to the spectrum.

Since Lyα photons resonantly scatter in neutral hydrogen, they will not travel far in the IGM until their frequencies are shifted away from the resonant frequency, $\nu_{\mathrm{Ly}\alpha}$. Photons initially scattered blueward of the line resonance will eventually cosmologically redshift back into the resonance. The result is that the IGM ultimately scatters all Lyα photons to the red side of the line resonance (broadening from motions inside a halo doesn't contribute to the line profile). For a homogeneous, expanding IGM, the resulting scattered line profile, $\phi(\nu, z)$, was simulated by Loeb & Rybicki (1999), and we fit their result with

$$\phi(\nu, z) = \begin{cases} \nu_*(z)\, \nu^{-2} \exp\left[\frac{-\nu_*(z)}{\nu}\right] & \text{if } \nu > 0 \\ 0 & \text{if } \nu \leq 0 \end{cases}, \tag{2.15}$$

$$\nu_*(z) = 1.5 \times 10^{11} \text{ Hz } \left(\frac{\Omega_{\mathrm{B}} h^2}{0.019}\right) \left(\frac{h}{0.7}\right)^{-1} \frac{(1+z)^3}{E(z)}. \tag{2.16}$$

This profile results in a strong, asymmetric Lyα emission line near $1220 - 1225$ Å with a scattering tail extending to long wavelengths.

In the He III region, He II recombinations sometimes produce more than one photon capable of ionizing H I. Recombinations directly to the $n = 2$ state produce He II Balmer continuum photons, which are capable of ionizing H I. In addition, two-photon decay from the $2s$ state produces 1.42 H I-ionizing photons per decay (Osterbrock, 1989). Since He II Lyα is also capable of ionizing hydrogen,

the mean number of hydrogen ionizations per He II recombination is

$$\frac{\alpha_2(\mathrm{He\,II}, T) + 1.42\alpha_{2s}^{\mathrm{eff}}(\mathrm{He\,II}, T) + \alpha_{2p}^{\mathrm{eff}}(\mathrm{He\,II}, T)}{\alpha_{\mathrm{B}}(\mathrm{He\,II}, T)} = 1.7, \tag{2.17}$$

where $\alpha_2(\mathrm{He\,II}, T)$ is the recombination coefficient for recombinations directly to the $n = 2$ state and $\alpha_{2l}^{\mathrm{eff}}(\mathrm{He\,II}, T)$ is the effective total recombination coefficient to the $2l$ state (Storey & Hummer, 1995).

Free-free emission. Free-free radiation accounts for about 1/4 of the cooling in the nebula. This energy is radiated in a continuous spectrum,

$$j_\nu^{\mathrm{ff}} = 7.2 \times 10^{-39} \sum_Z n_{\mathrm{e}} n_Z \, Z^2 \left(\frac{T}{\mathrm{K}}\right)^{-1/2}$$

$$\times \exp\left(\frac{-h\,\nu}{k\,T}\right) \ \mathrm{erg\ s^{-1} cm^3\,Hz^{-1} ster^{-1}}, \tag{2.18}$$

where j_ν^{ff} is the specific emission coefficient, ν is the frequency of emitted radiation and n_Z is the number density of ions of net charge Z (Ferland, 1980) A Gaunt factor of 1.3 has been assumed; this results in an error of less than 8 per cent over optical and UV frequencies (Karzas & Latter, 1961).

The luminosity per solar mass of Pop III stars from free-free radiation, l_ν^{ff}, is

$$l_\nu^{\mathrm{ff}} = \frac{4\pi}{M_*} \left(j_\nu^{\mathrm{ff},HeIII}\ V_{\mathrm{HeIII}} + j_\nu^{\mathrm{ff},HeII}\ V_{\mathrm{HeII}}\right). \tag{2.19}$$

Here He III and He II label the emission coefficients and volumes computed for the He III and He II *regions*; in each case the free-free emission is dominated by H II.

Emission from collisional excitation of H I. Collisions between a free electron and trace H I atom sometimes excite the H I atom, which then radiates away the excitation energy. The collisional

excitation rate coefficient for transitions from the $n = 1$ state to state u, q_{1u}, is

$$q_{1u} = \frac{8.629 \times 10^{-6}}{T^{1/2}} \frac{\Omega(1,u)}{\omega_1} \exp\left(\frac{-\chi(1,u)}{k\,T}\right) \ \mathrm{s}^{-1}\,\mathrm{cm}^3, \tag{2.20}$$

where T is in K, $\Omega(1,u)$ is the (temperature-dependent) effective collision strength for transitions from the $n = 1$ to state u, ω_1 is the statistical weight of the $n = 1$ state and $\chi(1,u)$ is the energy difference between the $n = 1$ and state u (Osterbrock, 1989).

We compute the collisional excitation to the $n = 2$ and $n = 3$ states using cross-sections from Callaway (1985) and Callaway et al. (1987). These excitations result in additional Lyα emission: excitations to the $2p$, $3s$ and $3d$ states radiatively decay to $n = 1$ via Lyα photons, and atoms in the $2s$ state (resulting from collisions to either the $2s$ or $3p$ state) may be additionally collisionally excited to the $2p$ state. The rate of collisional Lyα emission per unit stellar mass, $q^{\rm coll}_{\rm Ly\alpha}$, is

$$q^{\rm coll}_{\rm Ly\alpha} = n_{\rm e} n_{\rm HI}\ q^{\rm eff}_{1,2p} \left(\frac{V}{M_*}\right), \tag{2.21}$$

where

$$q^{\rm eff}_{1,2p} = \sum_u L(u)\, q_{1u} \tag{2.22}$$

and

$$L(u) = \begin{cases} 1 & \text{if } u \in (2p, 3s, 3d) \\ 0.33 & \text{if } u \in (2s, 3p) \end{cases} . \tag{2.23}$$

The factor of 0.33 accounts for $2s$ to $2p$ collisional excitation (Osterbrock, 1989). Summing over the He II and He III regions, $q^{\rm coll}_{\rm Ly\alpha} = 3.4 \times 10^{47}\ \mathrm{s}^{-1}\,\mathrm{M}_\odot^{-1}$.

Total spectrum. For the $f_{\rm esc} = 0$ case, the total specific luminosity per unit stellar mass emitted from a Pop III star and nebula and scattered in the IGM, $l^0_\nu(z)$, has three components that we treat: the truncated stellar spectrum, $l^{*\rm a}_\nu$, the free-free spectrum, $l^{\rm ff}_\nu$, and the scattered Lyα spectrum, $l^{\rm Ly\alpha}_\nu(z)$:

$$l^0_\nu(z) = l^{*\rm a}_\nu + l^{\rm ff}_\nu + l^{\rm Ly\alpha}_\nu(z). \tag{2.24}$$

The truncated stellar spectrum is

$$l_\nu^{*\mathrm{a}} = \begin{cases} l_\nu^* & \text{if } h\nu < 13.6 \text{ eV} \\ 0 & \text{if } h\nu \geq 13.6 \text{ eV} \end{cases} . \tag{2.25}$$

The $l_\nu^{*\mathrm{a}}$ spectrum is slightly modified between 912 and 1216 Å by scattering in the IGM, when the photons in that range are cosmologically redshifted into the Lyα resonance (e.g., Peebles, 1993). This effect is small when $1.75\,\nu_*(z) \ll 8.22 \times 10^{14}$ Hz, i.e., the width of the scattered Lyα line (Loeb & Rybicki, 1999) is small compared to the frequency difference between 912 Å and 1216 Å. Since $1.75\,\nu_*(z = 30) = 8.27 \times 10^{13}$ Hz, we ignore this correction to the spectrum.

The free-free spectrum is given by eq. (2.19). The Lyα spectrum is

$$l_\nu^{\mathrm{Ly}\alpha}(z) = q_{\mathrm{Ly}\alpha}\; h\nu_{\mathrm{Ly}\alpha}\; \phi(\nu_{\mathrm{Ly}\alpha} - \nu, z), \tag{2.26}$$

where $q_{\mathrm{Ly}\alpha}$ is the rate of Lyα photons produced per solar mass of the ionizing star. That rate is

$$q_{\mathrm{Ly}\alpha} = 0.75\,(q_{\mathrm{HI}} - q_{\mathrm{HeII}} + 1.7\,q_{\mathrm{HeII}}) + q_{\mathrm{Ly}\alpha}^{\mathrm{coll}}, \tag{2.27}$$

where $q_{\mathrm{HI}} \equiv Q_{\mathrm{HI}}/M_*$ and $q_{\mathrm{HeII}} \equiv Q_{\mathrm{HeII}}/M_*$. The factor 0.75 represents the fraction of hydrogen recombinations that result in Lyα photons, and the factor of 1.7 accounts for the number of hydrogen ionization per He II recombination, computed above.

Figure 2.3 show $l_\nu^{*\mathrm{a}}$, l_ν^{ff}, $l_\nu^{\mathrm{Ly}\alpha}(z)$ and $l_\nu^0(z)$ for $z = 15$. In this model no Pop III ionizing photons escape to the IGM, thus Pop III stars don't contribute to reionization of the Universe (cf. section 2.4.3).

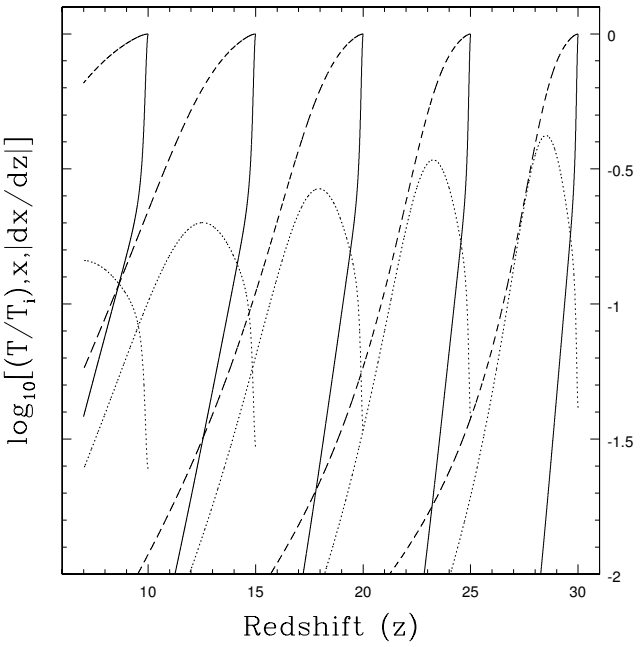


Figure 2.4 Properties of the ionized IGM near a Pop III star. There are sets of three curves for five values of $z_{\rm i} = (10, 15, 20, 25, 30)$. In each set of curves, the solid line is the temperature as a fraction of the initial temperature, the dashed line is the ionization fraction and the dotted line is $|{\rm d}x/{\rm d}z|$, followed from $z_{\rm i}$ to $z = 7$.

2.4.3 Complete escape of ionizing radiation into the IGM

2.4.3.1 Properties of the IGM

The second case we consider is that the nebula plays no role in reprocessing ionizing radiation from a Pop III star, $f_{\rm esc} = 1$. This may be because, in contrast to the assumptions we made in §2.4.2.1, the density of the nebula is low, or because the nebula is clumped into high-density regions with a small covering fraction, or because the nebula was blown away by the star(s). Because the timescales of important IGM processes extend beyond the lifetime of a Pop III star, in this section we will name the redshift of formation of a Pop III star $z_{\rm i}$, and then describe the evolution as a function of z.

We assume the IGM is uniform with baryon density $n_{\rm IGM}(z) = 1.7 \times 10^{-7}(1+z)^3$ cm^{-3} (which ignores the small fraction of baryons in collapsed haloes), $X = 0.75$ and $Y = 0.25$. Ionizing photons from a Pop III star stream into the IGM and form an ionized region. Because the density is low, for $z_{\rm i} \lesssim 30$ recombinations are of little importance on the timescale of the star's lifetime, $\tau \simeq 2 \times 10^6$ yr. For the purposes of calculating the properties of the ionized region of the IGM, we assume that all

of the ionizations occur immediately; at worst this contributes less than a 3 per cent error to the computed spectrum.

He III recombines more quickly than H II, and He I recombinations occur on a comparable timescale to those of H I. As discussed in §2.4.2.2, He II recombinations produce 1.7 H I-ionizing photons on average. In the low-density limit, a He I recombination ionizes 1 hydrogen atom (Osterbrock, 1989). We treat He I and He II ionizations as an extra 1 and 1.7 H I ionizations, respectively, and track only hydrogen recombinations.

For $z_{\rm i} \lesssim 30$, a Pop III star ionizes a volume of the IGM, $V(z)$, initially given by

$$V(z_{\rm i}) = \frac{(Q_{\rm HI} + 0.7\,Q_{\rm HeII})\,\tau}{n_{\rm IGM}(z)} = \frac{2.4 \times 10^4}{(1+z_{\rm i})^3} \left(\frac{M_*}{\rm M_\odot}\right) {\rm kpc}^3. \tag{2.28}$$

The initial temperature in this volume is determined from photoionization heating to be $T_{\rm i} \simeq 6 \times 10^4$ K. Given $z_{\rm i}$, we solve the temperature evolution of the ionized IGM region as a function of z until $z = 7$, when we assume that the Universe reionizes. If Pop III stars reionize the Universe before $z = 7$, their ionized nebulae will have overlapped, and the analysis of this section will no longer be appropriate. The post-reionization contribution of the IGM to the CIRB depends on the temperature and ionization state of the IGM, but is expected to be small under typical assumptions. But the most important effect that reionization would have on the Pop III contribution the CIRB may be the effect reionization has on the SFR. As noted in §2.3.5, reionization may stop or severely curtail the formation of Pop III stars.

To compute the temperature of the ionized IGM around a Pop III star as a function of redshift, we set the initial temperature to $T_{\rm i} = 6 \times 10^4$ K and follow the evolution of temperature considering cooling from Compton scattering of CMBR photons off of electrons, adiabatic cooling of the IGM from the expansion of the Universe, cooling from collisional excitation of H I, free-free cooling and recombination cooling (Cen, 1992). Since all of those processes except adiabatic cooling depend on

the ionization fraction, x, it is solved for simultaneously, using the approximation $x \simeq n_{\rm e}/n_{\rm IGM}$:

$$-\frac{{\rm d}x}{{\rm d}t} = \alpha_{\rm B}({\rm H\,I}, T) n_{\rm IGM} x^2 \tag{2.29}$$

(e.g., Peebles, 1993). Figure 2.4 plots $T(z)$, $x(z)$ and $|{\rm d}x(z)/{\rm d}z|$ for five values of $z_{\rm i}$.

Initially nearly all the hydrogen is ionized, and Compton cooling briefly dominates the thermodynamics. Once a small fraction, $\sim 10^{-4}$, of the hydrogen atoms have recombined, cooling via collisional excitation becomes the dominant process. As the temperature decreases, the recombination rate increases, which in turn increases the neutral fraction and thus the collisional cooling rate. When the temperature reaches about 1.5×10^4 K, Compton cooling dominates again. For lower values of $z_{\rm i}$, adiabatic cooling eventually dominates; for higher values of $z_{\rm i}$, recombination cooling becomes important. Free-free emission is important enough to be included.

2.4.3.2 Resulting spectrum

Recombination emission. As in §2.4.2.2, we consider only the Lyα component of the recombination spectrum. The rate of production of Lyα photons from recombinations, $q^{\rm rec}_{{\rm Ly}\alpha}$, is

$$q^{\rm rec}_{{\rm Ly}\alpha} = 0.63\,(q_{\rm HI} - q_{\rm HeII} + 1.7\,q_{\rm HeII})\,\tau \left|\frac{{\rm d}x}{{\rm d}t}\right|, \tag{2.30}$$

where 0.63 is a representative value for the fraction of hydrogen recombinations resulting in Lyα emission near the peak of the recombination rate.

Free-free emission. The free-free spectrum is

$$j^{\rm ff}_\nu = 7.2 \times 10^{-39}\, x^2\, n^2_{\rm IGM} \left(\frac{T}{\rm K}\right)^{-1/2}$$

$$\times \exp\left(\frac{-h\,\nu}{k\,T}\right)\ {\rm erg\ s^{-1} cm^3\,Hz^{-1} ster^{-1}} \tag{2.31}$$

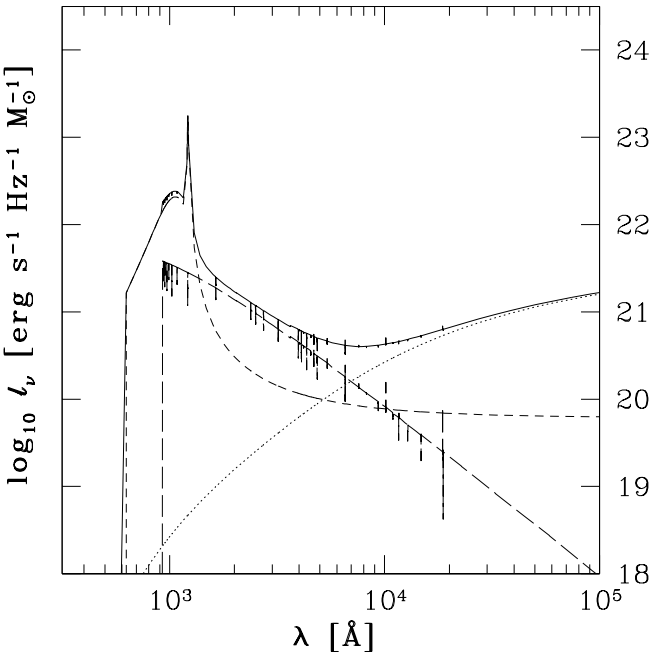


Figure 2.5 Spectrum of a $z = 15$ Pop III star plus emission from the IGM, for the $f_{\rm esc} = 1$ case of section 2.4.3. The long-dashed line is the spectrum of the star, cut off for absorption shortward of the Lyman limit. The dotted line is the spectrum of free-free emission from the IGM. The short-dashed line is emission from Lyα recombination in the IGM, corrected for scattering. The free-free and Lyα spectra are integrated from $z = 15$ to $z = 7$, and divided by the lifetime of a Pop III star to give a useful normalization. The solid line is the sum of the spectra. The sharp peak at 1216 Å is collisionally excited Lyα emission; the broad peak at 1000 Å is Lyα from recombinations. All spectra are in the rest frame of the star (see Section 2.4.3.2 for explanation).

(Ferland, 1980). The specific luminosity per unit stellar mass is

$$l_\nu^{\rm ff}(z, z_{\rm i}) = 4\pi\, j_\nu^{\rm ff}\, \frac{V(z)}{M_*} = 1.6 \times 10^{18}\, x^2\, (1+z)^3 \left(\frac{T}{\rm K}\right)^{-1/2}$$

$$\times \exp\left(\frac{-h\,\nu}{k\,T}\right)\ {\rm erg\ s^{-1}\,Hz^{-1}M_\odot^{-1}}, \tag{2.32}$$

where the second equality follows from equation (2.28). The $z_{\rm i}$ dependence of $l_\nu^{\rm ff}$ results from the implicit dependence of T and x on $z_{\rm i}$ and z.

Emission from collisional excitation of H I. The rate of Lyα photons produced by collisional excitations is given by eqs. (2.21) and (2.28):

$$q_{{\rm Ly}\alpha}^{\rm coll} = 2.0 \times 10^{55}\ x\,(1-x)\ (1+z)^3\ q_{1,2p}^{\rm eff}\ {\rm cm^{-3}\ M_\odot^{-1}}. \tag{2.33}$$

In the low density limit, though, collisions from the $2s$ to $2p$ state are unimportant, so

$$L(u) = \begin{cases} 1 & \text{if } u \in (2p, 3s, 3d) \\ 0 & \text{if } u \in (2s, 3p) \end{cases} . \tag{2.34}$$

Additionally, because the temperature in the IGM varies, we make the following fits to the data of Callaway et al. (1987):

$$\Omega(1,2p) \simeq -2.41 \times 10^{-3} \left(\frac{T}{10^4}\right)^2 + 0.148 \left(\frac{T}{10^4}\right) + 0.170 \tag{2.35}$$

and

$$\Omega(1,3s) + \Omega(1,3d)$$

$$\simeq -2.29 \times 10^{-3} \left(\frac{T}{10^4}\right)^2 + 0.0299 \left(\frac{T}{10^4}\right) + 0.116. \tag{2.36}$$

The fits are good to better than 5 per cent for temperatures from 1.6×10^4 to 6×10^4 K, where almost all of the collisional excitation occurs.

Total spectrum. For the $f_{\rm esc} = 1$ case, the total specific luminosity per unit stellar mass emitted from a Pop III star and IGM, including scattering, is $l_\nu^0(z)$. We treat the same three components as in the $f_{\rm esc} = 0$ case: the truncated stellar spectrum, $l_\nu^{*\rm a}$, the free-free spectrum, $l_\nu^{\rm ff}(\nu, z, z_{\rm i})$, and the scattered Ly$\alpha$ spectrum, $l_\nu^{\rm Ly\alpha}(\nu, z, z_{\rm i})$. The truncated stellar spectrum is given by eq. (2.25), and the free-free spectrum is given by eq. (2.32). Similar to eq. (2.26), the Lyα spectrum is

$$l_\nu^{\rm Ly\alpha}(z, z_{\rm i}) = q_{\rm Ly\alpha}\ h\nu_{\rm Ly\alpha}\ \phi(\nu_{\rm Ly\alpha} - \nu, z), \tag{2.37}$$

where

$$q_{\rm Ly\alpha}(z, z_{\rm i}) = q_{\rm Ly\alpha}^{\rm rec} + q_{\rm Ly\alpha}^{\rm coll}. \tag{2.38}$$

The total spectrum emitted by a Pop III star and the surrounding IGM from the formation of the star at $z_{\rm i}$ to the redshift at which the Universe is reionized, $z = z_{\rm reion}$, expressed in the rest frame of the star when it forms and normalized to the lifetime of the star, is

$$l_\nu^1(\nu, z_{\rm i}) = l_\nu^{*\rm a}(\nu)$$

$$+\frac{1}{\tau}\int_{z_{\rm reion}}^{z_{\rm i}} {\rm d}z\ \left|\frac{{\rm d}t}{{\rm d}z}\right| \left[l_\nu^{\rm ff}(\nu', z, z_{\rm i}) + l_\nu^{\rm Ly\alpha}(\nu', z, z_{\rm i})\right], \tag{2.39}$$

with

$$\nu' = \nu\left(\frac{1+z}{1+z_{\rm i}}\right). \tag{2.40}$$

Because this spectrum is in the frame of the star when it forms, IGM emission at later times (and thus lower values of z) appears in the spectrum at *higher* frequencies than it is emitted at; this is because we *blueshift* all of the IGM radiation back to redshift $z_{\rm i}$. The spectra are normalized by integrating the emission from $z_{\rm i}$ to $z = 7$ and dividing by the lifetime of the star, τ. The resulting spectrum isn't meaningful for the observation of a single Pop III stellar source plus ionized IGM region, because the photons are emitted at different times, but it is a useful spectrum for computing the CIRB. The spectrum of a single source at redshift $z < z_{\rm i}$ (such that the star no longer contributes

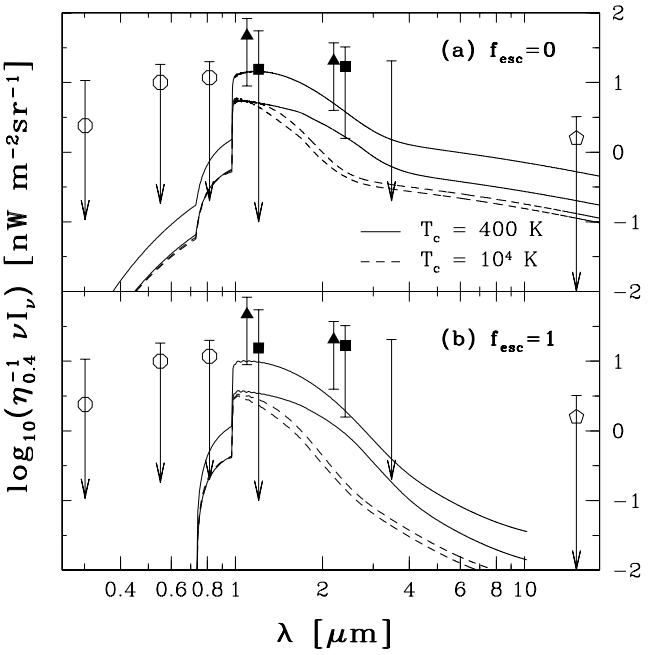


Figure 2.6 The cosmic infrared background from Pop III stars. The ordinate is observed frequency multiplied by observed specific intensity. Shown are curves for star formation in haloes with virial temperatures above critical temperatures $T_{\rm crit} = 400$ K (molecular-hydrogen cooling, solid lines) and 10^4 K (atomic-hydrogen cooling, dashed lines). (a) $f_{\rm esc} = 0$. (b) $f_{\rm esc} = 1$. For both $f_{\rm esc}$ cases, there are two sets of two curves: the upper set is for the ongoing star-formation model; the lower set is for the single-burst model. The properties of the curves are explained in §4.1; this figure shows $z_{\rm end} = 7$. The points show the excess CIRB, with 2σ errors, and are described in §4.2.

directly to the spectrum) is

$$l_\nu^1(\nu, z, z_{\rm i}) = l_\nu^{\rm ff}(\nu, z, z_{\rm i}) + l_\nu^{{\rm Ly}\alpha}(\nu, z, z_{\rm i}); \tag{2.41}$$

this spectrum is in the frame of the emission at redshift z.

Figure 2.5 shows $l_\nu^{*{\rm a}}$, $l_\nu^{\rm ff}(z_{\rm i})$, $l_\nu^{{\rm Ly}\alpha}(z_{\rm i})$ and $l_\nu^1(z_{\rm i})$ for $z_{\rm i} = 15$. The sharp emission peak at 1216 Å results from collisionally-excited Lyα emission shortly after the star forms. The broad peak centered near 1000 Å is Lyα emission from recombinations. At 608 Å the Lyα spectrum is cut off as a consequence of integrating the IGM spectrum to $z = 7$; the cut-off location is 1216 Å $(1+z_{\rm reion})/(1+z_{\rm i})$.

2.5 Cosmic infrared background

2.5.1 Model results

We now convolve the processed spectra of Section 2.4 with the star-formation histories of Section 2.3 to calculate the Pop III contribution to the CIRB. The background is evaluated with

$$\nu_{\rm obs} I_\nu = c \int_0^\infty {\rm d}z \left| \frac{{\rm d}t}{{\rm d}z} \right| \frac{\nu(z) j_\nu^{\rm c}(z)}{1+z} \tag{2.42}$$

(e.g., Peebles, 1993). Here $\nu_{\rm obs}$ is the observed frequency, I_ν is the observed specific intensity, $\nu(z) = (1+z)\nu_{\rm obs}$ and $j_\nu^{\rm c}(z)$ is the comoving specific emission coefficient. Both the star-formation history and the assumed recombination history determine $j_\nu^{\rm c}(z)$:

$$j_\nu^{\rm c}(z) = \frac{1}{4\pi} l_\nu \tau \psi(z), \tag{2.43}$$

where l_ν is either l_ν^0 or l_ν^1, $\tau = 2 \times 10^6$ yr is the fiducial main-sequence lifetime of a Pop III star (Bromm et al., 2001a) and $\psi(z)$ is either $\psi_{\rm on}(z)$ or $\psi_{\rm burst}(z)$, the SFR per comoving volume.

Results for $f_{\rm esc} = 0$ and 1 are shown in Fig. 2.6. There are a total of eight curves in the figure, generated by varying each of three parameters over two values: $f_{\rm esc} = 0$ or 1; $T_{\rm crit} = 400$ or 10^4 K; star-formation mode of ongoing or single-burst. Models with $f_{\rm esc} = 0$ always produce more CIRB at all wavelengths than $f_{\rm esc} = 1$ models. Models with $T_{\rm crit} = 400$ K produce more CIRB than $T_{\rm crit} = 10^4$ K models at all wavelengths except near 1 μm for single-burst star-formation models. The ongoing star-formation models produce more CIRB at all wavelengths than single-burst star-formation models, except near 1 μm for the $f_{\rm esc} = 0$ case.

The sharp edge at 1 μm in all of the curves is a result of our sharp truncation of Pop III star formation at $z_{\rm end} = 7$. The edge occurs at the redshifted wavelength of Lyα from stars at $z_{\rm end}$, i.e., $(1 + z_{\rm end})1216$ Å. The reason that the CIRB curves for $T_{\rm crit} = 400$ and 10^4 K have similar values at 1 μm for single-burst star-formation models is that the corresponding SFRs at $z = 7$ are similar (see Fig. 2.1).

The "bump" in the CIRB curves from 0.7 μm to 1 μm results from stellar emission between rest-frame 912 and 1216 Å (see Figs. 2.3 and 2.5) by sources at $z = 7$. From $\lambda = 1$ μm to roughly 2 μm the Pop III CIRB is dominated by Lyα emission from sources at $1 + z = \lambda/1216$ Å. The inflection points of the curves, particularly clear in the $f_{\rm esc} = 0$ panel, occur at the wavelength where the Pop III CIRB transitions from being dominated by Lyα emission to being dominated by continuum emission from sources at $z = z_{\rm end}$. The continuum radiation for the $f_{\rm esc} = 0$ spectrum is dominated by free-free emission (see Fig. 2.3), whereas stellar continuum is important as well for the $f_{\rm esc} = 1$ continuum spectrum (see Fig. 2.5). Consequently, the Pop III mid-IR background is determined predominantly by the choice of $z_{\rm end}$ rather than the SFR at extremely high redshifts. The Pop III contribution to the optical background (for $z_{\rm end} \geq 7$) is due to the stellar "bump" described above and the high-energy tail of free-free emission, which is only significant for the $f_{\rm esc} = 0$ spectrum; this is a consequence of the temperature of the gas, which is effectively much higher in the $f_{\rm esc} = 0$ case than for the $f_{\rm esc} = 1$ case (see §3).

2.5.2 Observational data

The points with error bars in Fig. 2.6 show the difference between the total extragalactic background and the extragalactic background due to resolved sources. At wavelengths, e.g., 0.3 μm, where the lower limit is an arrow, the contribution from resolved sources is sufficient to explain *all* of the background, at the 2σ level. At wavelengths where data points have lower limits plotted, e.g., 2.2 μm, the light from resolved sources is not enough to account for the measured background. It is this unexplained excess, measured by Cambrésy et al. (2001) at 1.25 μm and by Cambrésy et al. (2001) and Wright & Johnson (2001) at 2.2 μm, that we are trying to fit with our models, while conforming to upper limits at other wavelenths.

Near-IR data points The triangles at 1.25 and 2.2 μm (J and K band, respectively) correspond to the unexplained CIRB as measured by Cambrésy et al. (2001): they measured the total extragalactic background with *COBE*/DIRBE data, and subtracted from that the contribution due to resolved galaxies measured by near-IR surveys. The squares at J and K (offset slightly in

wavelength for clarity) are our computation of the unexplained CIRB from the CIRB measurement of Wright & Johnson (2001), using their measurement of the total extragalactic background from *COBE*/DIRBE data (but using a different foreground model from Cambrésy et al. 2001), and the same galaxy contribution subtraction and error propagation as Cambrésy et al. (2001). The Wright & Johnson (2001) point at J band is consistent at the 2σ level with no unexplained extragalactic background. All error bars are 2σ. The 2σ upper limit at 3.5 μm (L band) represents a measurement of the total extragalactic background at that wavelength (Wright & Johnson, 2001). *SIRTF* is expected to accurately measure the contribution of galaxies to the L-band CIRB, and thus allow a determination of the excess L-band CIRB (W. T. Reach, private communication).

Other data points The upper limits at 0.3, 0.6, and 0.8 μm are 2σ upper limits on the unexplained optical extragalactic background light (Bernstein et al., 2002). The open circle data points show the estimated unexplained optical extragalactic background, which is consistent with zero (2σ) at all three bandpasses. Bernstein et al. (2002) measured the total optical extragalactic background from *HST* images (utilising simultaneous ground-based spectra for absolute zodiacal-light subtraction calibration). From published number counts and their own careful photometry of the Hubble Deep Field observations, they find no evidence for convergence in the integrated light from galaxies; this suggests that deeper galaxy photometry will lower the upper limits.

The upper limit at 15 μm is our computation of the 2σ upper limit on the unexplained mid-IR background light, using the CIRB upper limits obtained by Renault et al. (2001) from gamma-ray observations, and the *ISO*/ISOCAM 15 μm galaxy counts reported by Elbaz et al. (2002). The open hexagonal data point shows the estimated excess CIRB, which is consistent with zero (2σ).

2.6 Discussion

Our models identify a very narrow range of parameter space in which Pop III stars may explain the excess CIRB:

1. The Pop III stars must be very massive, so that most of their energy is radiated in photons energetic enough to ionize hydrogen.

2. Cooling in low-mass haloes is possible due to H_2.

3. Approximately 40 per cent of the eligible baryons in the star forming halo must be converted into Pop III stars.

4. The escape fraction of ionizing photons from the nebula surrounding a Pop III star must be near zero, so that most of the ionizing photons are converted into Lyα photons.

5. Pop III star formation must begin by $z \simeq 25$ and persist until $z_{\rm end} \simeq 7$, so that the Lyα emission (Fig. 2.6) extends through the observed J and K bands.

6. Negative feedback effects must not inhibit Pop III star formation. In particular, we require that Pop III stars do not radiatively or mechanically destroy star-forming material, and that they they do not enrich their surroundings with a sufficient amount of metals to end Pop III star formation.

The amplitudes of the curves in Fig. 2.6 scale simply with η, and they do not change dramatically with small changes to the cosmological parameters. The characteristic breaks in the spectra are located at 912 Å$(1 + z_{\rm end})$ and 1216 Å$(1 + z_{\rm end})$, where $z_{\rm end}$ is the low-redshift limit to Pop III star formation; if $z_{\rm end}$ were much higher than 7, the Pop III CIRB peak would lie longward of the J band.

If all of the above conditions are met, then *all* of the observed near-IR CIRB deficit is due to Pop III star formation. Future observations may demonstrate that other sources contribute significantly to the near-IR background; e.g., part of the deficit could be due to the faint wings of galaxies that are unaccounted for in current surveys (Totani et al., 2001). The parameter space explored in this paper easily accomodates lower values of the unexplained CIRB, through decreasing the efficiency, η. The shape of the Pop III CIRB spectrum in the near-IR is primarily determined by the shape of the star-formation rate as a function of redshift.

The formation of very massive stars probably requires that the star-forming gas is not enriched with heavy elements to a mass fraction $Z_{\rm crit} \gtrsim 10^{-3} Z_{\odot}$, where $Z_{\odot}$ is the solar value (Bromm et al., 2001b). We require a large fraction of the mass processed through Pop III stars to end up in massive black holes, so as not to overproduce metals compared to the observed levels in the high redshift IGM and to avoid enriching the primordial gas too quickly to $Z_{\rm crit}$. Recently, Schneider et al. (2002b) have pointed out the possible problem that if *all* Pop III stars collapsed into massive black holes, the IGM would always have zero metallicity. It seems natural to assume, however, that a small fraction of the Pop III stellar mass, $\Omega_Z = Z\Omega_{\rm B} = \epsilon_{\rm pi}\Omega_{\rm III}$ is ejected into the IGM through pair-instability supernovae, even if the majority of the mass is permanently locked up in black holes. Assuming a ratio $\Omega_{\rm III}/\Omega_{\rm B} \sim 0.1$, the required pair-instability fraction to produce the critical level of metallicity $Z_{\rm crit}$ would be $\epsilon_{\rm pi} \sim 10^{-2}$ (see also Oh et al., 2001).

We thus conclude that if Pop III stars explain a large fraction of the near-IR background, then almost all Pop III material must end up in massive black holes. In that case we would predict a significant contribution, ~ 10 per cent, by massive black holes to the total baryonic mass budget in the Universe. Although such a prediction may seem somewhat extreme, it will ultimately be tested by observations (e.g., Agol & Kamionkowski, 2002; Schneider et al., 2002b). We emphasize that at present there are no observations that rule out such a scenario (e.g., Carr, 1994). Moreover, there are good physical reasons to seriously consider the existence of massive black holes resulting from Pop III star formation (e.g., Bromm et al., 2002; Madau & Rees, 2001; Schneider et al., 2002b).

In §2.3.4 we discussed $\Omega_{\rm III}$, the cumulative matter density processed through Pop III stars. In Fig. 2.2 the horizontal lines labelled H I, He I and He II correspond to the $\Omega_{\rm III}$ required to produce 10 ionizing photons per particle of each species (Bromm et al., 2001a). For the $f_{\rm esc} = 1$ model, the intersection of those lines with the $\Omega_{\rm III}$ star curves is expected to be closely related to the redshift of reionization. For the $f_{\rm esc} = 0$ model, though, Pop III ionizing photons are all absorbed in the haloes in which they are emitted, and thus Pop III stars make no contribution to the reionization of the Universe. We discussed above that if Pop III stars enriched their immediate environment with a sufficient amount of metals, Pop III stars would no longer form there (producing a SFR similar to

our single-burst model). If metals from Pop III stars efficiently mixed throughout the Universe, the era of Pop III stars might therefore come to an end (see Bromm et al., 2001b).

A single Pop III star produces $\sim 10^{37}$ erg s^{-1} M$_\odot^{-1}$ in the Lyα line, assuming $f_{\rm esc} = 0$. At $z = 7$ that would result in a flux of $\sim 10^{-23}$ erg s^{-1} cm^{-2} M$_\odot^{-1}$. For a discussion of the size of the Lyα emitting region, see Loeb & Rybicki (1999). For the $f_{\rm esc} = 1$ case the luminosity and size of the emitting region of a Pop III star depend in more detail on redshift and IGM parameters.

Future theoretical work should improve our understanding of the efficiency of Pop III star formation, as well as the physical conditions of the haloes in which Pop III stars form. We finally expect the launch of *NGST* to open an observational window into the earliest epochs of star formation, possibly including Pop III stars, less than a decade from now.

Acknowledgments

VB thanks the TAPIR group at Caltech for its hospitality during the completion of this work. We thank L. Cambrésy, W. T. Reach, Y. Lithwick, A. Benson, J. Silk, A. Ferrara, M. Rees and P. Natarajan for stimulating discussions. We thank the anonymous referee for comments that improved the presentation of this paper. MRS acknowledges the support of NASA GSRP grant NGT5-50339. This work was supported in part at Caltech by NSF AST-0096023, NASA NAG5-8506 and DoE DE-FG03-92-ER40701 and DE-FG03-88-ER40397.

Chapter 3

Probing Reionization with Lyman α Emission Lines[1]

3.1 Abstract

Lyman α emission from high-redshift galaxies may be a powerful probe of the ionization history of the IGM at $z > 6$: the observed Lyman α emission line is sensitive to the neutral fraction of IGM hydrogen in the range 0.1–1. We present calculations of observed Lyman α emission lines from $z > 6$ galaxies, illustrating the effect of varying the many free parameters associated with the emitting galaxy, its halo, and the IGM around the galaxy. In particular, we use a dynamic model of the IGM that includes the effect of IGM infall toward the emitting galaxy. Galactic winds may play a crucial role in determining observed Lyman α line fluxes. We compare our model predictions with observations of two $z = 6.5$ galaxies and conclude that, if galactic winds are allowed for, existing observations place no constraint on the neutral fraction of the IGM at $z = 6.5$. Future space-based observations will constrain the importance of galactic winds; if winds are unimportant for the observed $z = 6.5$ galaxies, our models suggest that the IGM neutral fraction at $z = 6.5$ is $\lesssim 0.1$.

[1]This chapter appeared as Santos (2004a), and is reproduced here by permission of the copyright holder, the Royal Astronomical Society.

3.2 Introduction

Strong Lyman α emission is present in many distant galaxies, and is an important signpost for discovering high-redshift galaxies and, especially, measuring redshifts for them. The presence of strong Lyman α emission has been crucial in confirming the highest redshift galaxies known at $z > 5$ (e.g., Hu et al., 1999, 2002a,b; Ajiki et al., 2003; Kodaira et al., 2003; Rhoads et al., 2003). See Taniguchi et al. (2003) for a review. It also plays an important role in identifying the redshift of some galaxies at $z \gtrsim 2$ because its strength, even in some apparently extremely dusty galaxies, allows for redshift determinations when little continuum light is visible (e.g., Steidel et al., 2003; Chapman et al., 2003).

In this paper we concentrate on Lyman α emission powered by star formation. Hot stars emit photons capable of ionizing hydrogen. If the ionized gas density is sufficient, hydrogen will recombine on relevant timescales. The recombination cascade produces a rich spectrum of lines, but usually includes a Lyman α photon (see Osterbrock, 1989, for a review). Because most recombinations produce Lyman α photons, it is predicted to be the strongest line emitted from recombination emission. AGN activity can also provide ionizing photons that ultimately generate a strong Lyman α line. Additional sources of Lyman α photons are atomic-hydrogen cooling of gas (e.g., Haiman et al., 2000; Fardal et al., 2001) and fluorescence of gas clouds illuminated by a strong, non-local source of ionization (e.g., Reuland et al., 2003).

The predicted strength of the Lyman α line, and the assumed prevalence of star formation at early times (inferred from the abundance of old stars in the local universe) led to an early prediction of an abundant population of high-redshift galaxies with very strong Lyman α emission lines (Partridge & Peebles, 1967). Though subsequent revisions to the theory of galaxy formation, primarily that it was a hierarchical process, explained the relative lack of Lyman α detections compared to those early predictions (Haiman & Spaans, 1999), searching for galaxies based solely on strong Lyman α emission has been validated recently as an effective technique for identifying galaxies up to $z = 6.5$.

Lyman α is a resonant transition with a large cross section. As a consequence, small quantities of neutral hydrogen scatter away Lyman α photons in direction and frequency. In particular, if

the intergalactic medium (IGM) is not highly ionized, as is expected for at least some of the time between recombination ($z \sim 1100$) and $z \simeq 6$ (Fan et al., 2002; Kogut et al., 2003), then the neutral hydrogen in the IGM can easily scatter the Lyman α line from galaxies. The Lyman α photons are not destroyed, but the scattering process diffuses the line in areal coverage and frequency (Loeb & Rybicki, 1999), and the resulting low surface brightness emission is currently unobservable.

A simple consideration of the IGM shows that the blue side of the emergent Lyman α line would be scattered away by a neutral IGM, and only the red side of the line would be observed. This follows from considering that most of the scattering by neutral hydrogen comes at wavelengths near the Lyman α transition in the rest-frame of the gas. However, for a completely neutral IGM at high redshifts, the Lyman α cross section far from resonance becomes important, since the natural line profile has 'damping wings' that fall off only like $\nu^{\pm 2}$. Thus even if the IGM has no component at the velocity of corresponding to resonant scattering of a particular Lyman α photon, that photon may still be scattered. This observation led to the conclusion that observations of Lyman α emission lines (or their absence) at high redshifts may probe the ionization state of the IGM in a regime poorly tested by the Gunn-Peterson trough measurements, $x_{\rm HI} \sim 1$ (Miralda-Escude & Rees, 1998; Miralda-Escude, 1998; Haiman, 2002; Barkana & Loeb, 2004; Cen, 2003a).

In particular, since the *WMAP* satellite discovered a high optical depth to the last scattering surface (Kogut et al., 2003), the ionization history of the universe at $z > 6$ is of great interest. Though the *WMAP* results indicate that the universe was mostly ionized for a substantial history of the universe at $z > 6$, it does not constrain the exact ionization level or the history in detail. There are now many predictions for the ionization history at $z > 6$ based on semi-analytic and numerical modelling of early star formation (e.g., Cen, 2003b; Wyithe & Loeb, 2003b; Ciardi et al., 2003; Haiman & Holder, 2003). If there is a population of strong Lyman α emitters at the redshifts in question, $6 \lesssim z \lesssim 20$, and if the effect of IGM scattering on their observed properties can be calculated, then future surveys for Lyman α emission may provide valuable information on the reionization history at $z > 6$.

Here we use a more sophisticated treatment of the dynamics of the IGM around galaxies with

strong intrinsic Lyman α emission to investigate the effect of the IGM on the observed line. The IGM model we use will be presented in more detail separately, but is summarized here and nearly identical to that described in Barkana (2004). The key feature is that infall of the IGM toward the galaxies is included. Barkana & Loeb (2003) have showed that such an IGM model makes a prediction for an absorption feature in the spectra of high-redshift QSO Lyman α lines. Here we consider normal galaxies, where the effect of an infalling IGM is even stronger.

Throughout this paper we assume a cosmology based on the recent results from *WMAP* and other work (Spergel et al., 2003, and references therein): $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km s^{-1}, and $\Omega_{\rm b} = 0.0469$. In general we quote distances in physical Mpc, denoted by 'pMpc.'

This paper is organized as follows. In Section 3.3 we present the Lyman α scattering cross section. Section 3.4 describes the model for the intrinsic Lyman α line profile and the observed profile modified by IGM scattering. Section 3.5 presents Lyman α line profiles as a function of the (many) input parameters that influence the observed line profile. We introduce the possible role of galactic winds in Section 3.6. Section 3.7 discusses our Lyman α flux predictions with comparison to current observations, and also reviews model assumptions. Section 3.8 summarizes.

3.3 Lyman α scattering

A Lyman α photon is emitted when a hydrogen atom makes a transition from the $n = 2$ level to the $n = 1$ ground-state. The emitted photon has an energy of 10.199 eV and a wavelength of $\lambda_\alpha = 1215.67$ Å. The transition is the strongest spontaneous transition of a hydrogen atom, with an Einstein A value of $A_{21} = 6.265 \times 10^8$ s^{-1}. Correspondingly, the absorption cross section of the transition is relatively large as well.

A Lyman α photon travelling through a region of neutral hydrogen will be repeatedly absorbed and re-emitted. This process by itself does not destroy Lyman α photons, but if photons are scattered over a sufficiently large region such that an observation spatially resolves the emission, then the Lyman α line will be difficult to observe in practice (Loeb & Rybicki, 1999). In this work we assume that Lyman α photons scattered in the IGM, that is, outside of the virial radius of the

emitting galaxy, are missed by observation. For a $z \gtrsim 6$ galaxy, the virial radius is larger than $1''$ (the size of a typical spectroscopic slit) for any galaxy more massive than $\sim 3 \times 10^9$ $M_\odot$.

So under the assumption that photons emitted in the Lyman α emission line of a galaxy are lost from subsequent observation if they are scattered by the IGM, we constructed a model of the IGM (see Section 3.4.2.1) and computed the absorption due to neutral hydrogen. The natural absorption cross section $\sigma_{\rm N}$ is (Peebles, 1993)

$$\sigma_{\rm N}(\nu) = \frac{3\lambda_\alpha^2 A_{21}^2}{8\pi} \frac{(\nu/\nu_\alpha)^4}{4\pi^2(\nu-\nu_\alpha)^2 + (A_{21}^2/4)(\nu/\nu_\alpha)^6}, \tag{3.1}$$

where $\nu = c/\lambda$ and $\nu_\alpha = c/\lambda_\alpha$.

The total absorption cross section due to a parcel of IGM gas is determined by the kinetic properties of the gas, both bulk and thermal, as well as the natural cross section. Bulk motions just introduce a frequency shift, but thermal motions broaden the profile, modelled as the convolution of the natural cross section with a Maxwellian velocity distribution,

$$\sigma_{\rm V}(\nu) = \int_{-\infty}^{\infty} M(v)\sigma_{\rm N}(\nu - \nu_\alpha v/c){\rm d}v, \tag{3.2}$$

$$M(v) = \left(\frac{m_{\rm H}}{2\pi kT}\right)^{1/2} \exp\left(-\frac{m_{\rm H}v^2}{2kT}\right), \tag{3.3}$$

where $m_{\rm H}$ is the mass of the hydrogen atom and T is the temperature of the IGM gas (modelled in the next section). For the usual Lorentzian approximation of $\sigma_{\rm N}$, the convolved cross section would be a Voigt profile, hence we label our convolved cross section $\sigma_{\rm V}$. Figure 3.1 shows $\sigma_{\rm V}(\lambda)$, assuming a temperature of $T = 10^4$ K for the scattering gas. The profile has a Doppler core with a half-width of ~ 10 km s^{-1} and power-law tails called 'damping wings.' If the IGM at the redshift of interest is or has previously been reionized, then $T = 10^4$ K is probably a good approximation. If the IGM is still thermally pristine, then the temperature may be as low as ~ 20 K, resulting in a narrower Doppler core of the scattering profile. In practice this would have a small effect on any near-future observation.

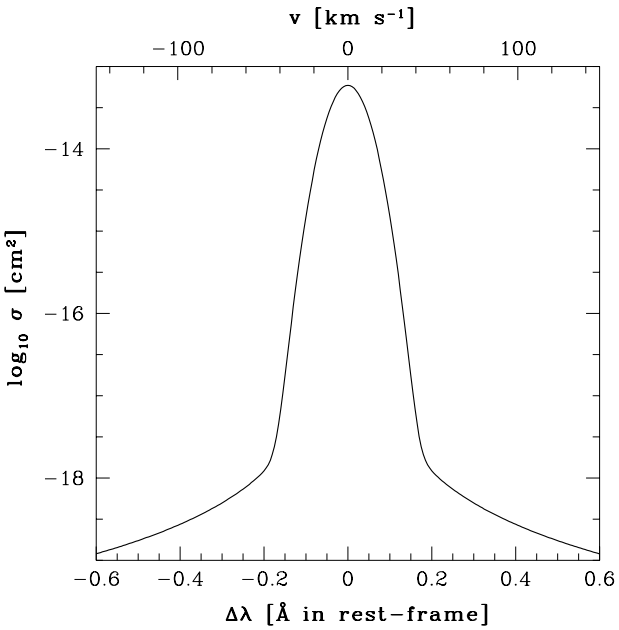


Figure 3.1 Cross section for Lyman α absorption, as a function of wavelength difference from resonance in the frame of the absorbing atom. The top axis shows the velocity (in the frame of the absorbing atom) where the Lyman α resonance corresponds to the wavelength on the bottom axis.

3.4 Lyman α line profiles

We will refer to two types of Lyman α lines from galaxies. The **intrinsic line** is the line produced in the galaxy, specifically, the line that reaches the virial radius of the halo containing the emitting galaxy. The **observed line** is the line after propagation though the IGM to $z = 0$.

3.4.1 Intrinsic Lyman α emission line

A Lyman α emission line from a high-redshift galaxy is typically a consequence of hydrogen recombinations in nebular gas. The ionization of the H II region may be maintained either by hot stars or by gas accretion onto a black hole. Though luminous QSOs have been discovered at $z > 6$ (Fan et al., 2003, and references therein), deep x-ray data indicates that most galaxies at those redshifts do not contain powerful AGNs (Barger et al., 2003). Thus in this work we will assume recent ($\lesssim 10$ Myr ago) star formation powers Lyman α emission lines.

Type O stars emit hydrogen-ionizing photons.[2] Some fraction, $f_{\rm esc}$, of these photons may escape

[2]Approximately 1 Gyr after the onset of star formation hot white dwarfs form, and these may emit substantial ionizing radiation (Charlot & Fall, 2000); however, for $z \gtrsim 5.5$ the universe is not old enough to have formed these

the galaxy without absorption. The rest of the photons are absorbed by either neutral hydrogen or dust within the galaxy. An ionizing photon absorbed by a hydrogen atom initiates a cascade of recombination emission, ultimately resulting in a Lyman α photon about 2/3 of the time, and two-photon emission the other 1/3 of the time (Osterbrock, 1989). A Lyman α photon subsequently scatters through absorption and reemission until the optical depth of the galaxy at its wavelength is $\lesssim 1$; the optical depth is determined both by the position and direction of the photon, and also its wavelength. If dust is present, it may absorb photons at any time in this process.

The radiative transfer problem for Lyman α depends sensitively on the geometry of neutral hydrogen and dust within the galaxy (Neufeld, 1991), and also the dynamics of the hydrogen (Kunth et al., 1998). For simplicity we will assume that the intrinsic Lyman α emission line is centered at the systemic redshift of the emitting galaxy, with a Doppler profile shape described by a characteristic velocity, $v_{\rm D}$,

$$\phi_\nu^{\rm in}(\nu) = \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}} \exp\left[\frac{-(\nu-\nu_\alpha)^2}{(\Delta\nu_{\rm D})^2}\right], \tag{3.4}$$

where $\phi_\nu^{\rm in}$ is the line profile per unit frequency and

$$\Delta\nu_{\rm D} = \frac{v_{\rm D}}{c}\nu_\alpha. \tag{3.5}$$

In Section 3.6 we address other possible intrinsic line profiles.

In this simplified model, two parameters characterize an intrinsic Lyman α emission line: its luminosity, L_α, and $v_{\rm D}$. L_α depends on the rate of production of ionizing photons, $\dot{Q}$, the ionizing photon escape fraction, $f_{\rm esc}$, and dust absorption. The production of ionizing photons in turn depends on the star-formation rate and the metallicity and initial mass function (IMF) of the stars formed. None of these quantities may be reliably predicted from first principles for a galaxy at high redshift. Therefore, any use of the *observed* Lyman α line of a high-redshift galaxy to deduce the radiative transfer properties of the IGM will be complicated, unless these quantities can be measured independently of the Lyman α line.

stars.

The estimation of $v_{\rm D}$ is simpler. The likely minimum velocity scale for motion in a galaxy is the circular velocity of the parent halo, $v_{\rm c}$; this velocity may describe the gas motion is the star-formation is mediated by a violent process, such as the merging of two equal-mass galaxies. If star-formation is a more quiescent process in the galaxy, it may occur in a disk whose peak velocity is, roughly, between $v_{\rm c}$ and $2v_{\rm c}$, for realistic halo and disk properties (Mo et al., 1998; Cole et al., 2000). Thus we assume $v_{\rm c} \lesssim v_{\rm D} \lesssim 2v_{\rm c}$.

3.4.2 Observed Lyman α emission line

The optical depth for photons to be absorbed in the IGM is

$$\tau(\nu) = \int_{R_{\rm vir}}^{R_{\rm obs}} x_{\rm HI}(r) n_{\rm H}(r) \sigma_{\rm V}(\nu') {\rm d}r, \tag{3.6}$$

$$\nu' = \nu \left[1 + \frac{v(r)}{c} \right], \tag{3.7}$$

where $R_{\rm obs}$ is the physical distance to the observer (though there is almost no contribution from $r \gg R_{\rm vir}$ because the Hubble flow eventually redshifts the IGM Lyman α resonance far from the emitted Lyman α line [in the frame of the galaxy]). We describe our model for $n_{\rm H}(r)$, $v(r)$ and $x_{\rm HI}(r)$ in Section 3.4.2.1.

The observed Lyman α emission line, in the rest-frame of the emitting galaxy, is then

$$\phi_\nu^{\rm obs}(\nu) = \phi_\nu^{\rm in}(\nu) \exp\left[-\tau(\nu)\right]. \tag{3.8}$$

3.4.2.1 IGM model

Density and velocity We will present a full description of our model for the density and velocity of IGM material in another paper (Santos & Adelberger, in preparation). The model is very similar to one described in Barkana (2004). Here we summarize our model as it applies to this paper.

The evolution of the density and velocity of the IGM follow from the small, linear, initial perturbations imprinted on the universe at very high redshift. We model the IGM starting from the

statistical description of linear perturbations by applying a simple description of the non-linear evolution of the perturbations.

Specifically, we start by considering a Lyman α emitting galaxy in a halo of total mass M at redshift z. We then constrain the initial linear overdensity on the scale of M to produce the halo at redshift z. Starting from that initial linear overdensity constraint, we use the excursion set formalism (Bond et al., 1991; Lacey & Cole, 1993) to compute the mean[3] initial linear overdensities, averaged in spheres centered on the location of the halo, on scales larger than the scale of M. The mean initial linear overdensity decreases monotonically with increasing scale; this is not externally imposed. Our approach also implies that the halo is not substructure of a larger collapsed halo (see also Barkana 2004).

Once we have the initial linear overdensities of spheres enclosing the halo, we evolve those overdensities using the spherical top-hat model (Partridge & Peebles, 1967). Because the mean initial linear overdensity decreases monotonically with increasing scale, there is no shell crossing, and the evolution of each IGM matter shell is independent of the evolution of the other shells.

Our dynamical evolution treats all matter equally, and we assume the baryons trace the dark matter in these shells. This description would be inappropriate if, e.g., the Lyman α emitting galaxy blows a strong wind into its surrounding IGM (see Adelberger et al., 2003, for evidence of this at $z = 3$), and we will return to this in Section 3.6.

A realistic distribution of initial linear overdensities around a halo, in contrast to the mean initial linear overdensities averaged over many halos, would not in general be spherically symmetric or even monotonic with radius within spherically averaged shells. In addition to collapsing toward the galaxy, there would be structures collapsed and collapsing within the 'IGM.' The most straightforward way to assess the effect of these complicated dynamics would be with numerical simulation. Here we appeal to analytic results to demonstrate the effect of structure with the collapsing IGM will not severely affect our conclusions.

Scannapieco & Barkana (2002) computed the bivariate probability distributions of two halos

[3]i.e., averaged over a large ensemble of identical halos

forming at different positions. Their results suggest that outside the virial radius of a halo, the typical enhancement in the number density of halos is, at maximum, only a factor of a few, for any mass. Thus we conclude that the fraction of matter collapsed into halos (above a given mass) in the IGM surrounding our galaxy of interest is at most a few times the universal collapse fraction. At $z = 6$ up to half the mass of the IGM may be in the form of collapsed halos with virial temperatures above 10^4 K, that is, halos massive enough to collisionally ionize their hydrogen. That fraction decreases with increasing radius. The enhancement is also a decreasing function of decreasing mass. Thus we conservatively expect that our simple picture gives at most a factor of two overestimate of the optical depth, averaged over a reasonable sample of halos. At large optical depths, this error doesn't affect the observed Lyman α line properties, and at low optical depths we may underestimate the observed line by up to that same factor of two.

The practical output of our model is, given a halo mass M at redshift z, the radius, velocity, and mass of shells of IGM matter surrounding the galaxy from its virial radius to a large radius.

Ionization state To calculate the scattering of a Lyman α emission line by the IGM, we need to know the ionization state of the hydrogen, in addition to its density and velocity profile. We solve the ionization balance of the IGM gas assuming two possible contributions, direct ionization by the galaxy emitting the Lyman α line and, if the universe is ionized, a mean ionizing background.

After calculating the density distribution around a halo using the prescription described in the previous section, we solve the ionization balance as a function of distance from the galaxy. In general, the recombination rate of ionized hydrogen in the IGM is relatively small. As a consequence, in the pre-reionized universe H II regions around galaxies are not in equilibrium, but began growing when star-formation turned on and expand at the rate that ionizing photons can ionize more neutral IGM (rather than balancing recombinations within the ionized region). The H II region is still expected to have a relatively sharp boundary, as the mean free path of ionizing photons in the neutral IGM is small. After reionization, all of the IGM is highly ionized by the mean ionizing background, but the direct ionizing flux of the galaxy provides an additional proximity effect on the ionization balance

immediately around it.

We divide the IGM around into one or two regions: either the universe is reionized, in which case all of the IGM is ionized; or the universe is not reionized, in which case there is an H II region in the IGM immediately surrounding the galaxy, and the IGM is fully neutral outside of that region. In that case the radius $R_{\rm S}$ of the H II region is calculated by solving

$$\int_{R_{\rm vir}}^{R_{\rm S}} \frac{4\pi r^2 n_{\rm H}(r)}{\dot{Q} - \Lambda(r)} {\rm d}r = t_{\rm SF}, \tag{3.9}$$

$$n_{\rm H}(r) = \frac{\left[\delta(r)+1\right]\rho_{\rm H,0}(1+z)^3}{m_{\rm H}}, \tag{3.10}$$

$$\dot{Q} = \int_{\nu_0}^{\infty} \frac{L_\nu}{h\nu} {\rm d}\nu, \tag{3.11}$$

$$\Lambda(r) = \int_{R_{\rm vir}}^{r} 4\pi \tilde{r}^2 n_{\rm H}^2(\tilde{r}) \alpha_{\rm B}(T) {\rm d}\tilde{r}, \tag{3.12}$$

where $R_{\rm vir}$ is the virial radius of the galaxy's halo, $t_{\rm SF}$ is the age of the star-formation activity[4], $\delta(r)$ is the overdensity with respect to the mean IGM density (calculated with the model in the previous section), $\rho_{\rm H,0}$ is the comoving (i.e., $z = 0$) density of hydrogen, z is the redshift of the galaxy, $h\nu_0$ is the ionization energy of hydrogen, L_ν is the specific luminosity of the galaxy, and $\alpha_{\rm B}(T)$ is the hydrogen Case B recombination coefficient. Since the H II region is almost completely ionized, here we assumed the electron physical density, $n_{\rm e}$, is $n_{\rm e} = n_{\rm H}$, and ignored the contribution to $n_{\rm e}$ from helium. We have not included a clumping factor modification to the recombination rate; however, we do account for the mean overdensity of the IGM matter. Additional clumping of this material would decrease the size of the true H II region compared to our calculation.

The ionization rate for a hydrogen atom within an ionized region a distance r from the galaxy is

$$\Gamma(r) = \int_{\nu_0}^{\infty} \left(\frac{L_\nu}{4\pi r^2} + 4\pi J_\nu \right) \frac{\sigma_{\rm V}(\nu)}{h\nu} {\rm d}\nu, \tag{3.13}$$

where J_ν is the mean specific intensity of the intergalactic background radiation. If the universe is

[4]We model star formation with a constant rate for an age $t_{\rm SF}$.

not reionized, then we set $J_\nu = 0$. The hydrogen neutral fraction at any radius, $x_{\rm HI}(r)$, is solved for by setting $\Gamma(r)$ equal to the recombination rate at that radius,

$$x_{\rm HI}(r) = 1 + \frac{\chi}{2} - \frac{1}{2}\left(\chi^2 + 4\chi\right)^{1/2} \tag{3.14}$$

$$\chi = \frac{\Gamma(r)}{n_{\rm H}(r)\alpha_{\rm A}(T)}, \tag{3.15}$$

where $\alpha_{\rm A}(T)$ is the hydrogen Case A (total) recombination coefficient. Additional clumping of the IGM would decrease χ linearly, increasing $x_{\rm HI}$.

We assume that when IGM gas falls across the virial radius it shocks and ionizes completely, and has no impact on the Lyman α line profile.

3.5 Lyman α line results

As described in Section 3.4.1, the strength and shape of the intrinsic Lyman α emission line depend on many properties of the source galaxy. Additionally, the properties of the observed line are strongly modified by scattering in the IGM, which depends on the properties of the galaxy, its halo, and the IGM model. To illustrate these dependencies, we will adopt a fiducial model for a Lyman α emitting galaxy, then show the dependence of the intrinsic and observed Lyman α line on the properties of that galaxy and the IGM.

3.5.1 Fiducial galaxy/halo model

Our fiducial Lyman α emitter model is a galaxy at rest in the center of a 10^{11} $M_\odot$ halo at $z = 6.5$. 'Fiducial galaxy' will refer to this total system, galaxy plus halo. We assign the fiducial galaxy a constant star-formation rate (SFR) of 10 $M_\odot\,{\rm yr}^{-1}$ for an age of $t_{\rm SF} = 10^8$ yr. The virial radius of the halo is $R_{\rm vir} = 19.8$ kpc, and the circular velocity at the virial radius is $v_{\rm c} = 148$ km s^{-1}. The halo dynamical time is then 1.3×10^8 yr. Given our cosmology, and assuming the universal baryon to dark matter ratio holds in the halo, the halo contains 1.5×10^{10} $M_\odot$ of baryons. Thus at

$10\ \mathrm{M}_\odot\,\mathrm{yr}^{-1}$ the gas supply could last up to 1.5×10^9 yr, about twice the age of the universe at that redshift. The chosen SFR also implies that 9 per cent of the baryons are converted into stars per halo dynamical time.

We use the following prescription to convert the fiducial galaxy properties into the emitted Lyman α line. We convert the SFR to a hydrogen-ionizing photon luminosity using

$$\dot{Q} = 3.5 \times 10^{53}\ \mathrm{s}^{-1} \left(\frac{\mathrm{SFR}}{\mathrm{M}_\odot\,\mathrm{yr}^{-1}} \right), \tag{3.16}$$

appropriate for a Salpeter IMF from 1 to 100 $\mathrm{M}_\odot$ with 1/20th solar metallicity (Leitherer et al., 1999; Schaerer, 2003). We assume that 2/3 of the ionizing photons that do not escape are converted into Lyman α photons (Osterbrock, 1989), so that

$$L_\alpha = \frac{2}{3}(1 - f_{\mathrm{esc}})\dot{Q}h\nu_\alpha. \tag{3.17}$$

Note that this assumes there is no absorption by dust; we return to the issue of dust in Section 3.7.3. The other Lyman α line parameter that needs to be specified is the line width, characterized by $v_\mathrm{D} \equiv f_v v_\mathrm{c}$. For our fiducial model, we set $f_v = 1.5$.

Additionally we need to model the effect of photoionization of the galaxy on the surrounding IGM. We assume the escaping photon luminosity of ionizing photons is simply

$$\dot{Q}_{\mathrm{esc}} = f_{\mathrm{esc}}\dot{Q}. \tag{3.18}$$

Our baseline choice for the escape fraction is $f_{\mathrm{esc}} = 0.1$, giving

$$L_\alpha = 3.4 \times 10^{42}\ \mathrm{erg\ s}^{-1} \left(\frac{\mathrm{SFR}}{\mathrm{M}_\odot\,\mathrm{yr}^{-1}} \right). \tag{3.19}$$

3.5.2 IGM model dependence

Now we illustrate the IGM scattering of the Lyman α line emitted by our fiducial galaxy, highlighting the importance of using our dynamic IGM model. At $z = 6.5$, the ionization state of the IGM is unknown. Limits from QSO spectra at redshifts approaching 6.5 indicate that $x_{\rm HI} > 10^{-3}$ averaged over the volume of the IGM, and $x_{\rm HI} > 10^{-2}$ averaged over the mass of the IGM (Fan et al., 2002). Simulations of reionization suggest that the $z \sim 6$ transition from $x_{\rm HI} \simeq 1$ to $x_{\rm HI} \simeq 0$ was very sudden (Gnedin, 2000). Combined with the QSO observations, these suggest that $x_{\rm HI}$ could have almost any value at $z = 6.5$. The *WMAP* results constrain the IGM ionization fraction to be ~ 1 over a substantial fraction of the evolution of the universe at $z > 6$, but do not provide detailed constraints on models of $x_{\rm HI}(z)$.

We start by assuming our fiducial galaxy ($z = 6.5$) is embedded in a fully neutral IGM. Haiman (2002) argued that the detection of reasonably strong Lyman α emission from $z = 6.5$ galaxies did not rule out a fully neutral IGM, as had been previously suggested (Hu et al., 2002a; Rhoads et al., 2003). However, the calculations of Haiman (2002) did not account for the infall profile of the IGM around a $z = 6.5$ galaxy.

First we illustrate the difference between our model of the IGM, and a smooth, Hubble flow model. We call our model the **dynamic IGM model** because we allow the IGM to evolve away from the comoving solution. We call the mean-density, comoving model the **simple IGM model**. The simplest difference between the models is the density of the IGM as a function of distance from the center of the halo. Figure 3.2a shows the density and density contrast of the IGM outside the halo virial radius for our dynamic IGM model, and a simple IGM. Within about 10 virial radii, or 0.2 physical Mpc (pMpc), the density enhancement is between 2 and 10 times the mean IGM density. Integrating the density out from the virial radius gives the curves shown in Fig. 3.2b, after converting total density to hydrogen density. Though Fig. 3.2 shows that there is a substantial difference in the density profile at small radius, the increased density as a function of radius is not the most important effect of the dynamic IGM model.

The most important difference between the IGM models, with respect to scattering of Lyman α

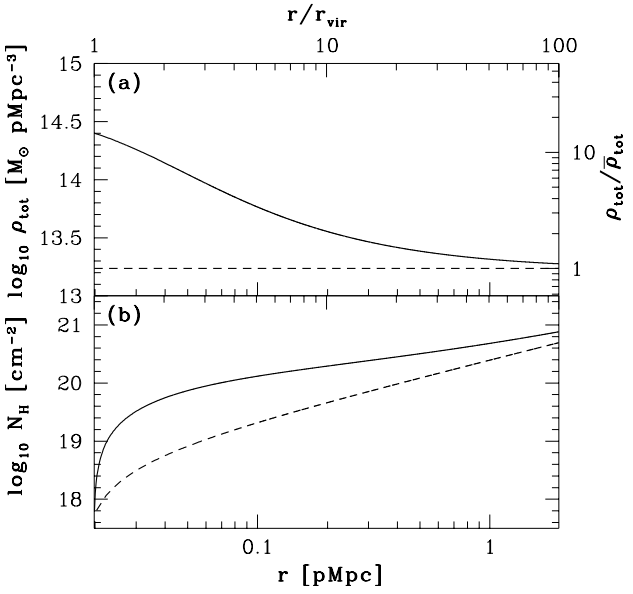


Figure 3.2 IGM density as a function of radius from the fiducial galaxy. The left edges of the plots are at the virial radius. (a) The total IGM mass density and density contrast as a function of radius in physical Mpc. (b) The cumulative IGM hydrogen column density to radius r. In both plots the solid curve is for the dynamic IGM model and the dashed curve is for the simple model.

line photons, is the distribution of density as a function of velocity. Figure 3.3 shows the relationship between radius and velocity, for the dynamic and simple IGM models. With this relation we replot the densities of Fig. 3.2 against velocity in Fig. 3.4. The local Hubble parameter is $H(z = 6.5) =$ 790 km s^{-1} pMpc^{-1}. For the IGM at positive velocities with respect to the galaxy, that is, the blue side of the line, there is little difference in density between the dynamic and simple models. However, while the simple model has no gas at negative velocities, corresponding to the red side of the line, the dynamic model has reached a column density of hydrogen of 10^{20} cm^{-2} by the time the velocity has reached 0. The absorption profile (Fig. 3.1) has a cross section of greater than 5×10^{-14} cm^2 over the central 10 km s^{-1}. Thus the average optical depth within about 100 km s^{-1} of the red side of the line is greater than $5 \times 10^4 x_{\rm HI}$. Immediately we can see that for $x_{\rm HI} \sim 10^{-5}$, a typically expected value for the IGM, optical depths of order unity and greater may be expected. This optical depth on the red side of the observed line due to scattering by highly (but not completely) ionized gas gives rise to a line profile distinctly different from a simple IGM model.

Figure 3.5 presents a schematic illustration of the difference between the dynamic and simple

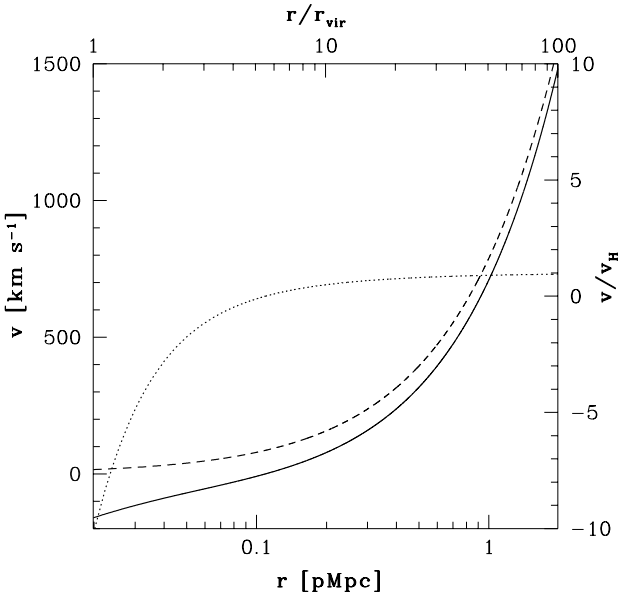


Figure 3.3 Velocity of IGM gas as a function of radius from the fiducial galaxy. The solid curve is the dynamic IGM model; the dashed curve is the Hubble flow relation of the simple model. The dotted curve is the ratio of the dynamic model to the simple model, in units marked on the right axis.

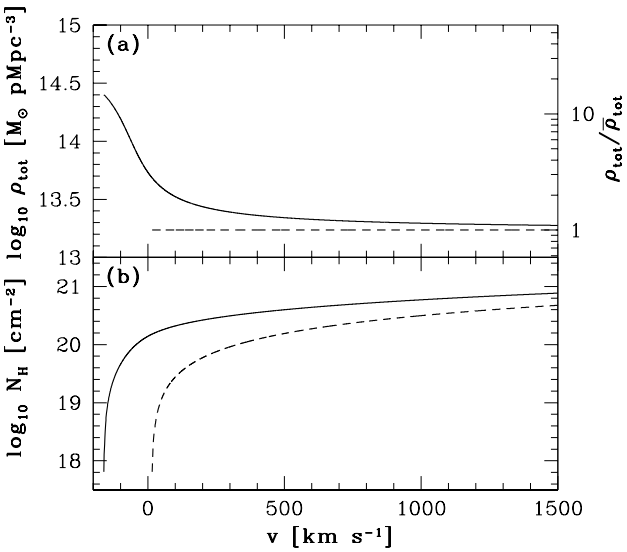


Figure 3.4 IGM density as a function of velocity in the frame of the fiducial galaxy. (a) The total IGM mass density and density contrast as a function of velocity. (b) The cumulative IGM hydrogen column density to velocity v. In both plots the solid curve is for the dynamic IGM model and the dashed curve is for the simple model.

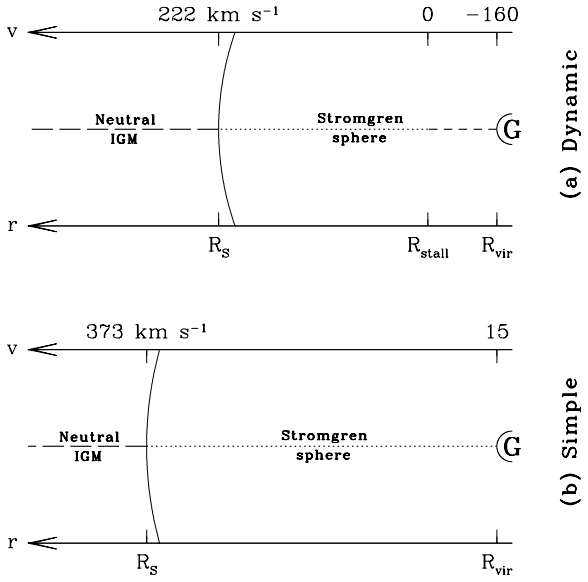


Figure 3.5 Schematic diagram of location and velocity of the IGM around the fiducial galaxy. (a) Dynamic IGM model. (b) Simple model. The galaxy is marked by a 'G' at right. The curve immediately surrounding it shows the virial radius of the halo. The next curve, at $R_{\rm S}$, shows the Stromgren sphere. The velocity axis marks the velocities (in km s^{-1}) corresponding to these radii. In (a) $R_{\rm stall}$, the radius where IGM gas has zero velocity with respect to the galaxy, is also marked. The dashed and dotted lines key the physical location of the IGM material to the contribution it makes to the scattering optical depth, shown in Fig. 3.6.

IGM models. As described above, we see that the dynamic model has gas with negative velocity, in contrast to the strictly positive velocity IGM in the simple model. We also see that the Stromgren radius is both larger and extends to larger IGM velocities. In Fig. 3.5 we have keyed different physical regions with different line types; Fig. 3.6 shows the effect each of these regions has on the observed line profile.

In Fig. 3.6 we show the optical depth as a function of observed wavelength, subdivided by the velocity of the IGM gas responsible for the optical depth, for the fiducial galaxy with the dynamic and simple IGM models. We label the velocity of the IGM at radius r from the galaxy $v(r)$; $v(R_{\rm S}) = 222$ km s^{-1} for the dynamic model, and $v(R_{\rm S}) = 373$ km s^{-1} for the simple model. When examining the observed line profile at a particular wavelength λ, it is useful to define the velocity

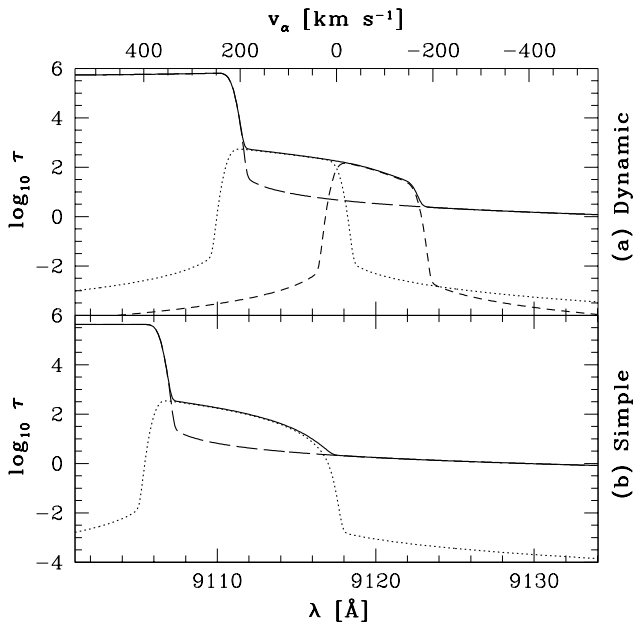


Figure 3.6 Optical depth due to IGM scattering as a function of observed wavelength. (a) Dynamic IGM model. (b) Simple model. The solid curves are the total optical depth due to all IGM H I. The short-dashed curves are the optical depth due to gas with negative velocity (in the frame of the galaxy). The dotted curves are the optical depth due to gas with positive velocity inside the Stromgren sphere. The long-dashed curve is the optical depth due to gas outside the Stromgren sphere.

in the frame of the galaxy where the Lyman α resonance corresponds to that wavelength,

$$v_\alpha = c\left[1 - \frac{\lambda}{(1+z)\lambda_\alpha}\right]. \tag{3.20}$$

The total optical depth has three prominent regions. At $v_\alpha > v(R_{\rm S})$, the optical depth is dominated by Doppler core scattering from neutral IGM at $v \simeq v_\alpha$. At $v(R_{\rm vir}) < v_\alpha < v(R_{\rm S})$ the optical depth is still due to Doppler core scattering, but the IGM at $v \simeq v_\alpha$ is mostly ionized. The third region is at $v_\alpha < v(R_{\rm vir})$, where the optical depth is due to IGM gas at $v > v(R_{\rm S})$ (the same as the first region), but now the scattering results from the damping wing of the scattering cross section. Figure 3.6 illustrates that the central part of the observed line profile depends on the ionization state of the IGM gas in the Stromgren sphere, and the red wing of the line profile depends on the size of the Stromgren sphere and the IGM outside of it.

Figure 3.7 quantitatively compares the IGM density computed for both the dynamic and simple

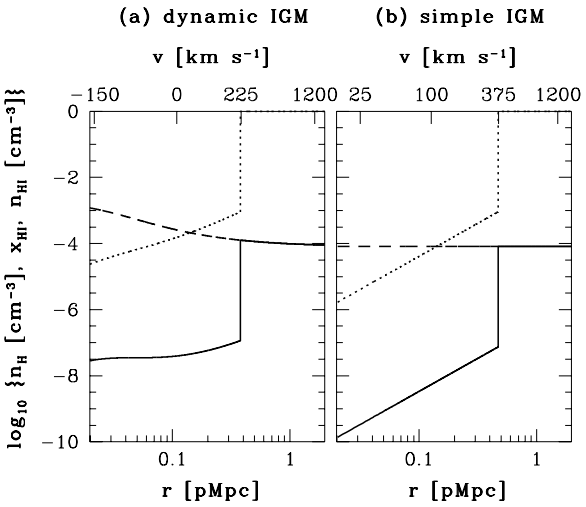


Figure 3.7 Density and ionization state of the IGM as a function of radius and velocity. (a) Dynamic IGM model. (b) Simple model. The dashed curves are $n_{\rm H}$, the total (H I+H II) hydrogen number density; the dotted curves are $x_{\rm HI}$, the neutral hydrogen fraction; and the solid curves are $n_{\rm HI} \equiv x_{\rm HI} n_{\rm H}$. Both plots are for the fiducial galaxy with radius in physical Mpc and velocity in the frame of the galaxy.

IGM models. In the dynamic model, $R_{\rm S} = 0.38$ pMpc, versus $R_{\rm S} = 0.47$ pMpc for the simple model. The larger Stromgren sphere in the simple model is a consequence of the lower IGM density.

Figure 3.8 shows the observed Lyman α line profiles per unit wavelength, ϕ_λ, for the dynamic (solid curve) and simple (dashed curve) IGM models, compared to the intrinsic line profile (dotted curve) of the fiducial galaxy. The line profiles are plotted against observed wavelength and also v_α.

For both the dynamic and simple models, the IGM outside of the virial radius, though highly ionized out to the Stromgren radius, has sufficient optical depth to completely suppress the Lyman α line with resonant scattering from the Doppler core of the absorption cross section. Thus the wavelength range over which the line is completely suppressed corresponds to $v_\alpha \gtrsim v(R_{\rm vir})$, where $v(R_{\rm vir})$ is the velocity of the IGM at the virial radius. For the dynamic model $v(R_{\rm vir}) = -160$ km s^{-1}, in contrast to 15 km s^{-1} for the simple model.

Because IGM gas is assumed to ionize completely as it crosses the virial radius, the line redward of $v(R_{\rm vir})$ is not scattered by the Doppler core of the absorption profile of any IGM gas. However, the observed lines are still depressed compared to the intrinsic profile at $v_\alpha < v(R_{\rm vir})$ due to scattering by the fully neutral IGM outside the Stromgren sphere, via the red damping wing of the absorption

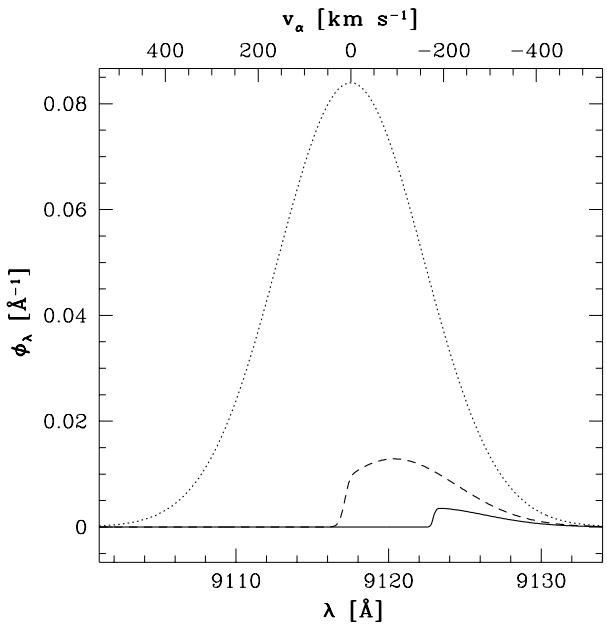


Figure 3.8 Lyman α line profiles. The dotted curve is the intrinsic Lyman α line profile of the fiducial galaxy. The solid curve is the observed Lyman α line profile using the dynamic IGM model. The dashed curve is the observed line using the simple IGM model. The velocity where the Lyman α resonance corresponds to the wavelength on the bottom axis is shown as v_α, on the top axis.

cross section. In summary, the optical depth at some wavelength (and thus v_α) can be thought of as a sum of the optical depth due to Doppler core scattering by IGM gas at $v = v_\alpha$, and optical depth due to damping wing scattering by IGM gas at $v \geq v(R_{\rm S})$. In Fig. 3.8, the difference between the dynamic and simple model predictions at $v_\alpha = -100$ km s^{-1} is primarily a consequence of the presence of neutral gas at $v = -100$ km s^{-1} in the dynamic model. At $v_\alpha = -300$ km s^{-1}, the difference is entirely a consequence of neutral gas between +225 and +375 km s^{-1} (the velocity range over which the IGM is neutral in the dynamic model but ionized in the simple model).

The integrated observed line flux for the dynamic model is only 1.6 per cent of the intrinsic line, compared to 11 per cent for the simple model. Thus incorporating a dynamic model for the IGM is an important consideration in connecting theoretical predictions for Lyman α emitters into observed counts, and for interpreting the ionization state of the IGM through Lyman α emission line observations.

3.5.3 Galaxy/halo parameter dependence

In the previous section we showed that our dynamic IGM model has a substantial impact on the observed Lyman α line from a galaxy at high redshift. In this section we illustrate how the observed line depends on properties of the galaxy and its halo. The parameters we explore are the halo mass, galaxy star-formation rate, age of star formation, escape fraction of ionizing radiation, and relation between the velocity width of the emitted line and the halo circular velocity. All figures are for the dynamic IGM model only; varying the parameters would have analogous effects under the simple IGM model.

For each parameter, we show plots analogous to Fig. 3.7 and Fig. 3.8, as appropriate. From these plots the virial radius and Stromgren radius can be read off, as well as the corresponding IGM infall velocities. The line profile figures show the effect of our model parameters, if any, on the intrinsic line profile, and the effect on the observed line profile.

Halo mass In Fig. 3.9 we recompute the IGM properties for the fiducial model, except that we set the halo mass to 10^{10} or 10^{12} $M_\odot$. Though more massive halos have larger values of $R_{\rm vir}$ (the radii at which the curves terminate), they have more negative values of $v(R_{\rm vir})$. More massive halos also have slightly higher IGM densities around them. The combination of larger virial radii (ionizing flux falls off as r^{-2}) and slightly higher IGM densities leads to larger neutral fractions within the Stromgren spheres of more massive halos. The Stromgren spheres of more massive halos are slightly smaller (due to larger densities within them), but $v(R_{\rm S})$ decreases more dramatically with increasing halo mass.

In Fig. 3.10 we show the corresponding intrinsic and observed Lyman α line profiles. More massive halos have broader intrinsic and observed lines (because $v_{\rm D} \equiv f_v v_{\rm c}$ and $v_{\rm c}$ increases for more massive halos). A new prediction of the dynamic IGM model is that the offset between the intrinsic line center and the blue edge of the observed line (or, similarly, the centroid of the observed line) increases with increasing halo mass. This conclusion is somewhat dependent on the assumption that the intrinsic Lyman α emission line is centered on the systemic velocity of the galaxy. In practice

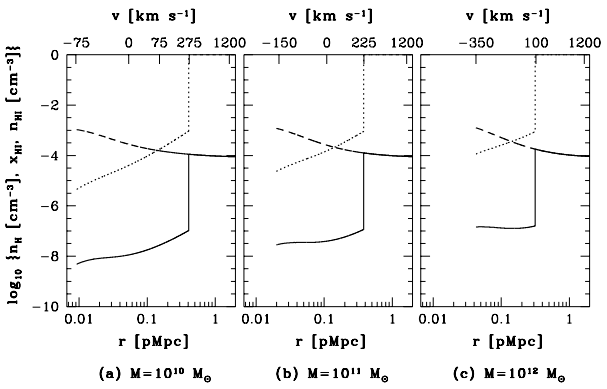


Figure 3.9 As in Fig. 3.7, but note different radial scale. The curves are truncated at $R_{\rm vir}$. (a) Halo mass $M = 10^{10}$ M$_\odot$. (b) $M = 10^{11}$ M$_\odot$ (fiducial model). (c) $M = 10^{12}$ M$_\odot$.

obtaining the intrinsic Lyman α line shape may be impossible, but a more complicated model of the intrinsic line may be passed through our IGM model to produce a predicted observed line profile, which will always have a lower limit in wavelength corresponding to $v_\alpha = v(R_{\rm vir})$.

Star-formation rate In Figs. 3.11 and 3.12 we recompute the IGM properties and Lyman α line profiles for the fiducial model, except that we set the SFR to 1, 100, or 1000 M$_\odot$ yr^{-1}. Increasing the SFR increases both the size of the Stromgren sphere and the ionization fraction inside. If we decrease the SFR from the fiducial value of 10 M$_\odot$ yr^{-1} to 1 M$_\odot$ yr^{-1} then the neutral IGM encroaches all the way to $v(R_{\rm S}) = 25$ km s^{-1} (compared to $v(R_{\rm S}) = 225$ km s^{-1} in the fiducial case), and thus the damping wing almost completely wipes out the observed line. Increasing the SFR to 100 M$_\odot$ yr^{-1}, the Stromgren sphere expands out to $v(R_{\rm S}) = 650$ km s^{-1}. Though there is a decrease in the neutral fraction within the Stromgren sphere, this effect is not sufficient to lift the complete suppression of the observed line at $v_\alpha \gtrsim v(R_{\rm vir})$. However, the greater distance (in velocity) to the fully neutral IGM means that the observed line redward of of $v_\alpha \simeq v(R_{\rm vir})$ is less damped than in the fiducial case. A further increase of the SFR to 1000 M$_\odot$ yr^{-1} expands the Stromgren sphere yet further, resulting in very little damped absorption at $v_\alpha \lesssim v(R_{\rm vir})$. Moreover, the neutral fraction inside part of the Stromgren sphere is now small enough that scattering by the cross section Doppler core no longer completely suppresses the line at $v_\alpha > v(R_{\rm vir})$. Note the notch in the spectrum at $v_\alpha = v(R_{\rm vir})$: this is the feature predicted in QSO spectra by Barkana & Loeb (2003).

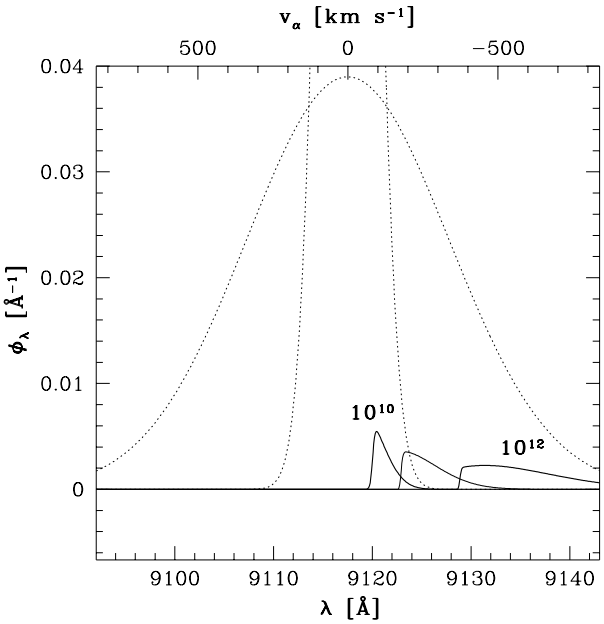


Figure 3.10 As in Fig. 3.8, but note different axis scales. The labels show M (in $M_\odot$) for the corresponding intrinsic and observed line profiles. The fiducial model $M = 10^{11}$ $M_\odot$ is the lighter curve, and the intrinsic fiducial line profile is not shown.

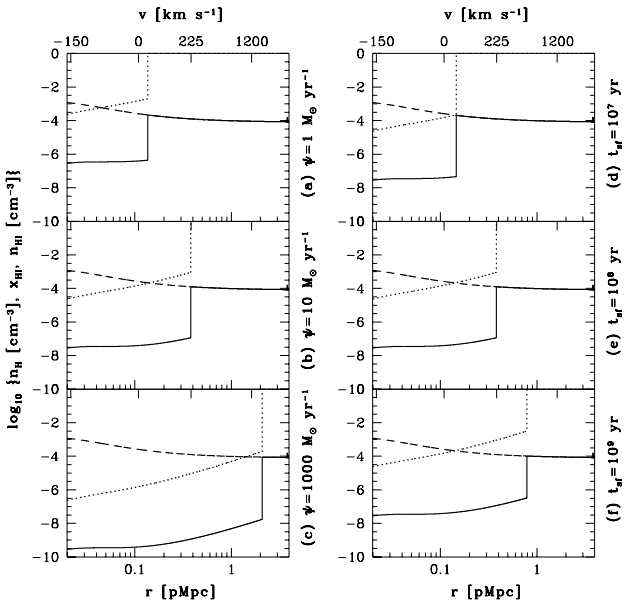


Figure 3.11 As in Fig. 3.7. (a) Star-formation rate $\psi = 1$ $M_\odot\,yr^{-1}$. (b) $\psi = 10$ $M_\odot\,yr^{-1}$ (fiducial model). (c) $\psi = 1000$ $M_\odot\,yr^{-1}$. See Fig. 3.14c for $\psi = 10$ $M_\odot\,yr^{-1}$. (d) Age of star formation $t_{SF} = 10^7$ yr. (e) $t_{SF} = 10^8$ yr (fiducial model). (f) $t_{SF} = 10^9$ yr.

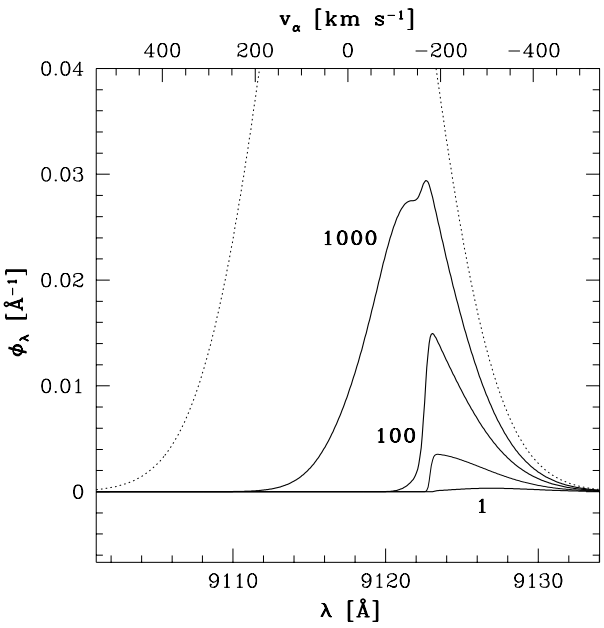


Figure 3.12 As in Fig. 3.8, but different vertical scale. The labels show star-formation rate (in $M_\odot\,yr^{-1}$) for the corresponding observed line profiles. The fiducial model $\psi = 10\ M_\odot\,yr^{-1}$ is the lighter curve, and the intrinsic fiducial line profile is the dotted curve.

Age of star formation In Figs. 3.11 and 3.13 we recompute the IGM and line profile properties for the fiducial model, except that we set the age of star formation to 10^7 or 10^9 yr. Changing $t_{\rm SF}$ changes the size of the Stromgren sphere, but does not change the ionization balance within it. Thus the effect on the observed line profile is only to change the amount of damping of the line at $v_\alpha < v(R_{\rm vir})$.

Escape fraction of ionizing photons In Fig. 3.14 we recompute the IGM properties for the fiducial model, except that we set $f_{\rm esc} = 0$ or 1. Scaling $f_{\rm esc}$ has the same effect as scaling the star-formation rate by the same factor, so the fiducial model modified to $f_{\rm esc} = 1$ is the same as the fiducial model modified to $\psi = 100\ M_\odot\,yr^{-1}$ (see above). In the case of vanishing escape fraction, for any reasonable value of the SFR the line will unobservable, because a neutral IGM down to $v(R_{\rm vir})$ allows only 4×10^{-5} of the intrinsic line to be observed. See also Section 3.7.2.

Velocity width factor In Fig. 3.15 we recompute line profile properties for the fiducial model, except that we set $f_v = 1$ or 2 (f_v does not influence the IGM properties). Larger values of f_v

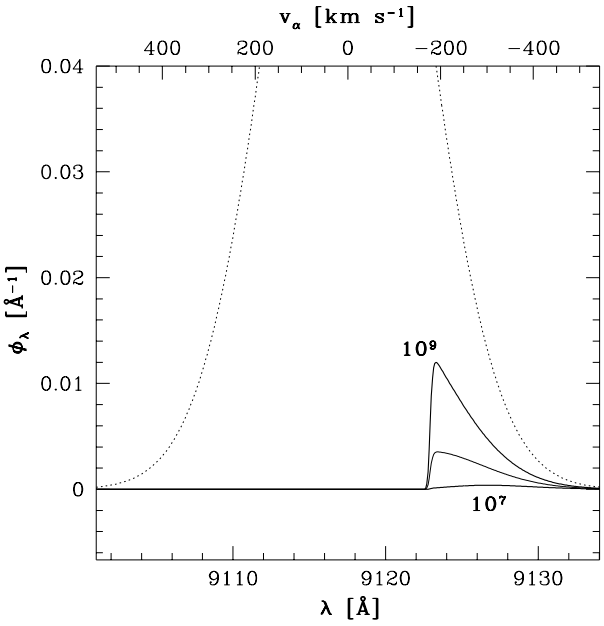


Figure 3.13 As in Fig. 3.8, but different vertical scale. The labels show the age of star-formation (in yr) for the corresponding observed line profiles. The fiducial model $t_{\rm SF} = 10^8$ yr is the lighter curve, and the intrinsic fiducial line profile is the dotted curve.

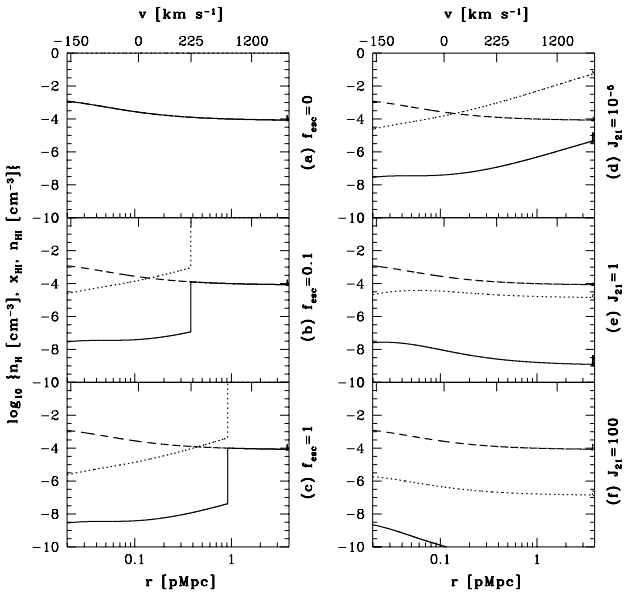


Figure 3.14 As in Fig. 3.7. (a) Ionizing photon escape fraction $f_{\rm esc} = 0$. (b) $f_{\rm esc} = 0.1$ (fiducial model). (c) $f_{\rm esc} = 1$. (d) Ionizing background $J_{21} = 10^{-5}$. (e) $J_{21} = 1$. (f) $J_{21} = 100$.

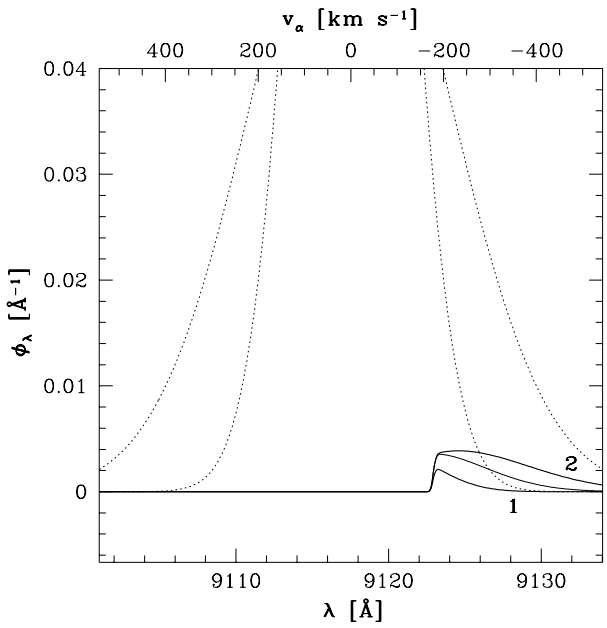


Figure 3.15 As in Fig. 3.8, but different vertical scale. The labels show the conversion factor f_v between Doppler line width and halo circular velocity for the corresponding intrinsic and observed line profiles. The fiducial model $f_v = 1.5$ is the lighter curve, and the intrinsic fiducial line profile is not shown.

broaden the line, resulting in a greater observable fraction of the intrinsic line. In the limit of much larger values of f_v, more of the line would be transmitted, but the line width would be so large that spectroscopic detection of the line would become more difficult than for a narrower line with less flux. However, for the values of f_v we consider, which should bound the actual values, the observed line is still relatively narrow and larger values of f_v produce more easily observed lines.

3.5.4 IGM ionization dependence

In this section we vary the ionization state of the IGM to show how this influences the observed Lyman α line. First we will assume, in contrast to the IGM models considered so far, that there is a mean ionizing background in the IGM that contributes to the ionization of the IGM around the halo, in addition to the contribution from the galaxy itself. We assume that the universe was reionized before star formation turned on in the halo (that is, reionization occurred longer than $t_{\rm SF}$

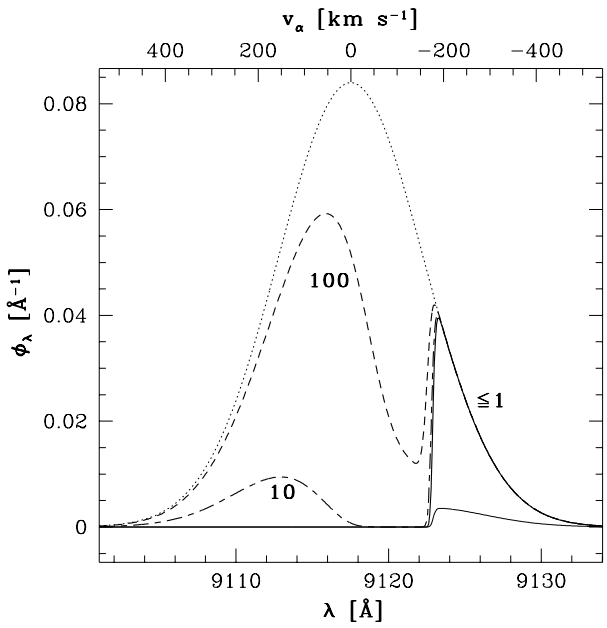


Figure 3.16 As in Fig. 3.8. The labels show ionizing background strength J_{21} for the corresponding observed line profiles. The fiducial model with no ionizing background is the lighter curve, and the intrinsic fiducial line profile is the dotted curve.

ago), and since then the background has been fixed at

$$J_{21} \equiv \frac{J_\nu}{10^{-21}\ \mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}}}. \tag{3.21}$$

For the fiducial values of $z = 6.5$ and $t_{\rm SF} = 10^8$ yr, that corresponds to reionization before $z = 7.2$.

In Figs. 3.14 and 3.16 we show the properties of the IGM and Lyman α line profiles for values of the ionizing background that bracket the physically interesting regimes. First we discuss $J_{21} = 10^{-5}$, a value so small that the IGM ionization state inside the galaxy's light sphere is completely dominated by the galaxy's ionizing radiation. The light sphere has a radius of $c\,t_{\rm SF} = 31$ pMpc for $t_{\rm SF} = 10^8$ yr, corresponding to an IGM velocity of almost 2.5×10^4 km s^{-1}. Consequently, the IGM ionization state outside of the light sphere is irrelevant.

Even though the ionizing background does not contribute at all to the ionization rate (see eq. 3.13) at any radius of interest, the pre-reionization of the universe before the galaxy turned on allows its Stromgren sphere to expand at near light speed. Consequently even the IGM at the velocity of 1200 km s^{-1} has a neutral fraction of 0.01. By the Gunn-Peterson trough test, this IGM wouldn't

even qualify as reionized. However, this level of ionization completely eliminates any damping wing optical depth. Thus the observed line profile follows the intrinsic line profile from the red wing up to $v_\alpha = v(R_{\rm vir})$, then goes to zero as in the fiducial case. This is the observed behavior for any ionizing background with $J_{21} \lesssim 1$, the range observed at $z < 5$ (Scott et al., 2002).

Ionizing backgrounds in the range $J_{21} \sim 10$–100 strongly ionize the mean density IGM. Since the IGM density decreases from a local maximum at $R_{\rm vir}$ toward the cosmic mean as a function of increasing radius, IGM gas just outside of the virial radius still strongly scatters the intrinsic line at v_α corresponding to the velocity of the overdense gas. As the density decreases, though, the neutral fraction drops as well, and the combination leads to greater transmission of the intrinsic line *toward the blue side of the line*. This would generate an unusually-shaped observed line with two peaks, one of which lies to the blue of the intrinsic line center.

At yet higher values of $J_{21} \gtrsim 10^3$, the IGM is so strongly ionized that the entire intrinsic line would be observed. We note again that we expect $J_{21} \lesssim 1$, though we are unaware of any predictions for J_{21} at $z > 6$ in models that reproduce the *WMAP* optical depth value.

Alternately, rather than considering a fixed ionizing background in a reionized universe, we can fix the neutral fraction of the IGM outside of the galaxy's Stromgren sphere, $x_{\rm HI}^{\rm IGM}$. This approach is more useful for considering a weakly-ionized IGM with $x_{\rm HI}^{\rm IGM} \sim 0.5$ (e.g., Cen, 2003b, where a fully reionized universe almost entirely recombines). Figure 3.17 shows the observed Lyman α line profiles of the fiducial galaxy when $x_{\rm HI}^{\rm IGM}$ is set to 0.5 and 0.05. As expected, the decreased H I abundance in the 'neutral' IGM means that the red wing of the Lyman α line is less suppressed by damping wing scattering. There is no effect on the Doppler scattering that scatters the center of the intrinsic line. See also Section 3.7.2.

3.5.5 Redshift dependence

Last, we turn to the effects due solely to changes in redshift. In the future there will (hopefully) be surveys for strong Lyman α emission lines at several different redshifts at $z > 6$. We study redshift dependence by keeping the other free parameters fixed; realistically, other parameters, such as star-

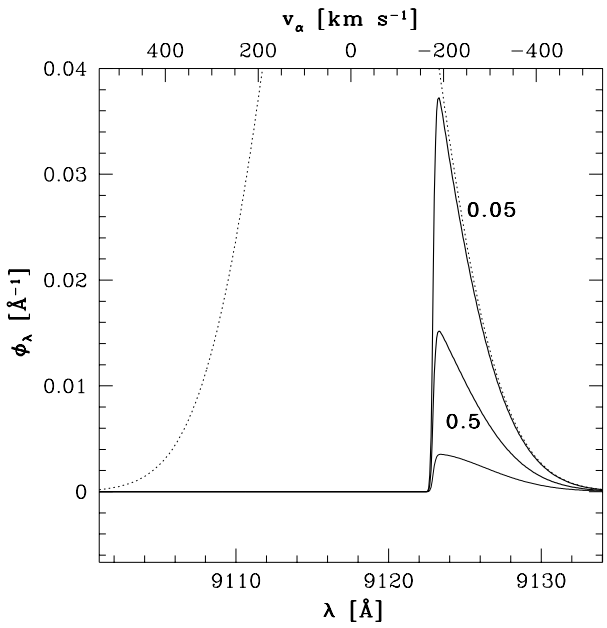


Figure 3.17 As in Fig. 3.8, but with different vertical scale. The labels show the neutral fraction of the IGM outside of the galaxy's Stromgren sphere for the corresponding observed line profiles. The fiducial model with a neutral IGM is the lighter curve, and the intrinsic fiducial line profile is the dotted curve.

formation rate, age of star formation, escape fraction, and the typical halo mass, may also depend on redshift. Comparison of the galaxy/halo parameter dependencies described above in conjunction with the effects shown here can be used to predict trends for a specific model of the evolution of the galaxy/halo parameters. Disentangling the redshift dependence of the IGM ionization state from the other redshift dependences may still be very difficult in practice.

Proceeding with simple cases to understand the expected effects, we place the fiducial halo at $z = 8.8$ and $z = 17$. Figures 3.18 and 3.19 show the IGM and Lyman α line profile properties. Because we keep the mass of the halo fixed, the virial radius shrinks at increasing redshift. This results in an increasingly negative velocity of the IGM at the virial radius. The IGM is also denser at higher redshift, resulting in higher neutral fractions when the ionization balance is calculated. An additional effect is the increased circular velocity of the halo, which broadens the intrinsic line.

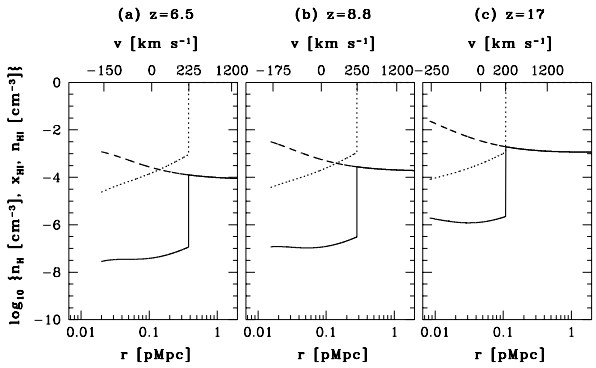


Figure 3.18 As in Fig. 3.7, but note different radial and velocity scales. The curves are truncated at $R_{\rm vir}$. (a) Redshift $z = 6.5$ (fiducial model). (b) $z = 8.8$. (c) $z = 17$.

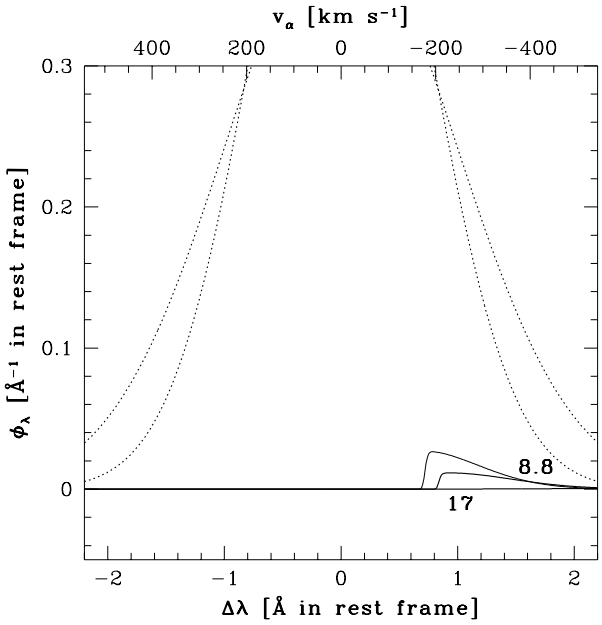


Figure 3.19 Lyman α line profiles. The wavelength scale is expressed as an offset from the Lyman α resonance in the frame of the galaxy, chosen so that the v_α scale is the same as in Fig. 3.8. The labels show the redshift for the corresponding observed line profiles. The narrower intrinsic profile (dotted curve) is for $z = 8.8$; the wider intrinsic profile (dotted curve) is for $z = 17$. The fiducial model $z = 6.5$ is the lighter curve, and the intrinsic fiducial line profile is not shown.

3.6 Galactic winds

We now consider another, possibly crucial, alteration to the model of Lyman α emitter plus IGM scattering described so far. High velocity (~ 360 km s^{-1}) redshifts of Lyman α emission compared to other interstellar features are observed in $z = 3$ Lyman Break Galaxies (Shapley et al., 2003); QSO absorption line cross-correlation suggests that these winds may have blown to large distances (~ 0.2 pMpc, Adelberger et al., 2003).

Galactic winds, if present in galaxies at $z > 6$, may have two important consequences for the interpretations of Lyman α emission from galaxies during reionization. The first is that as the wind blows through the IGM, it may collisionally ionize the hydrogen it passes, increasing the ionization fraction of this gas (Adelberger et al., 2003). The second effect is that the intrinsic Lyman α line may not have a Doppler shape centered on the galaxy's redshift. Instead, it may already display an asymmetric profile centered considerably to the red of the systemic redshift (Shapley et al., 2003, where they find a mean redshift of 360 km s^{-1}). Inspection of Fig. 3.6 shows that the IGM optical depth for the fiducial galaxy is only ~ 1 at redshifts of more than 200 km s^{-1}.[5]

Since we lack a good physical model to calculate wind properties, we adopt a very simple assumption to illustrate the possible effects of large-scale winds on the IGM (the first effect described above). We assume that a wind blows out to the Stromgren radius of the galaxy (see Fig. 3.11), completely ionizing all of the gas within that radius. Figure 3.20 shows the observed Lyman α line profiles including such winds. The result for the fiducial model is that the observed line becomes roughly symmetric, though still centered to the red of the intrinsic line center, and the observed flux doubles. If $t_{\rm SF}$ is changed to 10^9 yr, the Stromgren sphere is much larger (see Fig. 3.11) and the observed flux increases by almost a factor of 5.

The required wind velocity can be calculated by dividing the Stromgren radii (illustrated in previous plots) by $t_{\rm SF}$. For the fiducial halo, the Stromgren radius is 0.38 pMpc, giving a wind velocity over 10^8 yr of 3700 km s^{-1}. Almost any wind model would predict a smaller wind velocity than this, so though Fig. 3.20 shows the qualitative effect, even the 10^8 yr curve is an overestimate

[5]Note that in observations velocity (and redshift) is measured in the observer's frame, whereas in our plots v_α is in the frame of the emitting galaxy, hence a line redshifted by 360 km s^{-1} falls at $v_\alpha = -360$ km s^{-1}.

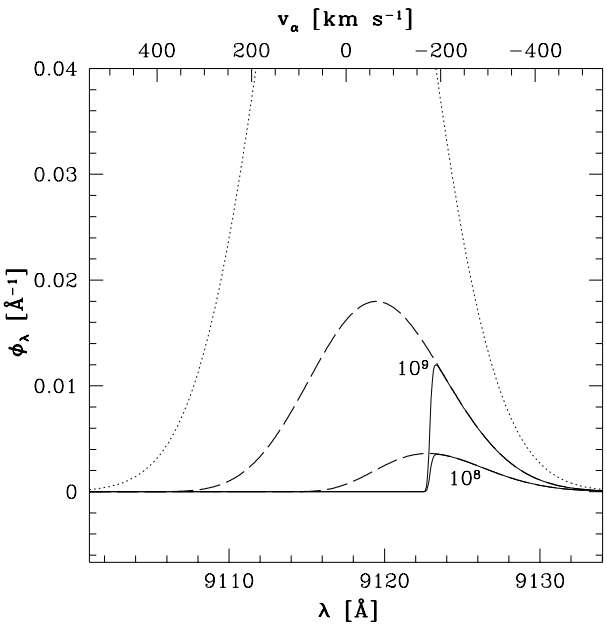


Figure 3.20 As in Fig. 3.8, but different vertical scale. The labels show star-formation timescale (in yr) for the corresponding observed line profiles. For each $t_{\rm SF}$, the light solid curve is the profile without assuming any winds (as in Fig. 3.13), and the long-dashed curve is the observed line profile assuming a wind blows to the Stromgren radius. The intrinsic fiducial line profile is the dotted curve.

of the true effect, assuming galactic winds blow to 0.2 pMpc, as in Adelberger et al. (2003).

Winds don't affect the damping wing suppression of the line from absorption by neutral (in the case of a non-reionized IGM) gas outside of the Stromgren sphere (the solid and long-dashed curves match at $v_\alpha < v(R_{\rm vir})$). Thus the depression of the observed line profiles compared to the intrinsic line profiles of plots like Fig. 3.8 can be used to estimate the transmission of a highly-redshifted Lyman α line. Since the damping wing optical depth is set by the IGM outside of the Stromgren sphere, parameters that affect the size of the Stromgren sphere (e.g., $t_{\rm SF}$, Fig. 3.11), and the ionization state of the IGM (e.g., J_{21}, Fig. 3.14, and $x_{\rm HI}^{\rm IGM}$), have a large influence on the transmission at $v_\alpha < v(R_{\rm vir})$. In the next section we illustrate the effect galactic winds may have on the observed Lyman α line for all of the parameter variations discussed in Section 3.5.

3.7 Discussion

3.7.1 Lyman α line fluxes

First we present three plots to summarize the results of the previous sections. Figure 3.21 shows the intrinsic Lyman α line flux (calculated as described in Section 3.4.1) as open triangles for each model listed on the bottom axis. The solid circles show the integrated flux of the observed Lyman α line calculated from the intrinsic Lyman α line flux and observed line profiles presented in Section 3.5.

Figure 3.22 shows the intrinsic Lyman α line flux as open triangles for each model, as in Fig. 3.21. Here the open squares are the integrated flux of the observed Lyman α line calculated assuming a galactic wind completely ionizes the IGM out to the Stromgren radius of the galaxy (see Section 3.6; this is probably an overestimate of the expected effect, but comparison with Fig. 3.21 shows that for many models this is a small effect regardless). The solid diamonds are the integrated flux of the observed Lyman α line calculated assuming the intrinsic Lyman α line is emitted entirely at 360 km s^{-1} redward of the systemic Lyman α wavelength (see Section 3.6). The consequence of moving the intrinsic line to the red is that the observed line flux is generally at least 20 per cent of the intrinsic line flux for most of the models (independent of whether of not the wind blows into the IGM). This is illustrated more explicitly in Figure 3.23. However, more massive galaxies with $v(R_{\rm vir}) \lesssim 360$ km s^{-1} (such as our $M = 10^{12}$ M$_\odot$ model) would require Lyman α lines even more redshifted than 360 km s^{-1} to make a difference compared to the no-wind model. There is no solid diamond for $M = 10^{12}$ M$_\odot$ in either Fig. 3.22 or Fig. 3.23 because the IGM transmission at $v_\alpha = -360$ km s^{-1} is formally zero. A realistic line profile (not concentrated entirely at $v_\alpha = -360$ km s^{-1}) would result in a flux similar to the no-wind model.

Figure 3.23 shows the transmission of the Lyman α line, T_α, defined as the ratio of the observed to intrinsic Lyman α line flux. The shapes of the symbols match their shapes in Figs. 3.21 and 3.22. Quite generally, an intrinsic Lyman α line redshifted by 360 km s^{-1} produces an observed line suppressed by less than a factor of 5, for galaxies at $z = 6.5$ in a neutral IGM. If the IGM is reasonably ionized, a redshifted Lyman α line may survive almost unscattered (e.g., solid diamond

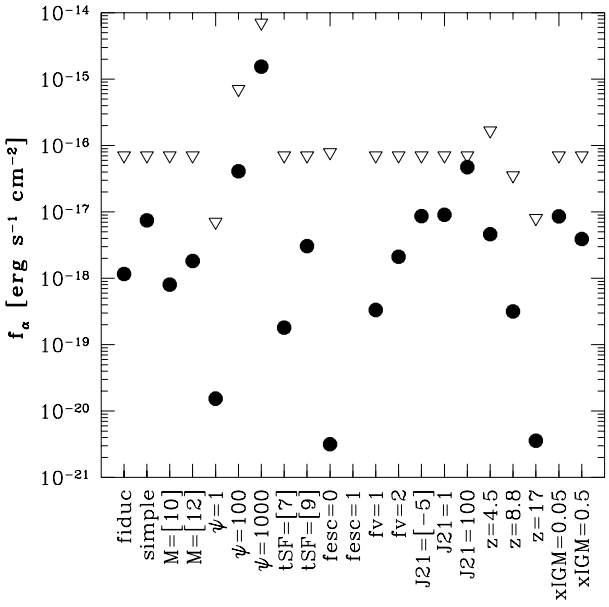


Figure 3.21 Intrinsic and observed Lyman α line flux as a function of model parameters. Each model plotted is described by a label on the bottom axis that refers to the fiducial model in the case of 'fiduc,' and refers to the parameter varied from the fiducial model in the rest of the cases (see Section 3.5). Numbers in square braces are the $\log_{10}$ of the parameter value. The open triangles are the intrinsic Lyman α line flux, and the solid circles are the observed Lyman α line flux, based on the line profiles calculated in Section 3.5.

for $x_{\rm HI}^{\rm IGM} = 0.05$ model), whereas if the intrinsic line is at the systemic redshift in those models, the observed line is almost an order of magnitude weaker (e.g., open circle for $x_{\rm HI}^{\rm IGM} = 0.05$ model).

3.7.2 Observed $z = 6.5$ Lyman α emitters

To compare observed Lyman α line fluxes to our predictions, we need an estimate of the intrinsic line flux. The observed Lyman α line flux provides a lower limit to the intrinsic flux, but current observations of $z = 6.5$ do provide one avenue to estimate the intrinsic flux. The UV continuum to the red of the Lyman α is a star-formation rate indicator (Kennicutt, 1998); near-IR photometry of $z = 6.5$ galaxies provides an estimate of their rest-frame UV continua. Two of the three confirmed $z = 6.5$ galaxies have SFRs estimated in this way: 10 $M_\odot\,{\rm yr}^{-1}$ for HCM 6A (hereafter 'H,' Hu et al., 2002b), and 36 $M_\odot\,{\rm yr}^{-1}$ for SDF J132415.7+273058 (hereafter 'K,' Kodaira et al., 2003) (in our cosmology). As discussed in Section 3.4.1, converting the SFR to an intrinsic Lyman α line flux depends on many parameters, but for our fiducial choices we estimate intrinsic Lyman α

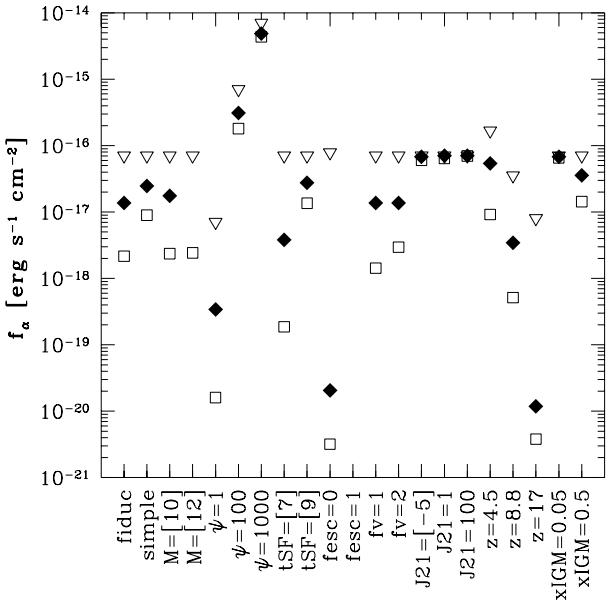


Figure 3.22 Intrinsic and observed Lyman α line flux as a function of model parameters, including possible effects due to galactic winds. The model labelling is described in Fig. 3.21. The open triangles are the intrinsic Lyman α line flux, as in Fig. 3.21. The open squares are the integrated flux of the observed Lyman α line calculated assuming a galactic wind completely ionizes the IGM out to the Stromgren radius of the galaxy. The solid diamonds are the integrated flux of the observed Lyman α line calculated assuming the intrinsic Lyman α line is emitted entirely at 360 km s^{-1} redward of the systemic Lyman α wavelength. See Section 3.6.

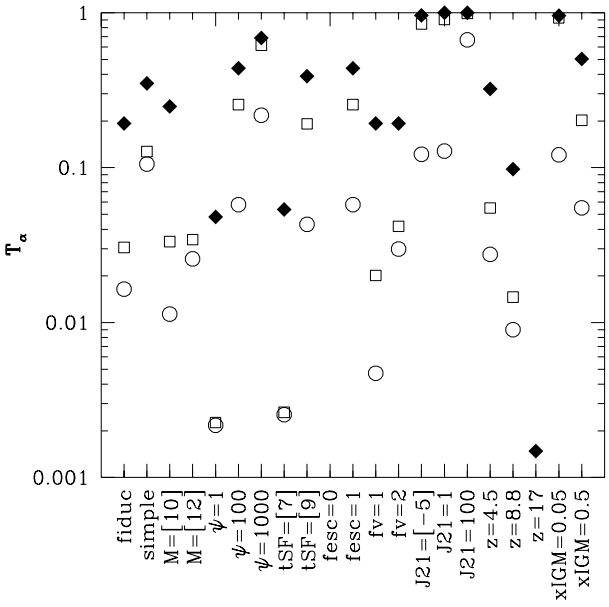


Figure 3.23 Ratio of observed to intrinsic Lyman α line flux as a function of model parameters, for both no wind and wind models. The model labelling is described in Fig. 3.21. The point shapes are the same as in Figs. 3.21 and 3.22: open circles show no-wind models, open squares show models where a strong wind clears the Stromgren sphere of IGM, and solid diamonds show models where the intrinsic Lyman α is concentrated at 360 km s^{-1} redward of the systemic redshift.

line luminosities of 3.4×10^{43} erg s^{-1} for H and 1.2×10^{44} erg s^{-1} for K. Note that both Hu et al. (2002a) and Kodaira et al. (2003) estimate Lyman α lower intrinsic luminosities by using an empirical relation between Hα luminosity and SFR (Kennicutt, 1998) and converting Hα luminosity to Lyman α luminosity. The corresponding intrinsic Lyman α line fluxes are 6.9×10^{-17} erg s^{-1} cm^{-2} for H and 2.4×10^{-16} erg s^{-1} cm^{-2} for K. Comparison with the observed Lyman α fluxes (corrected for lensing in the case of 'H') gives estimated observed to intrinsic ratios of $T_\alpha = 0.09$ for both H and K.

From Fig. 3.23 we see that appropriate models with a neutral IGM and some velocity offset between the intrinsic Lyman α line and the systemic galaxy redshift can produce values of $T_\alpha \sim 0.1$. Alternately models with a partially or mostly ionized IGM without any wind effect also have $T_\alpha \sim$ 0.1.

The redshift of the intrinsic Lyman α line with respect to the systemic redshift of the galaxy is a clearly key parameter for interpreting observed Lyman α lines. Fortunately it may be estimated

from non-resonant recombination lines, such as Hα or Hβ. For galaxies at $z \gtrsim 6.5$ these lines fall at $\lambda \gtrsim 3.6$ μm, unobservable from the ground in the near-future. The *James Webb Space Telescope* (*JWST*) is projected to have sufficient sensitivity to detect Hα and Hβ from the $z = 6.5$ galaxies discovered so far. Measurements of Hα and Hβ will also provide another SFR estimate (with $f_{\rm esc}$ no longer entering into the conversion between SFR and intrinsic Lyman α line flux), as well as a dust extinction estimate.

An example of an optimistic future observation would be a measurement of Hα and Hβ line redshifts and fluxes for a $z = 6.5$ galaxy with a measured Lyman α line and UV continuum. If the ratio of Hα to Hβ line flux matches the prediction for no extinction, and the SFR deduced from from those Balmer lines agrees with the SFR estimated from the UV continuum, then a reasonable conclusion would be that dust extinction is not very important in the galaxy[6]. Proceeding under the assumption of a normal IMF and no extinction, the intrinsic Lyman α line flux may then be estimated. If the Lyman α line were found to be centered near the Balmer line redshift (assumed to be the systemic redshift of the galaxy), then we could compare the measured value of T_α with the circles in Fig. 3.23; at that point a T_α could be converted into an estimate of the neutral fraction of the IGM, though with plenty of uncertainty associated with unknown parameters such as M, $t_{\rm SF}$, and f_v. A more clear-cut conclusion may be reached if the Lyman α line is found to be substantially redshifted compared to the Balmer redshift and the measured T_α is close to 1. This would be a strong signal that the IGM was mostly ionized at that redshift (see below), and may moreover place an interesting limit on the mass of the halo, depending on the shape of the observed Lyman α line.

In advance of *JWST*, Fig. 3.23 makes it clear that drawing any firm conclusions regarding the ionization fraction of the IGM based only on the observed Lyman α lines and UV continua of $z = 6.5$ is very difficult.

Before moving on, we present two more plots of interest regarding the observability of Lyman α lines. We have seen that large Stromgren spheres around galaxies allow more of the intrinsic Lyman α line to be observed (e.g., Fig. 3.13). The ionization of the IGM is directly proportional to the

[6]This conclusion could be strengthened by a stringent upper limit on, e.g., He II 1640 Å, a line that would be present for an IMF weighted toward very massive stars, (Schaerer, 2003).

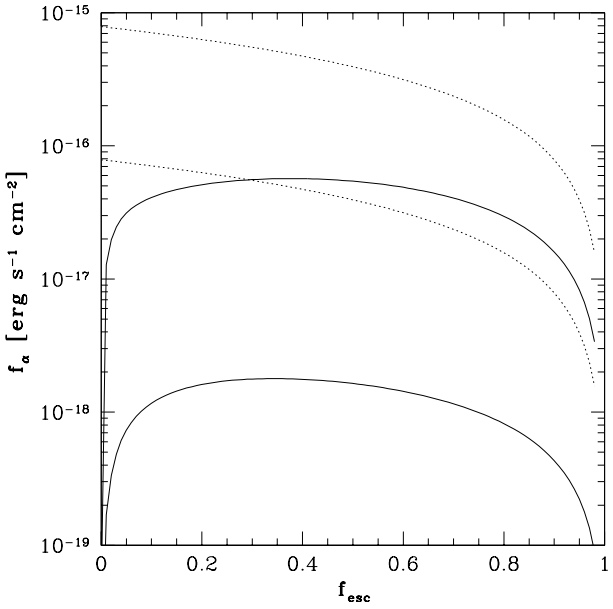


Figure 3.24 Intrinsic and observed Lyman α line flux as a function of ionizing photon escape fraction. The dotted curves are the intrinsic Lyman α line flux; the solid curves are the observed line fluxes. The lower curves are for the fiducial model with $f_{\rm esc}$ varied; the upper curves are for the fiducial model modified to $\psi = 100$ M$_\odot$ yr^{-1}.

parameter $f_{\rm esc}$, so the largest possible Stromgren sphere is created when $f_{\rm esc} = 1$. However, since the intrinsic Lyman α line flux is proportional to $(1 - f_{\rm esc})$, the strongest intrinsic Lyman α line is produced when $f_{\rm esc} = 0$. Clearly the strongest *observed* line comes from some intermediate value of $f_{\rm esc}$. In Fig. 3.24 we plot the intrinsic and observed Lyman α line flux for two models, the fiducial model, and the $\psi = 100$ M$_\odot$ yr^{-1} model. The observed line flux in both models peaks near $f_{\rm esc} = 0.35$.

Once reliable estimates of T_α are available, they can be used to estimate the neutral fraction of the IGM in the interesting regime of $x_{\rm HI} \sim 0.5$. Figure 3.25 illustrates T_α as a function of the model parameter $x_{\rm HI}^{\rm IGM}$ (see Section 3.5.4), for the fiducial galaxy. The solid curve was calculated assuming a Doppler profile intrinsic Lyman α line centered at the galaxy's systemic redshift, and the long-dashed curve was calculated assuming the intrinsic Lyman α line is concentrated 360 km s^{-1} redward of the systemic velocity. While an accurate measurement of T_α can constrain the IGM neutral fraction, this is only possible with reasonable knowledge of the redshift of the intrinsic Lyman α line compared to the systemic velocity of the galaxy.

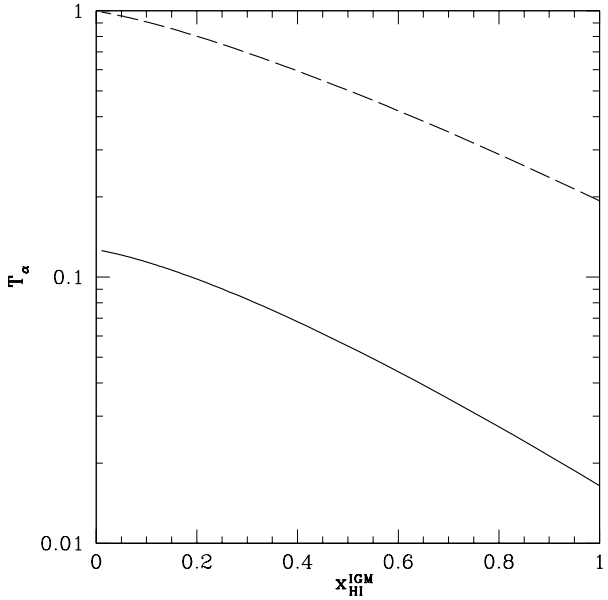


Figure 3.25 Ratio of observed to intrinsic Lyman α line flux as a function of $x_{\rm HI}^{\rm IGM}$. The solid curve assumes the intrinsic line profile described in Section 3.4.1. The long-dashed curve assumes the intrinsic Lyman α line is concentrated 360 km s^{-1} redward of the systemic velocity (see Section 3.6). Both curves are for the fiducial galaxy.

3.7.3 Model assumptions

Now we return to a discussion of the assumptions made in this paper and their possible consequences on our conclusions. First we discuss the possible effects of dust in a Lyman α emitting galaxy. If dust absorbs ionizing photons, then the star-formation rate in our models can simply be rescaled to take that into account. Alternately, dust may extinct the Lyman α line and UV continuum almost equally, but extinct the Balmer lines considerably less. In this case the Lyman α line extinction can be estimated and corrected for. Dust may also affect the UV continuum differently from the recombination emission, in the event that dust is distributed differently with respect to stars and H II regions. In this case the extinction value measured by the Balmer lines alone (with an extinction curve) may be used to correct the Lyman α line, assuming that resonant photons are extincted no differently than non-resonant photons.

Dust extinction can cause great ambiguity in the interpretation of Lyman α lines, even when UV continuum and Balmer lines are measured, because resonant Lyman α photons may be extincted differently from non-resonant UV continuum and Balmer line photons. Dust may extinct Lyman α

photons preferentially over UV continuum photons, due to the extra path-length through the galaxy that resonantly scattered photons travel before escaping. Certain gas and dust geometries, however, may mitigate or even reverse that effect (Neufeld, 1991). There is no direct observable to measure this possible effect, except that an extinction-corrected SFR estimated from Balmer lines does place an upper limit on the intrinsic Lyman α line flux.

More fundamental assumptions are associated with our model of the IGM around Lyman α emitting galaxies. We assumed the presence of an accretion shock at the virial radius of the halo containing the galaxy (see also Barkana & Loeb, 2004). If these galaxies have no accretion shock, as suggested by Birnboim & Dekel (2003), then our conclusions are greatly changed. It is the presence of the accretion shock that generates the characteristic wavelength scale where line transmission begins (see Section 3.5.2). In the absence of a virial shock, the infalling gas will eventually shock when it reaches the galaxy. The observed Lyman α line would be suppressed blueward of that velocity, implying only intrinsic Lyman α lines that were very redshifted with respect to the systemic redshift of the galaxy could be observed. Another consequence of the lack of a virial shock is that Lyman α emission from the cooling of the IGM gas would be located close to the galaxy (Birnboim & Dekel, 2003), thus any observed Lyman α line may have a contribution from contribution from cooling radiation as well as star-formation powered Lyman α. If future observations show an extremely redshifted Lyman α lines at high redshifts, this may be indicative of cooling radiation from IGM gas falling unshocked down to the center of a halo.

In Section 3.4.2.1 we discussed our assumptions about the density distribution of the IGM around the Lyman α emitting galaxy. We calculate the ensemble-averaged IGM density distribution; even for the spherically averaged density profile, deviations from this average are expected from galaxy to galaxy. Additionally, we estimated that structure near the emitting galaxy has only a small effect on the density distribution of the IGM. However, we did not address the effect of collapsed and collapsing structures on the velocity field of the IGM, which is crucial (more so than IGM density) to our predicted line profiles. Numerical simulations may be used to estimate both the variance in the IGM structure from galaxy to galaxy, and also the density and velocity distributions of the IGM

along the line of sight toward a galaxy accounting for substructure.

3.8 Summary

We have calculated observed Lyman α line profiles for high-redshift galaxies, with a focus on the prospects of using observed Lyman α lines to probe the ionization of the universe at $z > 6.5$. Toward this end we have investigated many of the parameters that influence the observed Lyman α line from a high-redshift galaxy. We conclude that future measurements with *JWST* may permit the estimation of IGM neutral fractions at $z > 6.5$, but current $z = 6.5$ data do not yet place a strong constraint on the neutral fraction at that redshift if we allow consideration of galactic winds. If we take at face value our estimates of the integrated Lyman α transmission of the two $z = 6.5$ galaxies for which estimation is possible, and assume our simplest model (with no winds), we would conclude that the IGM neutral fraction is $x_{\rm HI} \lesssim 0.1$ at $z = 6.5$.

Interpretation of an observed Lyman α line alone (or even with a measurement of the UV continuum) is extremely difficult, due to the host of parameters that influence it. However there is hope that with the rest-frame optical spectra that *JWST* will provide for high-redshift galaxies, Lyman α emission lines will yet prove fruitful for constraining the reionization history of the universe at $z > 6.5$

Acknowledgments

We acknowledge Kurt Adelberger for suggesting the IGM model used here. We thank Marc Kamionkowski, Avi Loeb, and Alice Shapley for helpful conversations. MRS acknowledges the support of NASA GSRP grant NGT5-50339.

Chapter 4

A Method to Infer the Stellar Population that Dominated the UV Background at the End of Reionization[1]

4.1 Abstract

We present an observational test of the spectrum of the ionizing background at $z \simeq 5$; the test is sufficiently sensitive to determine whether Pop II or Pop III stars are the dominant source of ionizing radiation. The ionizing background at $z \simeq 5$ may reflect the nature of the sources responsible for the final overlap phase of reionization. We find that rest-frame extreme-UV He I absorption will be detectable in deep spectral observations of some rare $z \simeq 5$ quasars; the ratio of He I to H I absorption reflects the shape of the ionizing background in the photon energy range between 13.6 and 24.6 eV. Most $z \simeq 5$ quasars have too much H I absorption along their line-of-sight for He I absorption to be observed. However, based on current measurements of H I absorber statistics, we use Monte Carlo simulations to demonstrate that the Sloan Digital Sky Survey (SDSS) will discover a sufficient number of $z \simeq 5$ quasars to turn up a quasar suitable for measuring He I absorption (and we illustrate a selection method to identify that quasar). From simulated observations of a suitable $z \simeq 5$ quasar with a 10-meter telescope, we show that a constraint on the spectral slope of the ionizing background at that redshift can be obtained.

[1]This chapter was submitted to Monthly Notices of the Royal Astronomical Society in 2003 as Santos and Loeb.

4.2 Introduction

Intergalactic hydrogen is almost completely ionized by $z \simeq 6$ (Becker et al., 2001; Djorgovski et al., 2001), but the sources responsible for reionization are still not known (see Barkana & Loeb 2001 and Loeb & Barkana 2001 for reviews of the theoretical possibilities). A low neutral-hydrogen fraction is found at lower redshift even in relatively over-dense systems, despite the short (compared to the Hubble time) recombination timescale in those systems. This implies a continued supply of ionizing photons after reionization completed. Measurements of the intensity of the ionizing background below $z \sim 3$ (Scott et al., 2002) are consistent with an ionizing background produced by observed quasars (Haardt & Madau, 1996), but analysis of a combination of the hydrogen and helium opacities suggests that galaxies and quasars contribute equally to the UV background at that redshift (Sokasian et al., 2003). At higher redshifts, however, the declining abundance of bright quasars (e.g., Fan et al., 2001b) suggests that they cannot provide an ionizing background sufficient to reionize the universe (Wyithe & Loeb, 2003a), and this conclusion is supported by deep x-ray surveys (Barger et al., 2003).

Recent results from the *WMAP* satellite imply that reionization was substantially underway by $z \sim 15$ (Kogut et al., 2003; Spergel et al., 2003). However, quasar absorption studies at $z \simeq 6$ (Becker et al., 2001; Djorgovski et al., 2001) suggest a dramatic evolution in the ionizing background around that redshift. Some authors have reconciled the *WMAP* evidence for early reionization with the quasar result that reionization is just finishing at $z \simeq 6$ by postulating two reionizations of the universe (Cen, 2003b,a; Wyithe & Loeb, 2003a,b), generally invoking a hard Pop III stellar spectrum for the early reionization and a Pop II stellar spectrum for the final reionization at $z \simeq 6$.

The most direct technique to ascertain the sources responsible for the ionizing background at high redshift is to take a census of all sources that contribute ionizing photons. The advantage of this technique is that the relative contribution of different types of sources are measured directly, as are the spatial distributions and other properties of the populations. Unfortunately, the large luminosity distance to high redshift means that only the most luminous sources can be detected with current observational methods. For example, Fan et al. (2001b) measure a power-law high-redshift

quasar luminosity function with a steep slope, but estimate that the traditional power-law break of the quasar luminosity function would occur more than one magnitude below their survey limit: there is little constraint on the faint end of the quasar luminosity function from direct source detection.

A complementary method to direct detection is measurement of the properties of the integrated ionizing background. Although this technique does not identify ionizing sources directly, it is effective at detecting ionization due to a very abundant population of low-luminosity sources. Moreover, the shape of the integrated ionizing background is dictated by the combination of sources responsible for it; different spectra arise from various possible contributors, such as quasars (Telfer et al., 2002), Pop II OB stars (e.g., Leitherer et al., 1999), very massive metal-free (Pop III) stars (Bromm, Kudritzki, & Loeb, 2001b), and x-rays from early stars (Oh, 2001).

This paper describes possible observations that would constrain the spectral shape of the ionizing background just after reionization finished. In Section 4.3, we introduce quasar absorption line spectroscopy in the context of this paper. Section 4.4 describes our parametrization of the ionizing background, and the resulting ionization state of absorption systems in the intergalactic medium (IGM). In Section 4.5 we apply those results to observed statistics of H I absorbers to generate a model for H I and He I opacity toward high-redshift quasars; the assumed properties of the quasars are given in Section 4.6. The method of simulating line-of-sight (LOS) absorption spectra and studying the sensitivity of the results to the shape of the ionizing-background spectrum is described in Section 4.7. Section 4.8 presents a method to select the best quasars for absorption line spectroscopy. We summarize our results in Section 4.9.

4.3 Absorption spectroscopy

The easiest way to make a crude measurement of the shape of the ionizing background is through the ionization state of photoionized IGM gas; in the next section we show explicitly how the background spectrum is related to the ionization state. Atoms and ions with at least one bound electron in the ground state have strong bound-bound transition cross-sections, typically located at UV or soft x-ray energies. Consequently, even small amounts of these species on the line-of-sight to a

background source produce strong absorption features. The most common extragalactic application of this technique uses UV-bright quasars as the background sources (for a recent review see Bechtold, 2003). Many species have been detected in absorption, ranging from H I and He II, to Fe II and Zn II, plus high-ionization ions such as C IV and Si IV (e.g., Prochaska et al., 2001).

We examine the first ionization states of hydrogen and helium; they are the most abundant elements in the IGM, and are responsible for most of the UV opacity toward high-redshift quasars. Specifically, we will investigate the relative neutral fractions of hydrogen and helium, which have first ionization potentials of 13.6 and 24.6 eV, respectively. Since the photoionization cross-sections for H I and He I are peaked at threshold, their relative neutral fractions constrain the effective slope of the ionizing background between 13.6 and 24.6 eV.

The relative abundances of H I and He I constrain the ionizing background—the challenge is to measure these abundances in the IGM at high redshift. Both atoms exhibit line and continuum absorption. For H I, the Lyman series lines begin at 1216 Å and continue down to 912 A, where continuous photoelectric absorption begins. The corresponding line transitions in He I are from $1s^2$ to $1s2p$ (584 Å), $1s3p$ (537 Å), etc., down to the first ionization threshold of helium, at 504 Å. These wavelengths are the rest-frame values; the observed wavelengths of these transitions depend on the redshift of the absorbing gas. If the neutral gas is distributed in many systems with a discrete redshift distribution, then the associated absorption lines appearing in a quasar spectrum form a 'forest' of absorption features against the continuum. Figure 4.1 shows two example absorption spectra with the general absorption regions labelled.

The absorption cross-sections of H I and He I are larger at the center of a line (at the IGM temperature of $\sim 2 \times 10^4$ K) than in the continuous absorption region. Despite this fact, the strongest absorption effect for the expected distribution of neutral IGM gas (see Section 4.5) is from the cumulative continuous absorption due to many absorbers at different redshifts, named the 'valley' by Moller & Jakobsen (1990). This is illustrated for H I in Fig. 4.1; once Lyman limit absorption begins (shortward of $912(1+z_{\rm Q})$ Å, where $z_{\rm Q}$ is the quasar redshift), the forest typically gives way to almost complete absorption. In such cases it is very difficult to observe helium forest

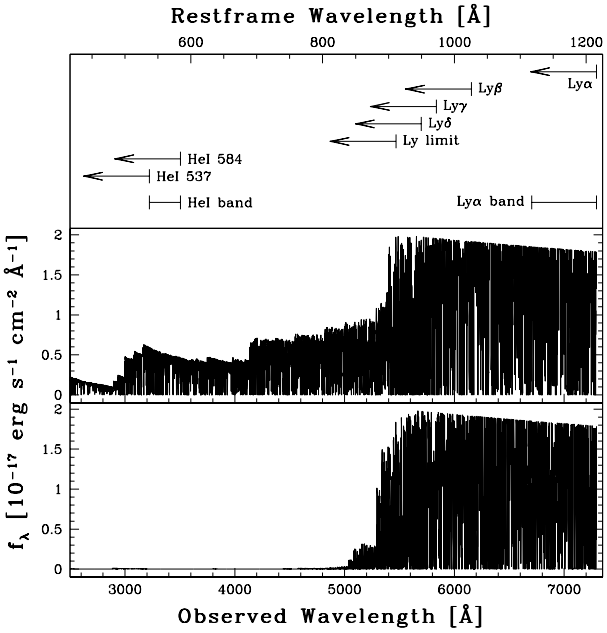


Figure 4.1 Simulated absorption spectra for $z = 5$ quasars with $\mathrm{AB}_{1450} = 20$. The top axis is labelled in rest-frame wavelength, and the bottom axis is labelled in the corresponding observed wavelength at $z = 5$. The top panel shows the wavelengths where the first four Lyman series line forests begin; it also shows where the the first two ground-state He I transition opacities begin. The He I band is shown, as is the region of the corresponding Lyα forest. The middle panel shows a quasar spectrum processed by IGM absorption, along a line-of-sight with relatively little IGM opacity. The bottom panel shows a quasar absorption spectrum along a typical line-of-sight.

lines because of the small residual continuum flux at short wavelengths. In rare cases the valley is not very pronounced, and the continuum is strong enough that the He I 584 Å forest region may be observed. We call the region of the He I 584 Å forest red-ward of He I 537 Åthe 'He I band.'

The identification of absorption features in spectra like those in Fig. 4.1 can be difficult, especially in low-resolution spectra, except in the region red-ward of the Lyβ forest, where the only possible absorption feature from hydrogen or helium is the hydrogen Lyα line. (Trace metals produce additional absorption features, though these can be discriminated based on linewidth or because they have a doublet feature, e.g., C IV.) The He I band is the analogous region for He I, where the He I 584 Å lines are the only source of helium opacity; however, this region also may contain absorption due to all of the H I Lyman transitions, confusing the assignment of any individual line to He I. In principle, unsaturated He I lines could be identified because they are twice as narrow as unsaturated Lyα lines. However, line blending and the expected weakness of the He I lines complicate a practical implementation of that criterion, and spectroscopy of sufficiently high resolution would require an extremely bright target.

Even after hydrogen reionization is complete at $z \simeq 6$, the filamentary nature of the clumpy IGM leads to significant Lyα opacity in regions with neutral hydrogen column density of $N_{\rm HI} \sim 10^{14}$ cm^{-2}. A line-of-sight through the universe pierces many of these filaments at different redshifts, resulting in the Lyα forest described above. Helium is expected to be singly reionized (24.6 eV ionization threshold) at a similar redshift to hydrogen, for almost any expected ionizing spectrum. Consequently, absorption by neutral helium at $z < 6$ should be confined to the same filaments that give rise to hydrogen absorption.[2] This justifies a one-to-one search for He I 584 Å lines at the same redshift as lines identified in the 'Lyα band' region, where the Lyα band is defined as the spectral region where absorbers with a He I 584 Å line in the He I band exhibit Lyα absorption (see Fig. 4.1).

The application of absorption line studies of the IGM is limited by the supply of suitable background sources, primarily quasars. The Sloan Digital Sky Survey[3] (SDSS) is a large photometric and spectroscopic survey of the northern sky that will identify and measure redshifts for almost all

[2]This is in contrast to doubly ionized helium, which is distributed more diffusely (that is, the absorption features traces smaller over-densities) than hydrogen (Kriss et al., 2001).

[3]http://www.sdss.org

quasars brighter than 20th magnitude over one quarter of the sky. At the current discovery rate of the survey, it will find approximately 1000 quasars at $4 < z < 5.2$, and about 200 at $4.8 < z < 5.2$ (Anderson et al., 2001; Fan et al., 2001b). As we will demonstrate (Section 4.8), rest-frame extreme ultraviolet (EUV) absorption line study of $z \simeq 5$ quasars requires a large sample of quasars, and could only be realized with the large catalog of $z \simeq 5$ quasars that the SDSS will provide.

4.4 Ionizing background

Next we consider the relationship between the ionization state of IGM absorbers and the ionizing background. The quantity of interest for an absorbing system is the ratio of column density of neutral helium, $N_{\rm HeI}$, to the column density of neutral hydrogen, $N_{\rm HI}$,

$$\eta \equiv \frac{N_{\rm HeI}}{N_{\rm HI}}. \tag{4.1}$$

As we will describe in Section 4.5, the abundance of absorbers as a function of redshift and $N_{\rm HI}$ is already observed. Thus predictions for He I absorption from these systems can be easily made from calculations of the relative ionization states of helium and hydrogen.

We will assume that an absorbing system is optically thin to ionizing radiation and ignore the presence of He III in the absorber. As we will show in Section 4.7, the most important absober systems for our method attenuate almost all of the light at the line center of their Lyα absorption lines. Which absorbers satisfy this criterion depends on two quantities, $N_{\rm HI}$, which specifies the intrinsic line strength, and the spectral resolution of the Lyα forest observations we consider, 1.7 Å, which is much larger than the intrinsic line widths (see Sections 4.5 and 4.7). Most of the saturated Lyα absorption is generated by lines with equivalent widths comparable to the spectral resolution, since higher column-density systems are less common (see Section 4.5). For a resolution of 1.7 Å, absorbers with $N_{\rm HI} \simeq 10^{17}$ cm^{-2} are the most important for the aims of this paper.

Absorbers with $N_{\rm HI} \simeq 10^{17}$ cm^{-2} are have optical depths near unity to photons at the H I ioninizing threshold; such absorbers are still optically thin to He I-ionizing photons. These absorbers

are expected to have densities of about 4×10^{-3} cm^{-3} at $z = 5$ (Schaye, 2001).

Thus the relative abundance of neutral helium to neutral hydrogen in an absorbing system with high H II and He II fractions[4] is described by the optically-thin limit (e.g., Miralda-Escude & Ostriker, 1992),

$$\eta_{\rm thin} \equiv \frac{N_{\rm HeI}}{N_{\rm HI}} = \frac{n_{\rm HeII}}{n_{\rm HII}} \frac{\alpha_{\rm HeI}}{\alpha_{\rm HI}} \frac{\Gamma_{\rm HI}}{\Gamma_{\rm HeI}} \simeq 0.083 \frac{\Gamma_{\rm HI}}{\Gamma_{\rm HeI}}, \tag{4.2}$$

where $n_{\rm He}/n_{\rm H} \simeq 0.083$ is the cosmic abundance ratio of helium to hydrogen (assuming a mass fraction of ^{4}He of 0.25, consistent with the combination of *WMAP* CMB results (Spergel et al., 2003) and big bang nucleosynthesis predictions (Burles et al., 2001)), $\alpha_{\rm HeI}/\alpha_{\rm HI} = 1.0$ is the ratio of the radiative recombination coefficients for He I and H I (Osterbrock, 1989), and

$$\Gamma_{\rm HI} = \int_{\nu_{\rm HI}}^{\infty} {\rm d}\nu \frac{4\pi J_\nu \sigma_{\rm HI}(\nu)}{h\nu}, \tag{4.3}$$

and similarly for $\Gamma_{\rm HeI}$. The angle-averaged specific intensity is J_ν, $\sigma_i(\nu)$ is the photoionization cross-section of species i, and the lower limit of the frequency integral is ν_i, the photon frequency at the ionization threshold of species i =H,He.

The most important absorbers for measuring η are systems where the observed equivalent widths of the Lyα forest absorption lines are roughly equal to the spectral resolution obtained in the Lyα forest spectrum. For a typical spectral resolution of 0.27 Å (a choice justified in Section 4.7.2), this corresponds to $N_{\rm HI} \simeq 3 \times 10^{14}$ cm^{-2} at $z \sim 5$. At such column densities, the optical depth to photons at the hydrogen ionization threshold is $\tau_{\rm HI} \lesssim 2 \times 10^{-3}$, thus the optically-thin approximation should be valid.

However, optically thick absorbers can have a substantial impact on the spectrum of the ionizing background, both through absorption and emission. Haardt & Madau (1996) find that the mean UV background is attenuated by roughly a factor of 3 at both $\nu_{\rm HI}$ and $\nu_{\rm HeI}$; additionally, much of the flux shortward of 228 Å is absorbed by He II and re-emitted at 304 Å (see Haardt & Madau,

[4]We neglect the possible presence of He III in the absorption systems. Observations suggest that the second reionization of helium did not occur until $z \simeq 3$ (Kriss et al., 2001; Theuns et al., 2002), and Wyithe & Loeb (2003b) predict that the full helium reionization happens over a narrow redshift interval. These studies suggest that most of the intergalactic helium was singly ionized at $z \simeq 5$.

1996, fig. 5c). These changes to the unabsorbed spectrum have important implications for the He III abundance, but do not affect the abundance ratio of He I to H I.

Because we are only concerned with the relative rate of neutral-helium ionizations to hydrogen ionizations, we approximate the ionizing background by a power law,

$$J_\nu \propto \nu^{-\alpha_\mathrm{b}}. \tag{4.4}$$

Using a continuous star-formation history with metallicity $Z = 10^{-3}$ and a Salpeter IMF from 1 to 100 $\mathrm{M}_\odot$ (Leitherer et al., 1999), the effective power-law index is $\alpha_\mathrm{b} = 2.05$. Also of interest is the spectrum of very massive metal-free stars, which represent the most extreme expectation of the stellar IMF at very high redshift: these stars are well approximated by a blackbody spectrum with temperature $T = 10^5$ K (Bromm et al., 2001b). Very massive metal-free stars yield a particularly hard power-law index of $\alpha_\mathrm{b} = 0.28$. Quasars produce a spectrum with a power-law shape matching their own EUV spectral slope; that value is not well constrained at high redshift, but $\alpha_\mathrm{b} = 1.6$ may be a reasonable guess (see Section 4.6). Finally, x-rays radiated from early stars have a near-zero spectral slope (Oh, 2001), but the effect of secondary ionizations should produce an effective slope slightly greater than zero (S.P. Oh, private communication). Figure 4.2 shows the value of η_thin as a function of α_b.

4.5 IGM absorption model

Our model for the H I absorption in the IGM is empirical, based on the data obtained in many quasar absorption line studies. Ongoing and planned observations are likely to constrain its parameters better in the future.

Absorption by H I is quantified by discretizing the absorption into individual absorbing systems, identified by their redshift, z, and neutral hydrogen column density, N_HI. The distribution of H I systems may be described reasonably well by power laws in both redshift and column density. We define H I systems with $N_\mathrm{HI} < 1.6 \times 10^{17}$ cm^{-2} as Lyα forest systems, and systems with $N_\mathrm{HI} \geq$

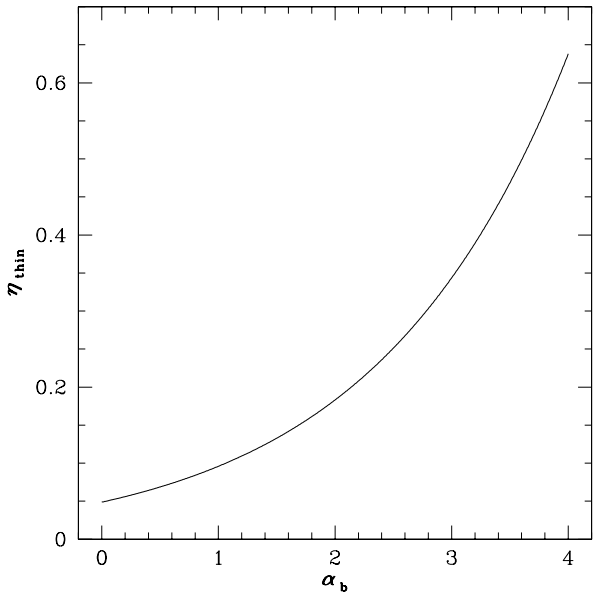


Figure 4.2 The ratio, $\eta_{\rm thin}$, of the column density of neutral helium to the column density of neutral hydrogen, as a function of the spectral slope of the ionizing background, $\alpha_{\rm b}$. We assume that the absorbing systems are optically thin to ionizing radiation, highly ionized in hydrogen and highly singly-ionized in helium. The effective power-law index of a normal stellar population is $\alpha_{\rm b} = 2.05$; the effective power law index of a population of very massive metal-free stars is $\alpha_{\rm b} = 0.28$.

1.6×10^{17} cm^{-2} as Lyman Limit Systems (LLSs). The distribution of H I systems in redshift and column density are parametrized by three constants: A, γ, and s,

$$\frac{{\rm d}N(z)}{{\rm d}z} = A(1+z)^{\gamma}, \tag{4.5}$$

$$f(N_{\rm HI}) \propto N_{\rm HI}^{-s}. \tag{4.6}$$

Here $N(z)$ is the total number of absorbers along a line-of-sight to redshift z, and $f(N_{\rm HI})$ is the number of absorbing systems per unit H I column density.

We chose $s = 1.5$ for the power law index of the $N_{\rm HI}$ distribution for all values of $N_{\rm HI}$, and assume that the absorber population extends up to $N_{\rm HI} = 10^{22}$ cm^{-2} (Storrie-Lombardi & Wolfe, 2000). For the Lyα forest, we adopt the redshift evolution determined by *HST* observations at low redshifts (Dobrzycki et al. 2002, cf. Weymann et al. 1998a) and from ground-based surveys at high redshifts (Bechtold, 1994). Those surveys covered the H I column density range of 10^{14} cm^{-2} $\leq$

$N_{\rm HI} \leq 1.6 \times 10^{17}$ cm^{-2} (i.e., from a rest-frame equivalent width of 0.24 Å up to the LLSs). We extend their results to $N_{\rm HI} = 10^{13}$ cm^{-2} by scaling up the number density using the $N_{\rm HI}$ distribution quoted above, to give

$$\begin{aligned} A = 106, \quad \gamma = 0.65, \quad \text{for } z < 2.4, \\ A = 29, \quad \gamma = 1.7, \quad \text{for } z \geq 2.4. \end{aligned} \tag{4.7}$$

The transition redshift, $z = 2.4$, was chosen to match the low-redshift and high-redshift fits smoothly; this is higher than the traditional value of about 1.5 (e.g., Weymann et al., 1998a), but is a consequence of the upward revision of γ by Dobrzycki et al. (2002) compared to Weymann et al. (1998a); given the large uncertainties in the values of A and γ, this is not alarming.

For the LLSs, we adopt $A = 0.2$ and $\gamma = 1.5$, which are consistent with the results of Storrie-Lombardi et al. (1994) and Stengler-Larrea et al. (1995). We note that Stengler-Larrea et al. (1995) do not present an analysis of their full sample of LLSs because they exclude absorption systems within 5000 km s^{-1} of the quasar, yet their results are consistent with the result of Storrie-Lombardi et al. (1994) for all LLSs. These parameter choices provide a relatively smooth intersection between the abundance of LLSs and Lyα forest systems at $z \geq 1$, assuming the column density distribution described above. The covariance of the uncertainties in A and γ implies that the mean number of LLSs toward a high-redshift quasar is constrained, but the LLS distribution with redshift is uncertain: the 1-σ uncertainty on γ is about 0.4 (Storrie-Lombardi et al., 1994; Stengler-Larrea et al., 1995).

The absorber distributions described above can be converted into a mean comoving H I number density using

$$\langle n^{\rm c}_{\rm HI}(z) \rangle = A(1+z)^{\gamma} \frac{H_0 E(z)}{c(1+z)^2} \langle N_{\rm HI} \rangle \,, \tag{4.8}$$

where H_0 is the Hubble constant, $\langle N_{\rm HI} \rangle$ is the mean H I column density of the absorber population, and

$$E(z) = \left[\Omega_{\rm m} (1+z)^3 + \Omega_{\Lambda} + (1 - \Omega_{\rm m} - \Omega_{\Lambda})(1+z)^2 \right]^{1/2} . \tag{4.9}$$

For a universe with density parameters $\Omega_{\rm m} = 0.3$ in matter, $\Omega_{\Lambda} = 0.7$ in a cosmological constant,

and $\Omega_{\rm b}h^2 = 0.02$ in baryons, and a Hubble constant $h \equiv H_0/(100 \text{ km s}^{-1}\text{Mpc}^{-1}) = 0.7$, this gives an H I neutral fraction at $z \gtrsim 4$ of,

$$\langle x_{\rm HI}(z)\rangle \simeq 9.8 \times 10^{-3} A (1+z)^{\gamma-1/2} \frac{\langle N_{\rm HI}\rangle}{4 \times 10^{19} \text{ cm}^{-2}}. \tag{4.10}$$

For LLSs (which have $\langle N_{\rm HI}\rangle = 4 \times 10^{19}$ cm^{-2}), $A = 0.2$ and $\gamma = 1.5$, thus $\langle x_{\rm HI}(z=6)\rangle = 0.01$. This is consistent with the limit derived by Fan et al. (2002) from H I absorption toward a $z = 6.28$ quasar.

We assume that there is no correlation between absorbing systems. There is observational evidence for clustering of the Lyα forest lines (e.g., Liske et al., 2000; Dobrzycki et al., 2002, and references therein), but our assumption is conservative for the purposes of this paper: we will show that to study He I absorption, we need quasars with fewer than average absorbers along the line-of-sight. Clustering of absorbers would skew the 'bad' quasar targets (i.e., those with many strong absorbers) worse, and skew the 'good' quasar targets better. This is not an important effect for low-$N_{\rm HI}$ absorbers, since they are very numerous, but could be a significant effect for LLSs if they are clustered (note, however, that Sargent, Steidel, & Boksenberg 1989 showed that LLSs followed Poisson statistics in their sample of 37 absorbers). The proximity effect generates an absorber deficit of ~ 30 per cent within 4 h^{-1} Mpc of the quasar, or about 200 Å observed from the quasar Lyα line (Scott et al., 2000). We also ignored this relatively small effect.

In addition to the absorber redshift and column-density distributions, our IGM model must also describe their absorption properties. We used the photoionization cross-sections given by Osterbrock (1989) for H I and Verner et al. (1996) for He I. Our model treats the first 11 line transitions from the ground state (which all atoms are assumed to occupy), using data from the NIST Atomic Spectra Database v2.0.[5] The line profiles were modelled as Doppler cores of width $b = 26$ km s^{-1} (Kim et al., 1997) with damping wings outside of the core (e.g., Peebles, 1993).[6]

[5] http://physics.nist.gov/cgi-bin/AtData/main_asd

[6] In Section 4.7.2 we assume instrumental resolutions considerably larger than 26 km s^{-1}, so the choice of exactly that value and our use of a single b value for all absorbers rather than the actual distribution should have a negligible effect on our results.

Our absorption model includes only hydrogen and helium; metal lines are not considered here. There are not likely to be many metal line systems in the quasars used to measure η because metal lines are associated with H I absorbers of high column density, and, due to the selection techniques employed (see Section 4.7), quasars with strong absorbing systems are unsuitable for measuring η. Moreover, the strongest metal lines that could pollute the Lyα band, such as the C IV and Mg II doublets, show only a few absorbers per unit redshift, compared to of order 100 Lyα absorbers in the same redshift interval. In the He I band, metal lines are again far less numerous than H I lines, and thus were ignored.

One complication to the application of absorber statistics determined from other quasar samples to the population of quasars discovered by SDSS is the quasar selection techniques employed. The SDSS quasar selection primarily uses observed optical colors to generate a well-defined selection function (Fan et al., 1999); by contrast, the quasars studied in absorption line surveys were culled from heterogeneous catalogs (e.g., Hewitt & Burbidge, 1987), and include radio-, x-ray-, and emission-line-selected quasars in addition to color-selected quasars. An analysis of the systematic errors introduced by using absorption lines in quasars discovered by several different techniques to predict absorption patterns in quasars selected by another technique is outside the scope of this paper; we simply analyze the colors of the quasars most important to this study to ensure they would meet the quasar color-selection criteria of SDSS (see Section 4.8).

4.6 Model quasar intrinsic spectrum

In order to simulate realistic observations of IGM absorption of background quasars, we need to make assumptions about the intrinsic properties of the quasar UV spectrum. The SDSS will provide a large catalog of high-redshift quasars, so we attempt to model the typical properties expected of SDSS quasars based on the results of the survey so far.

In a sample drawn from 182 square degrees of the SDSS, Fan et al. (2001a) presented 18 quasars with $z \geq 4$ and $i^* < 20$, where i^* is the preliminary SDSS determination of the quasar i' magnitude. We assume the quasars presented are a fair sample of the final SDSS results.

For each quasar in their sample, Fan et al. (2001a) compute the properties of the quasar near-UV (NUV) continuum, assuming a power law form, $f_\nu \propto \nu^{-\alpha_{\rm NUV}}$. They set the normalization to the quasar continuum at the observed wavelength corresponding to rest-frame 1450 Å. They then convert that specific flux into a magnitude on the AB system,

$$\mathrm{AB}_{1450} = -2.5 \log_{10}[f_\nu(\lambda_{\rm rest} = 1450\ \text{Å})] - 48.6, \tag{4.11}$$

where $f_\nu(\lambda_{\rm rest} = 1450)$ is the specific flux in erg s^{-1} cm^{-2} Hz^{-1} at rest-frame 1450 Å. Typically $\mathrm{AB}_{1450} \simeq i^*$ or z^*, depending on quasar redshift. Fan et al. (2001a) also estimate $\alpha_{\rm NUV}$ for each quasar; the mean for quasars with $z \geq 4$ is 0.6, with substantial uncertainty in the slope of any individual quasar. However, Telfer et al. (2002) found a break in the UV slope of their quasar composite spectrum near the wavelength of Lyα; the measured slope became softer blue-ward of Lyα. This confirmed earlier work by Zheng et al. (1997). The Telfer et al. (2002) sample is comprised almost entirely of quasars with $z < 2.5$; for radio-quiet quasars they find a mean EUV slope of about 1.6, with a break to a shallower NUV slope at about 1250 Å. Telfer et al. (2002) compare their radio-quiet sample to the radio-quiet SDSS quasar sample of Vanden Berk et al. (2001), and find an evolution toward harder NUV slope with samples at higher redshift; EUV comparison between the samples is extremely difficult due to the presence of strong IGM absorption in the high-redshift quasars.

Given the available observational evidence, we adopt a conservative model for the EUV properties of SDSS quasars: $\mathrm{AB}_{1450} = 20$ with a NUV slope of 0.6 between rest-frame 1450 and 1250 Å and an EUV slope shortward of 1250 Å given by $\alpha_{\rm EUV} = 1.6$.

4.7 Lines-of-sight toward SDSS quasars

Many $z \sim 5$ quasars will be discovered by SDSS; in this section we evaluate how good the best quasar target will be. We then simulate observations of the best quasar target and demonstrate our ability to constrain the spectrum of the ionizing background from such data.

Extrapolating from the results of the color-selected sample of Fan et al. (2001a), SDSS will discover ~ 1000 quasars with $4 < z < 5.2$. We assume the SDSS quasar selection function does not depend much on redshift over the range $4 < z < 5.2$; this is a conservative estimate for our purposes because the detection probability is higher for quasars with $4.7 < z < 5.2$ than for other $z > 4$ quasars (Fan et al., 2001a). Fan et al. (2001b) fit the redshift dependence of the high redshift quasar spatial density with $\rho \propto 10^{-0.5z}$. Thus about 200 of the high-redshift quasars will fall within the range $4.8 < z < 5.2$.

We would like to know how many of these quasars have spectra like the middle panel of Fig. 4.1, which would be useful for measuring He I lines, and how many have spectra like the bottom panel of Fig. 4.1, which would not be a suitable target for measuring He I lines. We quantify this by determining the distribution of quasar sightline 'suitability,' as measured by the flux of the quasar in the He I band. From that distribution and the expected number of SDSS quasars, we determined what the most suitable quasar discovered by the SDSS will be. Though for some applications an analytic approach to the subject is suitable (Zuo & Phinney, 1993), we require a Monte Carlo simulation approach (Moller & Jakobsen, 1990; Jakobsen, 1998) for this study.

4.7.1 Absorber Monte Carlo simulations

We started from the distribution of H I absorbing systems in redshift and column density described in Section 4.5. For a quasar at a given redshift $z_{\rm Q}$, we used Poisson statistics to generate the number of low-z Lyα forest, high-z Lyα forest and LLS absorbers along the line-of-sight. Once we generated the number of absorbers of each type, we assigned redshifts and column densities drawn from the absorber model described in Section 4.5. This list of absorbers was then passed through a routine to calculate the optical depth to H I at every sampled wavelength value.[7] In each simulated spectrum, the optical depth due to helium was stored separately from hydrogen; given the universal optically thin ratio of column densities, $\eta_{\rm thin}$ derived in Section 4.4, the helium optical depth scales simply as $\eta_{\rm thin}$.

[7]Wavelength sampling ranges from 0.08 to 0.28 Å, chosen so that there are three samples per Doppler FWHM of the spectral lines of interest (He I lines are twice as narrow as H I lines); all results were checked for sensitivity to spectral resolution and are converged at this resolution choice.

For each LOS, the H I optical depth data were converted into the IGM transmission as a function of wavelength. We integrated the IGM transmission over the He I band to derive the mean He I-band transmission along each LOS, $T(\mathrm{He\,I})$. Figure 4.3 shows the cumulative fraction of quasars with $T(\mathrm{He\,I})$ greater than a given value. For simplicity, we will assume the expected 200 SDSS quasars with $4.8 < z_{\mathrm{Q}} < 5.2$ all fall at $z_{\mathrm{Q}} = 5$. Then from Fig. 4.3 we can read off the value of $T(\mathrm{He\,I})$ corresponding to a fraction of 1/200; this is the largest expected $T(\mathrm{He\,I})$ value of the lines-of-sight toward the SDSS quasars. The result is $T(\mathrm{He\,I}) = 0.03$. Note that if we change the assumed redshift of the quasars by $\Delta z = 0.2$ (the dotted lines in Fig. 4.3), the expected maximum value of $T(\mathrm{He\,I})$ for 200 quasars changes by a factor of about 2. When the final SDSS quasar catalog is constructed, we may use the real distribution of quasar redshifts together with our simulation machinery to generate a $T(\mathrm{He\,I})$ histogram that depends only on the IGM absorber model. Thus measurements of the observed distribution of $T(\mathrm{He\,I})$ can be used to constrain a combination of the properties of the H I component of the IGM and the intrinsic quasar spectral shape.

Small changes to the absorber-model parameter s, the slope of the column-density distribution, have a strong effect on the expected maximum value of $T(\mathrm{He\,I})$ (as pointed out by Moller & Jakobsen 1990, and illustrated by the dashed lines in Fig. 4.3). In the next subsections we discuss what could be learned about the $z \simeq 5$ ionizing background under the assumption that SDSS will discover a quasar with properties given in Section 4.6 along a LOS with $T(\mathrm{He\,I}) = 0.03$, our 'expected' value for the best SDSS quasar, or $T(\mathrm{He\,I}) = 0.1$, which we consider to be a reasonable 'optimistic' value for the best SDSS quasar.

4.7.2 Quasar absorption line spectrum simulations

Based on the results of the previous subsection, we analyzed lines-of-sight with $T(\mathrm{He\,I}) = 0.03$ and 0.1, our 'expected' and 'optimistic' values for the best SDSS quasar LOS. For each observational realization of a quasar along the simulated LOS, a value of η was assumed. Then the optical depth data were used to construct the IGM transmission as a function of wavelength. This was multiplied by the intrinsic quasar spectrum from the model described in Section 4.6 to generate

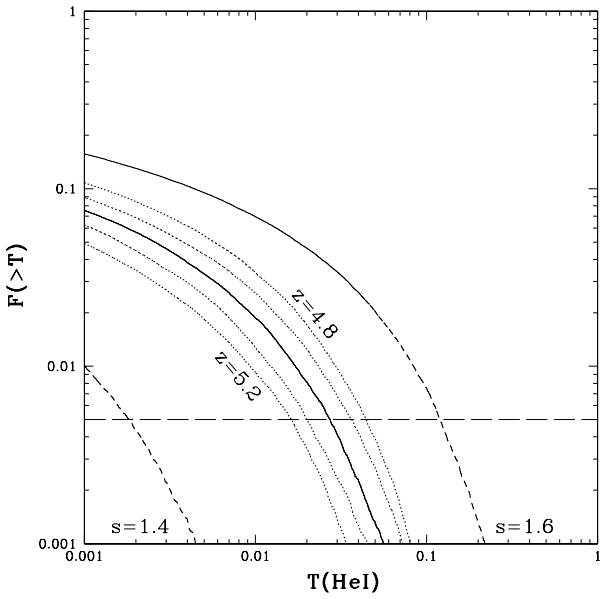


Figure 4.3 Cumulative histogram of the He I-band transmission along quasar lines-of-sight, as a function of redshift and absorber model. The solid curve shows the fraction of $z = 5$ quasars that have a mean transmission in the He I band greater than a given mean transmission value $T(\mathrm{He\,I})$, assuming $s = 1.5$ in the distribution of absorber column densities (Section 4.5). The lowest dashed curves are for $s = 1.4$ and $s = 1.6$, respectively, assuming $z = 5$ quasars. The two dotted curves closely above the solid curve are for $z = 4.9, 4.8$, and the two dotted curves closely below the solid curve are for $z = 5.1, 5.2$, all assuming $s = 1.5$. The horizontal dashed line at 1/200 intersects the curves at the expected $T(\mathrm{He\,I})$ of the best quasar in a sample of 200 quasars.

the observed spectrum. The Galactic extinction curve was taken from Scheffler & Elsaesser (1987). Atmospheric extinction was applied based on the sky transparency from Mauna Kea (the location of the Keck Observatory).[8] Our modelling of the telescope and instrument parameters included slit losses, system throughput, instrumental resolution, pixel sampling, spatial extraction of the spectra, and detector noise. Slit losses and spatial extraction both incorporate an assumed value for the seeing. A mean emission spectrum of the sky,[9] measured at Mauna Kea, was also passed through the telescope/instrument model.

Our model observations were performed assuming a Galactic extinction value of $A_{u'} = 0.15$ toward the target quasar, typical for SDSS quasars (Schneider et al., 2002a). We assumed the observations were made at a constant value of 1.2 airmasses and a constant seeing of 0.7 arcsec.

Our simulated observations of the H I Lyα forest were based on the properties of the Keck II 10-meter telescope and the Echellette Spectrograph and Imager (ESI; Sheinis et al. 2002), an intermediate resolution optical spectrograph, with ESI in echelle mode ($R \sim 4000$). We assumed a 1 arcsec slit, with 75 km s^{-1} resolution and 11.5 km s^{-1}pixel^{-1}. The spatial extraction window was 0.77 arcsec; the spatial pixel scale was 0.154 arcsec pixel^{-1}. We assumed no dark current, and a readnoise of 2.1 counts per pixel (the gain was 1.3 e$^-$ per count). We assumed each H I Lyα forest spectrum consisted of 40 co-added 1000 sec observations, for a total exposure time of 11 hours. This is similar to the exposure times of Keck observations of the highest-redshift SDSS quasars (White et al., 2003).

We assumed that the He I band was observed with the Keck I 10-meter telescope and the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995), using the blue channel of LRIS (LRIS-B; McCarthy et al. 1998) with the 1200 line grism. This configuration with a 0.7 arcsec slit delivers a resolution of 1.30 Å, sampled at 0.24 Å pixel^{-1}. The spatial extraction window was 1.08 arcsec, about 1.5 times the seeing; the spatial pixel scale was 0.2151 arcsec pixel^{-1}. We assumed no dark current, and a readnoise of 2.5 counts per pixel (the gain was 1.6 e$^-$ per count). Because readnoise otherwise made a non-negligible contribution to the noise, we binned the data by 2 pixels in the

[8] http://www2.keck.hawaii.edu/inst/lris/atm_trans.html

[9] http://www.cfht.hawaii.edu/Instruments/ObserverManual/ chapter5.html; note the sky will darken slightly as we approach solar minimum in 2008.

spatial and spectral directions. We assumed each He I 584 Å forest spectrum consisted of 72 co-added 2000 sec observations, for a total exposure time of 40 hours. This is substantially more observation time than is typically devoted to any one object with Keck. However, the total observation time devoted to study He II absorption in the $z = 2.885$ quasar HE2347-4342 was almost 150 hours with the *FUSE* satellite.

For each simulated exposure, we computed the expected signal, expected sky, and expected detector noise. The observed counts in each spectral pixel due to source, sky and detector were drawn from Poisson distributions with the means set to the expected counts, and the three contributions were summed. Then the expected sky and detector noise were subtracted (assuming photon-noise limited sky subtraction). The simulated exposures were then summed to create the final simulated observation.

4.7.3 Quasar absorption line spectrum analysis

Our goal is to estimate the ionization state of the IGM as measured by the relative abundance of He I to H I, η. Our analysis technique is to make use of the cross-correlation of the H I Lyα forest spectrum and the He I-band spectrum, which depends on η as illustrated in Fig. 4.4.

Each pixel in the He I-band spectrum was matched up with the Lyα-band spectrum pixel whose wavelength is closest to $\lambda_{\rm HI} = (1216/584)\lambda_{\rm HeI}$, where $\lambda_{\rm HeI}$ is the wavelength of the He I-band pixel. Thus an absorber at a given redshift will exhibit He I 584 Å absorption and Lyα absorption in matched pixels. Figure 4.5 shows a plot of the fluxes, normalized to the local quasar continuum, of pixel pairs matched in this manner, for two values of η. If He I 584 Å absorption and Lyα absorption were the only features in the spectra, and the observations were perfect, then the points would fall along a tight locus described primarily by η. In regions of little absorption, both continuum-normalized pixel fluxes would be near one: regions of strong absorption would have normalized H I pixel fluxes near zero and normalized He I pixel fluxes set by η. However, the presence of other lines and the effects of instrumental smoothing and observational noise scatter the points. There is little scatter in the H I-pixel-flux direction, both because these simulated observations have a good

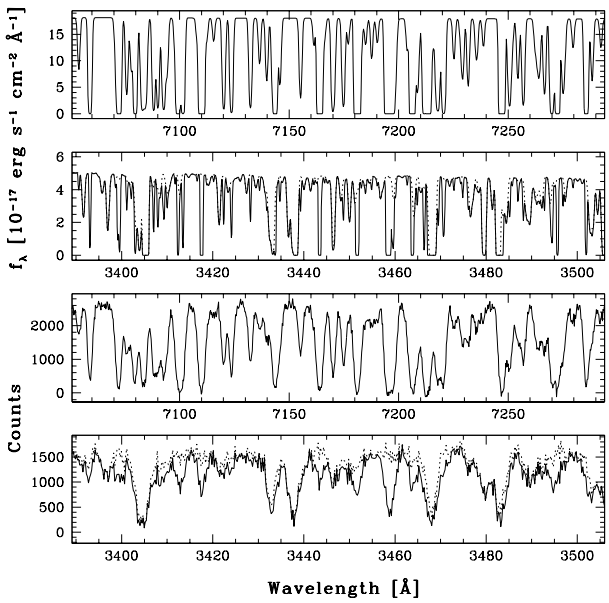


Figure 4.4 A simulated absorption spectrum for a $z = 5$ quasar with $\mathrm{AB}_{1450} = 20$. The top panel zooms in on the $z = 4.8$–5 H I Lyα forest of the spectrum from the middle panel of Fig. 4.1. The second panel shows the corresponding $z = 4.8$–5 He I 584 Å forest region of the same spectrum; the dotted curve is the absorption from H I only, and the solid curve assumes $\eta_{\mathrm{thin}} = 0.5$ for illustration. The lower two panels show mock observations of the quasar, using the observation model described in Section 4.7.2.

signal-to-noise ratio, and because there are no absorption features in the Lyα band besides Lyα. There is substantial scatter in the He I-pixel-flux direction, though, because of lower signal-to-noise in the observations, and also the presence of Lyα absorption (and, to a lesser extent, absorption due to the other Lyman series lines) from low-redshift absorbers. Despite the scatter, one can easily see that, at low values for the normalized H I pixel flux, the points Fig. 4.5 for the $\eta = 0.06$ case (solid squares) lie higher (less He I absorption) than the points for the $\eta = 0.20$ case (open squares).

In the He I-band spectrum, we would like to know what the quasar effective continuum level is after accounting for bound-free absorption by higher-redshift absorbing systems. For a $z_{\mathrm{Q}} = 5$ quasar, any strong absorber between $z = 2.53$ and $z = 2.84$ will cause a change in the effective continuum in the He I band, due to the H I bound-free absorption edge at $912(1+z)$ Å. Consequently, the H I Lyα forest spectrum should be searched for strong Lyα lines over the range of 4292 to 4669 Å (for $z_{\mathrm{Q}} = 5$). We assumed for our analysis that the effective continuum can be accurately estimated. In the analysis of real observations, one could apply the same technique used to estimate the effective

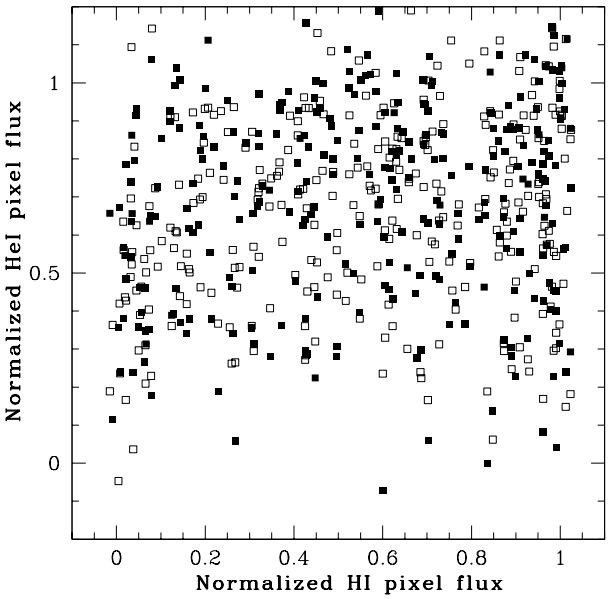


Figure 4.5 An example of the pixel-by-pixel cross-correlation of the absorption in the He I band and the H I Lyα band. The solid squares represent continuum-normalized flux pairs from an observation of a quasar LOS with $\eta = 0.06$; the open squares are from an observation of the same LOS but assuming $\eta = 0.20$.

continuum to our mock observations, and thus analyze the real observations in an unbiased way.

Using our large Monte Carlo library of simulated quasar LOSs, we generated the distribution of pixel pair fluxes, as a function of η, for multiple observational realizations of a quasar along each LOS. Figure 4.6 shows contours that represent the likelihood of finding a given pair of continuum-normalized fluxes for $\eta = 0.06$ and $\eta = 0.20$. This distribution of mock-observational results were then used to analyze an individual observation of a quasar LOS: we generated the matched pair data as shown in Fig. 4.5, then summed the value of the likelihood at that point over all the pixel pairs with continuum-normalized H I pixel fluxes less than 0.5 (cf. Fig. 4.6), for each of the two values of η under consideration. We termed each sum of the likelihood values (which depend on η) $L(\eta)$. The $L(\eta)$ value is larger when a set of pairs matches up well with the contours of Fig. 4.6 for η, and smaller when there is not a good match. Then we formed the ratio $R \equiv L(0.06)/L(0.20)$, which we find is better for recovering our input value of η than, for example, $L(0.06) - L(0.20)$. We have found that binning the He I-band spectra by 2 pixels (in addition to the 2×2 on-chip binning) reduces the scatter in R.

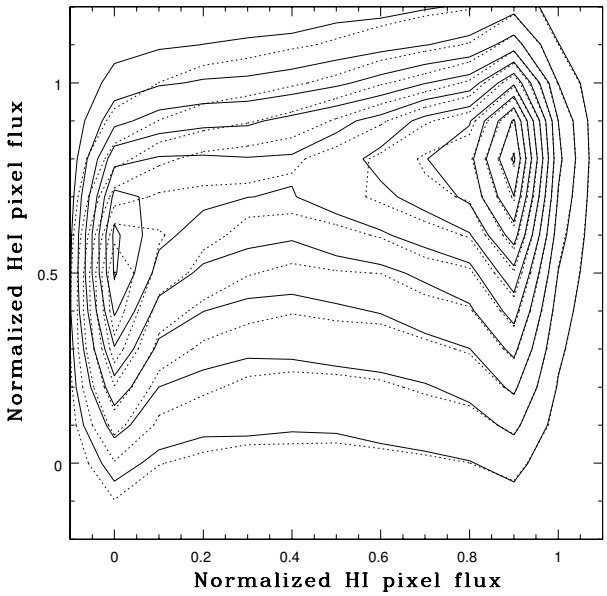


Figure 4.6 Contour plot of the ensemble continuum-normalized fluxes of pairs of corresponding He I-band and Lyα-band pixels (see Section 4.7.3). The solid contours assume $\eta = 0.06$, and the dotted contours are for $\eta = 0.20$. Pixel pairs with strong H I absorption have H I pixel fluxes near zero; for those pairs increasing η shifts the He I pixels to smaller fluxes, due to increased He I absorption associated with the H I absorption.

Figure 4.7 shows histograms of R for each of our two input values of η, assuming observations described in Section 4.7.2 of a quasar along a sightline with $T(\mathrm{He\,I}) = 0.03$, our *expected* value for the best SDSS LOS toward a $z \simeq 5$ quasar (see Section 4.7.1). The histograms are clearly separated, though they do show some overlap. Figure 4.8 is the same as Fig. 4.7, but for observations along a LOS with $T(\mathrm{He\,I}) = 0.1$, our *optimistic* value for the best quasar LOS in the SDSS. In this case the histograms are much more clearly separated, illustrating the importance of discovering a very good quasar LOS.

Figures 4.9 and 4.10 cumulate (and normalize) the histograms of Figs. 4.7 and 4.8. They illustrate that, for our model observations toward one $T(\mathrm{He\,I}) = 0.03$ LOS, if the true value of η is 0.06, then we have almost a 60 per cent chance of making an observation that will reject $\eta = 0.20$ at the 95 per cent confidence level. Conversely, if the true value of η is 0.20, then we have a 60 per cent chance of making an observation that will reject $\eta = 0.06$ at the 95 per cent confidence level. The situation along a $T(\mathrm{He\,I}) = 0.1$ LOS is much more optimistic: if the true value of η is either

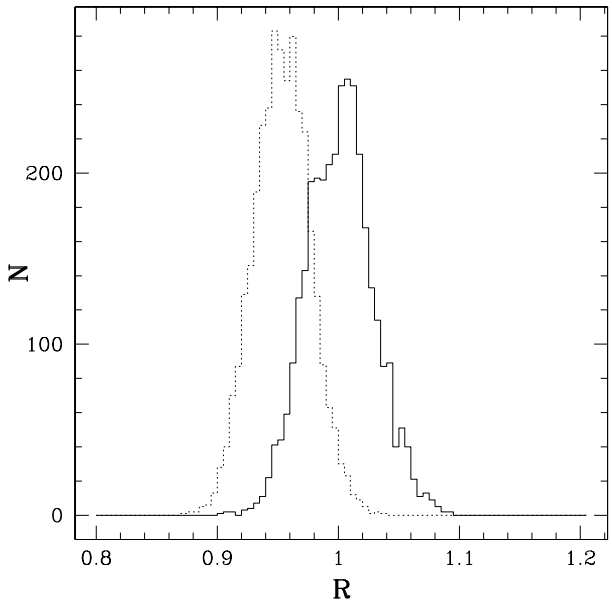


Figure 4.7 Histograms of R (see Section 4.7.3) for each of our two input values of η, 0.06 and 0.20, assuming observation of a quasar along a sightline with $T(\mathrm{He\,I}) = 0.03$. The solid (dotted) line shows the R histogram for analysis of quasar LOSs with $\eta = 0.06$ ($\eta = 0.20$).

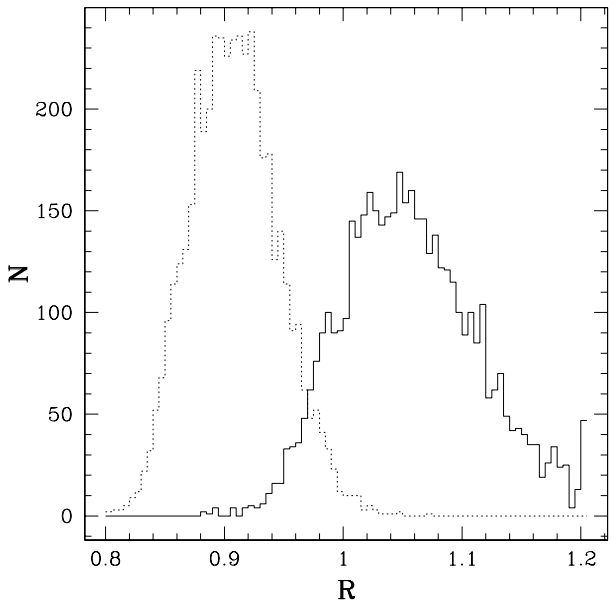


Figure 4.8 Same as Fig. 4.7 but for an LOS with $T(\mathrm{He\,I}) = 0.1$.

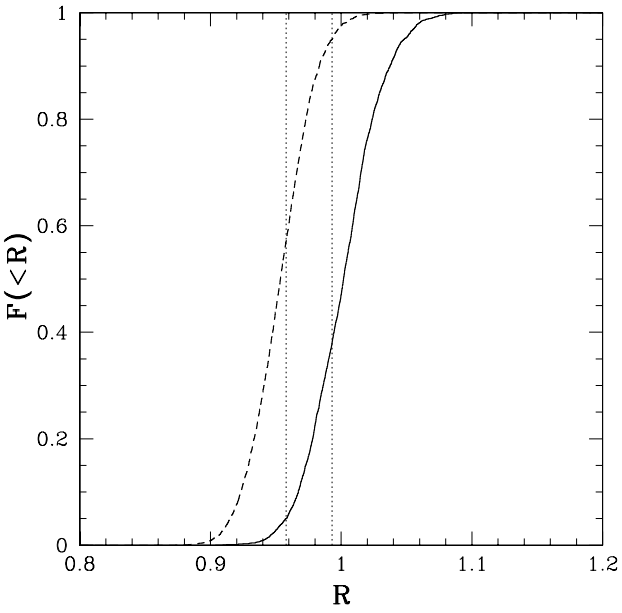


Figure 4.9 Normalized cumulative histograms of R from Fig. 4.7; 95 per cent of the LOSs have $R(0.06) > 0.958$, the position of the left vertical line, and 95 per cent of the LOSs have $R(0.20) < 0.993$, the position of the right vertical line. Thus if the intrinsic η value along a LOS is either 0.06 or 0.20, one observation has about a 60 per cent chance of rejecting the other value.

one of our model choices, we will reject the other value of η about 95 per cent of the time at the 95 per cent confidence level.

Using the observations we propose, it will be difficult to estimate a precise value of η from the data. If the true value of η is, for example, 0.11, then we will only be able to reject very extreme values of η. However, if the ionizing background is either hard or soft, it may well be possible to reject the other hypothesis. For example, if we measured an R value of 1, we could be confident that the ionizing background at $z \simeq 5$ is not dominated by Pop II stars.

4.8 Observational quasar selection techniques

In previous sections we estimated the likely properties of the best SDSS quasar for the measurement of He I absorption features, and what we could learn about the $z \simeq 5$ ionizing background from analysis of such a quasar. In this section we describe the final aspect of practical implementation, how to pick out the best SDSS quasar from the expected sample of 200.

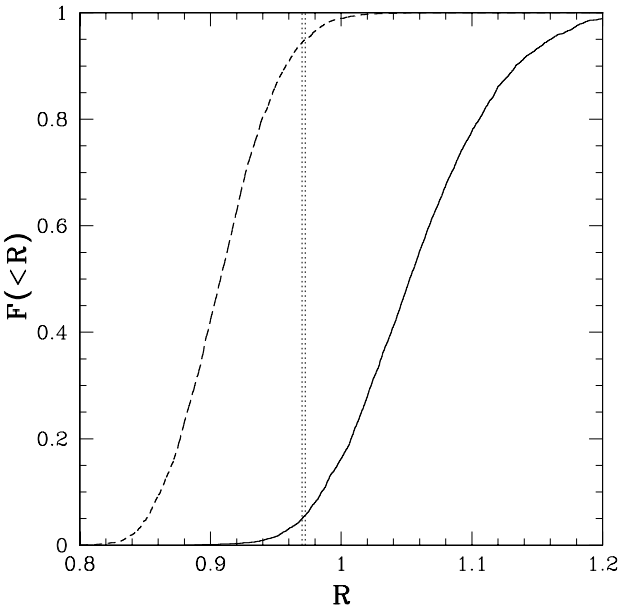


Figure 4.10 Normalized cumulative histograms of R from Fig. 4.8; 95 per cent of the LOSs have $R(0.06) > 0.970$, the position of the left vertical line, and 95 per cent of the LOSs have $R(0.20) < 0.972$, the position of the right vertical line. Thus if the intrinsic η value along a LOS is either 0.06 or 0.20, one observation has about a 95 per cent chance of rejecting the other value.

We would like to select the $z \simeq 5$ SDSS quasar with the highest flux in the He I band. So far we have primarily discussed $T(\mathrm{He\,I})$, the IGM-transmitted fraction in the He I band; however, this was always under the assumption of the intrinsic quasar model described in Section 4.6. The real SDSS quasar sample will certainly include quasars with a range of values for AB_{1450}, and these quasars may have a range of EUV spectral shapes. Due to these distributions, the quasar with the largest He I-band flux may not be the quasar with the largest $T(\mathrm{He\,I})$; we made conservative assumptions about AB_{1450} and the EUV spectrum, so we expect that the true distributions of quasar properties may only improve the suitability of the best SDSS quasar over our estimates.

For $z_{\mathrm{Q}} = 5$, the He I band covers 3222 to 3506 Å; this is within the SDSS u' filter, which runs from $\sim$ 3250 to 3750 Å (Fan et al., 2001a). As a consequence we expect the u' magnitude to serve as a good proxy for the He I-band flux. Figure 4.11 shows the relationship between $T(\mathrm{He\,I})$ and $T(u')$, the transmitted flux fraction in the u' band. The strong correlation illustrated in the scatter plot is independent of the intrinsic quasar properties. Assuming our standard quasar model, the transmitted fractions are directly related to fluxes; the cumulative histogram at the bottom of Fig.

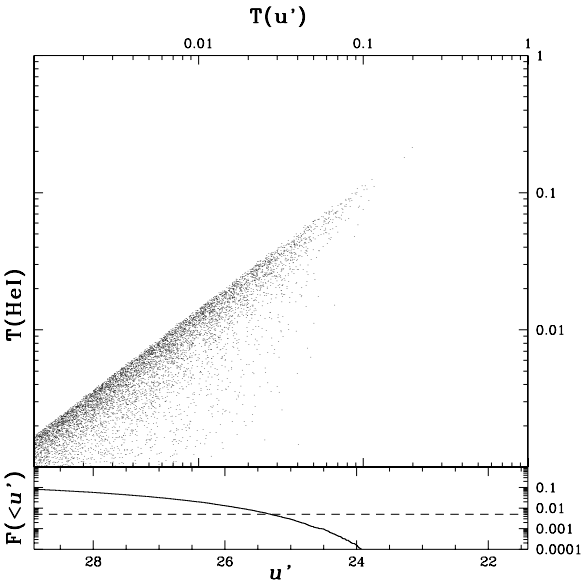


Figure 4.11 Plot of $T(\mathrm{He\,I})$ versus $T(u')$. This plot shows the mean transmissions in the He I band and in the u' band for each $z = 5$ quasar LOS in 10^5 LOS Monte Carlo simulations. The excellent correlation between $T(\mathrm{He\,I})$ and $T(u')$ at the high $T(\mathrm{He\,I})$ end of the distribution shows that u'-band photometry is an effective method for selecting the best quasar for follow-up He I-band spectroscopy. Under the assumption that the intrinsic quasar spectrum is given by the model of Section 4.6, we convert $T(u')$ into a u'-band AB magnitude, and plot the cumulative fraction of quasars brighter than that magnitude in the bottom panel. The horizontal dashed line at 1/200 intersects the curve at $u' = 25.3$, the expected u' magnitude of the brightest quasar in a 200 quasar sample.

4.11 shows the fraction of observed quasars brighter than a given value of u'. We expect that 95 per cent of SDSS $z \simeq 5$ quasars will be fainter than $u' = 27$, but the quasar with the highest He I-band flux will have $u' \sim 25.5$, with an optimistic value of $u' = 24$. The u'-band 1-σ limiting magnitude for the SDSS survey is expected to be about $u' = 24$ (Fan et al., 2001a), thus under optimistic assumptions there is some chance the SDSS survey itself would make a u' detection of a $z \simeq 5$ quasar, and thus identify an excellent candidate for spectroscopic follow-up observations. It is likely, however, that deeper u'-band photometry on the SDSS quasars will be desirable to locate the few brightest quasars in the u' band. The 2.5-meter SDSS telescope integrates for only 54.1 seconds per filter, so achieving deeper u' photometry on a small telescope equipped with a blue-sensitive CCD should not be difficult (extending the 1-σ limit to $u' = 25.5$ would only require 2.5 minutes per source with the SDSS telescope, or about 8 hours of total integration time for all 200 SDSS $z \simeq 5$ quasars).

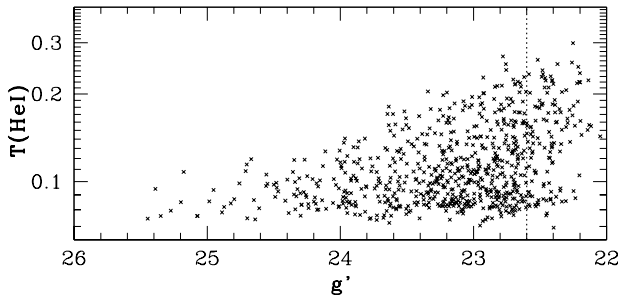


Figure 4.12 Plot of $T(\mathrm{He\,I})$ versus g' of quasars. These points are for the standard intrinsic quasar flux model of Section 4.6. Only quasars with $g' > 22.60$ (plotted to the left of the dotted line) are identified by the SDSS quasar color selection algorithm as quasars. Note that this plot focuses on quasars with very optimistic values of $T(\mathrm{He\,I}) \gtrsim 0.1$; quasars with lower values of $T(\mathrm{He\,I})$ generally all fall at $g' > 22.60$.

Once candidates have been identified from u'-band photometry, low-resolution spectroscopy should be performed to determine the flux in the He I band; because the u' band extends longward of the He I band, a quasar can be relatively bright in u' but faint in the He I band (note the points scattered below the main correlation in Fig. 4.11). The low-resolution spectroscopy will require a blue-sensitive spectrograph, but on only a 2-meter class telescope.

We now turn to our final point: the SDSS quasars are selected based on photometry in all of the SDSS bands. The primary SDSS photometric-selection criteria for $z \simeq 5$ quasars require $u' > 22.00$ and $g' > 22.60$ (Fan et al., 2001a). As illustrated in Fig. 4.11, we do not expect any $z \simeq 5$ quasars to be nearly as bright as $u' = 22.00$. In contrast, Fig. 4.12 shows that quasars with $T(\mathrm{He\,I}) = 0.1$ will sometimes be slightly brighter than 22.60 in g'. Though unlikely, it is possible that the SDSS quasar selection algorithm could miss not only a $z \simeq 5$ quasar, but the best one in the survey for our purposes. When the SDSS quasar survey is complete, and the SDSS data are publicly released, a full search of the SDSS catalog can be made to determine whether any objects meet all the SDSS $z \simeq 5$ quasar criteria except for $g' > 22.60$, but still have, e.g., $g' > 22.40$. These objects could be followed up independently.

4.9 Summary

We have developed a method to infer the dominant stellar population responsible for the ionizing background radiation at $z \simeq 5$, using absorption lines in the spectrum of a $z \simeq 5$ quasar. Specifically, we related the strength of relative absorption in He I lines compared to H I lines to whether Pop II or Pop III stars dominated the ionizing background shortly after reionization completed.

First we related the ionizing background shape to the ratio, η, of He I to H I in an optically thin absorbing system; the spectrum of Pop II stars gives rise to $\eta = 0.20$ while the spectrum of very massive Pop III stars yields $\eta = 0.06$. Then we adopted a model for the population of IGM absorbers based on absorption line observations of quasars. The existing observations do not tightly constrain absorber populations at $z \gtrsim 4$; our model does extrapolate to a neutral fraction of 1.4×10^{-2} at $z = 6.28$, consistent with measurements. We adopted a model quasar spectrum to shine through our model IGM: we assumed the quasar flux to be at the faint limit for the SDSS quasar survey (20th magnitude in the AB system at rest-frame 1450 Å), and modelled the EUV spectral shape based on the composite spectrum of lower-redshift quasars with a spectral slope of $\alpha_{\rm EUV} = 1.6$.

We extrapolated the early results of the SDSS quasar survey to find that SDSS will discover 200 quasars with $4.8 < z < 5.2$. We used Monte Carlo simulations of our IGM model to generate random realizations of lines-of-sight (LOSs) toward $z_{\rm Q} \simeq 5$ quasars. Based on those results, we concluded that a sample of 200 quasars should contain one quasar along a line-of-sight with 3 per cent transmission by the IGM in the He I band ($537(1+z)$ to $584(1+z)$ Å); given the uncertainty in the absorber model, we consider a LOS with 10 per cent transmission to be an optimistic possibility.

We simulated large numbers of quasar LOSs with 3 and 10 per cent transmission. We placed a model quasar at $z = 5$, and then simulated observations of the He I and H I forests. Based on the result for a large ensemble of observations, we derived an estimation technique that can discriminate (at 95 per cent confidence) a Pop II ionizing background from a Pop III ionizing background $\sim$ 50 per cent of the time for observations along our *expected* best LOS, and that can discriminate the backgrounds 95 per cent of the time for observations along our *optimistic* best LOS. These observations require very long integrations with a 10-meter class telescope instrumented with a

blue-sensitive low-dispersion spectrograph and an optical intermediate-resolution spectrograph.

Finally, we examined the selection of the best quasar/LOS combination on which to perform our proposed experiment from the expected 200 SDSS quasars at $z \simeq 5$. The He I forest at $z = 5$ falls in the u' filter; consequently u' imaging is an efficient way to select the best quasar/LOS combination. However, the SDSS u' photometry will not be deep enough to detect $z \simeq 5$ quasars (any detection would be beyond the optimistic expectation of our IGM model), so follow-up u' photometry of the 200 $z \simeq 5$ quasars will be required; these observations will be inexpensive. We also consider the possibility that a quasar/LOS combination will be so good for our purposes that it falls outside of the SDSS quasar color selection criteria. This is very unlikely for quasars with ~ 10 per cent or lower transmission, but about half of all quasars with 20 per cent transmission, should any exist, would be brighter in g' than the SDSS quasar color selection limit of 22.6. If quasars are discovered with 10 per cent transmission, it may be worthwhile to modify the SDSS quasar color selection criteria so as to discover quasars with even less IGM absorption in the He I forest.

Acknowledgments

MRS thanks the TA group at Harvard CfA for warm hospitality. We gratefully acknowledge helpful conversations with Wal Sargent, Rob Simcoe, Chuck Steidel, Peng Oh and Marc Kamionkowski. This work was supported in part by NASA GSRP grant NGT5-50339 and NASA grant NAG5-9821 (for MRS), and by NASA grant ATP02-0004-0093 and NSF grants AST-0071019 and AST-0204514 (for AL). AL acknowledges support from the Institute for Advanced Study and the John Simon Guggenheim Memorial Fellowship.

Chapter 5

The Abundance of Low-luminosity Lyman α Emitters at High Redshift[1]

5.1 Abstract

We derive the luminosity function of high-redshift Lyman α emitting sources from a deep, blind, spectroscopic survey that utilized strong-lensing magnification by intermediate-redshift clusters of galaxies. After observing sky near 9 clusters where the magnification factor was > 10 over most of the survey area, we located 12 emission line candidates in the range $2.2< z <6.3$ whose identification we justify as Lyman α, in most cases via further spectroscopic observations. We construct the selection function of our survey taking into account our varying intrinsic Lyman α line sensitivity as a function of wavelength and position on the sky. Due to the strong magnification factor over most of our survey, we provide meaningful number density constraints down to unprecedented Lyman α line luminosities of 10^{40} erg s^{-1}, corresponding to a star-formation rate of 0.01 M$_\odot$ yr^{-1}. We present a cumulative $z \simeq 5$ Lyman α luminosity function that is consistent with $n(> L) \propto L^{-1}$ over 10^{41} to $10^{42.5}$ erg s^{-1}. Our results, analyzed in the context of other Lyman α survey data and simple theoretical models, may show direct evidence of the suppression of star formation in low-mass halos, as predicted by theoretical models of galaxy formation.

[1]This chapter appeared as Santos et al. (2004), and is reproduced here by permission of the copyright holder, the American Astronomical Society.

5.2 Introduction

The epoch of cosmic reionization, when the intergalactic hydrogen in the universe transitioned from neutral to ionized, is the current frontier of observational cosmology. QSOs discovered by the Sloan Digital Sky Survey (SDSS) indicate that reionization was just finishing at $z \simeq 6$ (Becker et al., 2001; Djorgovski et al., 2001; Fan et al., 2002). Recent results from the *WMAP* satellite suggest that significant reionization of the universe took place by $z \sim 12$ (Spergel et al., 2003). The sources that reionized the universe, however, are still unknown: at $z \sim 6$ neither bright QSOs discovered by SDSS (Fan et al., 2001a) nor faint AGN from deep x-ray observations (Barger et al., 2003) produce enough photons to reionize the universe. Other evidence from the temperature and ionization state of the intergalactic medium (IGM) suggests that, though QSOs dominate the meta-galactic ionizing background at $z \sim 3$, the spectrum may be softer at reionization (e.g., Sokasian et al., 2003): hot stars may be the dominant source of reionizing photons. One goal of the forthcoming NASA/ESA *James Webb Space Telescope* (*JWST*), a 6-meter IR telescope scheduled for launch in 2010, is to study the formation of the first generations of galaxies and their contribution to reionization (Mather & Stockman, 2000).

Early galaxies played many important roles beyond their involvement with reionization. The IGM was enriched well above the primordial metal abundance by $z = 5$ (Songaila, 2001; Pettini et al., 2003); additional evidence for early metal production comes from metal-poor globular clusters in the Milky Way: age estimates imply a formation epoch of $z \gtrsim 4$ for current cosmological parameters (Krauss & Chaboyer, 2003), but the typical metallicity of these objects is 10^{-2} of the solar value (Harris, 1996). The stars responsible for reionization and early metal production may still be present in some form today. It is an important challenge to identify the transition between the very first, metal-free, stars, and Population II stars, because of the strong constraints on the metallicity of low-mass stars provided by studies of halo stars in the Milky Way. A complete understanding of the metallicity distribution of old Galactic stars will benefit from direct observation of very high redshift star formation in *proto-galactic systems that will evolve into galaxies like the Milky Way.*

In advance of *JWST*, which will use IR capabilities to observe galaxies before the end of reion-

ization in rest-frame UV and optical light, current ground-based facilities have the opportunity to discover and characterize star-forming galaxies near the epoch of reionization with rest-frame UV observations. In particular, the identification of Lyman α emission from star-forming regions of early galaxies has proven to be a powerful tool for discovering $z > 4$ galaxies and measuring their redshifts (see Section 5.3). The redshift range $5 < z < 7$ is of particular interest for two reasons. One is that the very detection of Lyman α emission may place a constraint on the progress of reionization (Rhoads & Malhotra, 2001; Haiman, 2002; Hu et al., 2002a; Rhoads et al., 2003): it is difficult to observe Lyman α emission from galaxies embedded in a neutral IGM, but the strength of the constraint derived from the successful detection of Lyman α depends on the assumed properties of the sources (Santos, 2004a). The second important reason to study galaxy formation during and after reionization is that an intense UV background and 10^4 K IGM is predicted to suppress star-formation in galaxies that form after reionization (Barkana & Loeb, 1999; Gnedin, 2000; Benson et al., 2002b). There is a discrepancy between some theoretical predictions of the abundance of dark matter halos on dwarf-galaxy mass scales and the number of observed dwarf satellites in the Local Group (Moore et al., 1999). Reionization may sterilize many dwarf galaxy-scale halos to star formation, so that the luminous satellites of the Milky Way are dwarf galaxies formed before reionization, and the remaining satellite halos are empty of stars and thus dark (Benson et al., 2002a; Somerville, 2002).

This paper presents the results of a spectroscopic Lyman α emission-line survey for galaxies up to $z = 6.7$ that utilizes the strong-lensing properties of intermediate-redshift clusters to magnify the surveyed regions. In Section 5.3 we review the use of Lyman α surveys as probes of early star formation. Section 5.4 motivates the importance of surveys for low-luminosity galaxies. Section 5.5 describes the advantages of a survey utilizing strong lensing and details our survey strategy, targets, observations, and data reduction. In Section 5.6 our Lyman α emission-line detections are presented. We compute our survey volume and source number density in Section 5.7. Section 5.8 compares the results of our survey to other surveys and theoretical models. In Section 5.9 we summarize.

Throughout this paper we use a ΛCDM cosmological model with $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, \sigma_8)$=$(0.3, 0.7, 0.043, 0.9)$

and $h \equiv H_0/(100 \text{ km s}^{-1}\text{Mpc}^{-1}) = 0.7$; these values are consistent with the values derived in Spergel et al. (2003).

5.3 Lyman α Surveys

The primary appeal of Lyman α emission as a signpost to high-redshift galaxy formation is that it traces star formation at a wavelength that conveniently redshifts into the visible and near-IR, where it is easiest to make sensitive, high-angular resolution observations. The Lyman α emission line may be quite strong, but its luminosity is quite sensitive to physical details of the nature and geometry of the star-forming regions. Because Lyman α emission traces hydrogen recombinations, it is intimately related to the production of ionizing photons by the stars present. Both the initial mass function (IMF) and metallicity of the stars affect their production rate of ionizing photons. However, if the IMF and metallicity are constrained to reasonable ranges, the ionizing photon production rate can be reliably connected to the star-formation rate.

The major complication for the interpretation of Lyman α line strengths is the effect of the nebula around the star-forming region. Hydrogen at low density does not recombine quickly, so, e.g., ionizing photons that escape into the IGM are "lost" for the purposes of producing a Lyman α emission line. Even after a hydrogen recombination produces a Lyman α photon, which happens for about two thirds of the recombinations (Osterbrock, 1989), there are many ways the photon may be destroyed prior to escape. The resonant nature of the Lyman α transition results in a very short mean free path, even in a mostly ionized nebula. Consequently, if dust is mixed in with the gas, then the chance of absorption by a dust grain may be higher for a Lyman α photon than a non-resonantly scattered photon at the same wavelength (but see Neufeld, 1991, for an alternative). The dust content of very high-redshift galaxies is still relatively unconstrained, and will likely remain so at least until *JWST*.

On the positive side, Lyman α is the intrinsically strongest recombination line from an H II region. Another meritorious aspect of Lyman α is that its emission strength does not strongly depend on the metallicity of the ionized gas (the only effect is from the temperature of the photoionized gas,

which depends on metallicity); consequently, it can be used as a tracer for truly primordial star formation, where dust extinction is also believed not to be a problem.

Partridge & Peebles (1967) introduced a model of galaxy formation "to assess the general possibility of observing distant, newly formed, galaxies." In their model they estimated that 6–7% of the luminosity of early galaxies would be emitted in the Lyman α emission line: they predicted line luminosities of 2×10^{45} erg s^{-1} over galaxy formation time-scales of 3×10^{7} yr. The predictions of Partridge & Peebles (1967) and others led to observational surveys for Lyman α emission from high-redshift galaxies, reviewed by Pritchet (1994, §4.5). Pritchet summarized the status of searches at that time with "no emission line primeval galaxies have been found" despite 16 cited surveys covering various redshifts ranges from $z = 2$ to $z = 5$. In striking contrast to these pioneering explorations, many high-redshift Lyman α emission line galaxies have been discovered and confirmed in the past 8 years. Stern & Spinrad (1999, §4) and Taniguchi et al. (2003) have reviewed recent progress, and we provide a brief summary here.

The search technique that has dominated recent success in the discovery of large numbers of Lyman α emission-line galaxies at $z > 4$ is narrow-band photometry. This approach uses a narrow (~ 100 Å) filter chosen to lie in a spectral region of low sky background; such surveys cover relatively large areas of sky with sensitivity to Lyman α emission over a small window in redshift, $\Delta z \sim 0.1$. Many groups have now performed successful blind narrow-band surveys for $z > 4$ galaxies: Hu, Cowie, & McMahon (1998) and Rhoads et al. (2000) at $z = 4.5$; Ouchi et al. (2003) at $z = 4.9$; Hu, McMahon, & Cowie (1999), Rhoads & Malhotra (2001), and Rhoads et al. (2003) at $z = 5.7$; and Hu et al. (2002a) and Kodaira et al. (2003) at $z = 6.5$.

Spectroscopic surveys provide a complementary technique to the narrow-band method. For equal observing time at one position, spectroscopic searches at optical wavelengths can cover a large range in redshift, $\Delta z \sim 3$, to better line flux sensitivity than a corresponding narrow-band survey. However, the area surveyed by a long slit is typically $\sim 5 \times 10^{-2}$ arcmin^2, in contrast to 20–2000 arcmin^2 for an imaging camera. Deep long slit observations of other targets have been searched for serendipitous detection of high-redshift galaxies. These techniques discovered the first confirmed $z > 5$ galaxy

(Dey et al., 1998) and subsequently turned up a few more $z > 4$ sources (Hu et al., 1998; Weymann et al., 1998b). Serendipitous surveys will likely continue to play a role in discovering high-redshift galaxies, since the deepest spectra on the largest telescopes are likely to be pointed observations rather than devoted emission-line surveys.

A few other hybrid techniques combine aspects of the narrow-band imaging and long-slit spectroscopy approaches. Maier et al. (2003) used an imaging Fabry-Perot interferometer to take sequences of narrow band images within the night-sky windows corresponding to $z = 4.8, 5.7, 6.5$, and have confirmed discoveries at $z = 4.8$ and 5.7. Recently Lilly et al. (2003) used a "slit-let" slit mask with a narrow-band filter to do a spectroscopic survey over a relatively large area and narrow redshift window. There is also an ongoing *HST* program to use slitless spectroscopy with the ACS camera grism to discover high-redshift Lyman α emission (J. E. Rhoads, priv. comm.).

5.4 Searches for distant low-luminosity galaxies

In this section we present the motivation for conducting a survey devoted specifically to low-luminosity $z \sim 5$ Lyman α emitting galaxies ($\lesssim 10^{42}$ erg s^{-1} in the Lyman α line), an unexplored region of survey parameter space.

At $z \sim 5$, the luminous Lyman α galaxies (§5.3) and QSOs (e.g., Fan et al., 1999, 2000, 2001a) discovered so far represent the rarest and most spectacular tail of the range of structure formation scales (e.g., Barkana & Loeb 2003 suggested that the high-z SDSS QSOs reside in 10^{12} M$_\odot$ virialized halos). They almost certainly evolve into the rarest and most massive environments in the local universe: rich clusters of galaxies. In striking contrast, the characteristic mass of virialized halos at $z \sim 5$ is only 10^9 M$_\odot$; if such an object steadily converts its $\sim 10^8$ M$_\odot$ of baryons into stars, it will have a star-formation rate of only ~ 0.1 M$_\odot$ yr^{-1}.

These objects, which we will refer to as low-mass halos, would not be detectable in any of the surveys cited above, but our understanding of galaxy formation depends crucially on constraining their properties for three important reasons: First, they represent the most common environment *by mass* of virialized halos (the peak of the mass-weighted mass function is always near the characteristic

mass); consequently, if low-mass halos form stars efficiently, they could dominate the star-formation rate at high redshift. Second, they are the progenitors of common galaxies in poor environments, like the Milky Way, under the current paradigm of structure formation. The detection of low-mass sources is a direct test of the "bottom-up" description of galaxy assembly. Third, low-mass objects have a unique link to the IGM: kinetic energy injected into the IGM by photoionization at $z \gtrsim 6$ is expected to raise the cosmic Jeans mass and inhibit gas cooling; this effect has no consequence for the brightest sources residing in deep potential wells, but may heavily suppress star formation in $10^9\,M_\odot$ objects (e.g., Barkana & Loeb, 1999; Gnedin, 2000). This Jeans-mass effect has been cited as the solution for the Cold Dark Matter "crisis" of over-predicting the number of Milky Way satellite galaxies compared to observation (Benson et al., 2002a; Somerville, 2002). Low-mass halos also place relevant constraints on the energy scale of dark matter in Warm Dark Matter models (Z. Haiman, priv. comm.).

A practical and strategic advantage in characterizing the luminosity function at low luminosities is to determine the optimum survey depth for future surveys that aim to discover large numbers of $4 \lesssim z \lesssim 7$ galaxies. If the luminosity function were very steep, then deep surveys such as ours would be more efficient than shallower, wider field surveys. Theoretical prejudice suggests that the luminosity function should have a steep effective slope in the region associated with the exponential cut-off of the number density of underlying halos, assuming some sort of mass-to-luminosity correspondence. Previous non-detections combined with recent successes seem to bear this out (Pritchet, 1994, and see below). That is, current surveys may be approaching the characteristic luminosity; however, the characteristic luminosity and luminosity function shape has yet to be well constrained. In the local universe, luminosity functions based on star-formation rate estimators such as Hα luminosity roughly follow the Schechter function (Schechter, 1976) form of power-law behavior at low luminosity, reflecting the underlying power-law of the mass function, albeit with possibly a different slope (e.g., Gallego et al., 1995).

The predicted suppression of star formation in halos with small potential wells suggests that the Lyman α luminosity function at $z \sim 5$ may have a modified shape. The reason is that the

characteristic mass scale where the exponential tail and power law regions of the mass scale meet, $\sim 10^9$ $M_\odot$, corresponds quite closely to the mass scale where a number of physical mechanisms may suppress star formation. We described the effect of a hot IGM above. In addition, energy injected in the ISM of star-forming galaxies by stellar winds and supernovae (called "negative feedback," or just "feedback") is predicted to heavily suppress star formation in halos with circular velocities below about 100 km s^{-1} (Dekel & Silk, 1986), corresponding to a mass scale at $z \sim 5$ of $\sim 10^{11}$ $M_\odot$. A current implementation of feedback by Benson et al. (2002b) shows the importance of feedback on the high-redshift star-formation rate. This effect complements the inhibiting effects of a hot IGM on star formation in low-mass halos. A third mechanism that may reduce star formation preferentially in low-mass halos is the effect of large-scale winds blown by star-forming galaxies. Scannapieco et al. (2000) and Scannapieco & Broadhurst (2001) computed the influence of winds blown out through the IGM by the first galaxies to form. They concluded that these winds effectively sweep gas out of nearby halos in the process of collapsing, meaning that even though the dark matter continues its collapse to virialization, there is little corresponding star formation because of the lack of baryons. In their model winds influence the entire star-formation history of the universe, but at $z \sim 5$ particularly suppress star-formation in halos smaller than $\sim 10^{10}$ $M_\odot$.

Figure 5.1 is a schematic illustration of the possible effect of the suppression of star formation in low-mass halos. The solid curve is the mass function of halos at $z = 5$ (Sheth & Tormen, 2002), and we have converted the mass scale on the top axis into a star-formation rate on the bottom axis using a simple prescription (see Section 5.8.2). In this simple model the Lyman α line luminosity function, if Lyman α luminosity is proportional to SFR, would have the shape of the solid curve. We have introduced the suppression of star formation in low mass halos by using the prescription of Gnedin (2000) to efficiently filter out gas from halos below a critical mass scale $M_{\rm F}$, the filtering mass (Gnedin, 2000). Each of the broken curves is the shape of the Lyman α line luminosity function we expect (again assuming it scales with SFR) after applying filtering on a different mass scale; the labels are the $\log_{10}$ of the filtering mass. Ultimately the filtering mass, and thus the physics described in the previous paragraph, may be constrained directly by a measurement of the shape

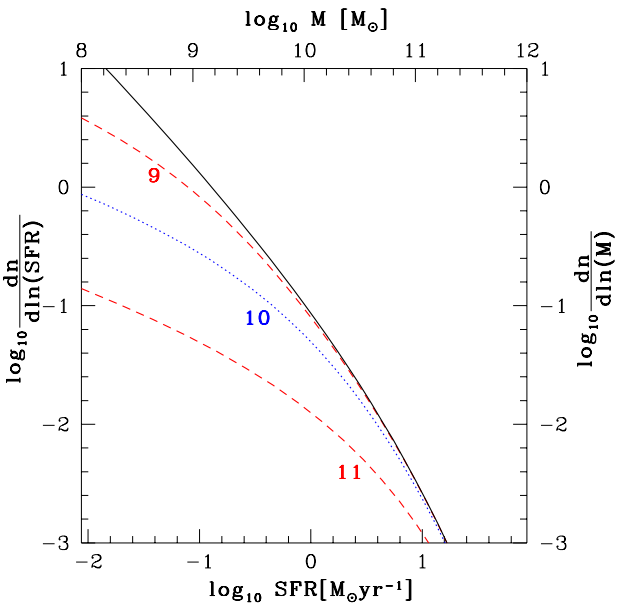


Figure 5.1 The star-formation rate function based on a dark matter halo model. The solid curve is the abundance (on the right axis) of halos at $z = 5$ as a function of the mass on the top axis. The bottom axis is a simple conversion of the halo mass into the expected star formation in that halo, so that the solid curve may also be read as a theoretical star-formation rate function using the bottom and left axes. The broken curves are also star-formation rate functions, but in each case star-formation in low-mass halos has been suppressed. The broken curves are labeled with the $\log_{10} M_{\rm F}$, where $M_{\rm F}$ is the halo mass scale below which star formation is suppressed. See Sections 5.4 and 5.8.2 for details.

of the star-formation rate function, along with estimates of the corresponding halo masses (see Section 5.8.2).

The current observational limit is just at the threshold of $z \sim 5$ Lyman α galaxy *detection*; the detailed form of the luminosity function will not be well-constrained in the immediate future. However, our low-luminosity Lyman α survey, and others like it, in concert with surveys at higher luminosities, may constrain or detect the break in the luminosity function associated with the characteristic halo mass and where star formation is suppressed.

5.5 Observations

5.5.1 Survey stategy

The goal of our survey is to extend the census of Lyman α sources at $z \sim 5$ to the faintest luminosities possible with existing observational facilities. We achieved this through deep spectroscopic exposures on areas of sky strongly magnified by gravitational lensing.

Strong gravitational lensing by rich clusters of galaxies at $z \sim 0.2$ is an invaluable resource to a survey covering a very small area of sky to great depth, as such clusters magnify background high-redshift sources by greater than a factor of 10 over regions of roughly 0.1 arcmin^2. This advantage comes at a price: lensing increases the apparent area of a background source at fixed surface brightness, so the pointing is deeper by a factor of the magnification, $\mathcal{M}$, but covers only $1/\mathcal{M}$ of the area of an equivalent unlensed pointing (assuming unresolved sources). This is a superior strategy for a deep, narrow survey. Achieving the same depth and area in an unlensed survey would require a factor of $\mathcal{M}$ more observing time, a huge difference for $\mathcal{M} \sim 10$.

Only a small area of the whole sky, ~ 100 arcmin^2, is magnified by a factor of 10 or more by clusters; this sets a fundamental limit to the areal coverage of any survey utilizing strong lensing by clusters. For a large survey that detected many sources, statistical information about the distribution of magnification over the survey area might be sufficient to construct a Lyman α emitter luminosity function. However, since we planned to survey a small area and detect only a few sources, we chose to observe clusters with *HST* imaging and spectroscopic redshifts for many of the identified arcs and multiple images (e.g., Kneib et al., 1996). These are currently available for only a small fraction of strong-lensing clusters. The positions of the identified arcs constrain the distribution of mass in the cluster which can, in turn, be used to predict a magnification map of the cluster for high-redshift sources. Since lensing depends on the angular-diameter distance between the source and the lens, and that distance changes slowly with redshift at $z \sim 5$, the sky area with large magnification is fairly independent of the source redshift for $z > 3$

Given that high-quality lens models are only available for about a dozen clusters observable from

Hawaii, the total area available to us for a survey is currently quite small. To make the most of this limited resource, we conducted a spectroscopic survey. The advantage of a spectroscopic survey was that we simultaneously surveyed for Lyman α emission over the redshift range $3 \lesssim z \lesssim 7$. The primary draw-back of a spectroscopic survey, as mentioned above, was that with the instrumentation available the most efficient technique was slit spectroscopy. The geometry of a long slit is not well matched to the lensed region of sky (see Figs. 5.2 and 5.3), thus some of the slit area covers area outside the cluster that is not strongly magnified.

A long-slit survey does have other advantages. We can expect many emission-line sources in our survey other than Lyman α; in particular, optical lines associated with strong star formation, such as [O II] 3726, 3729 Å; Hβ 4861 Å; [O III] 4959, 5007 Å; and Hα 6563 Å. Low-resolution spectroscopy with large wavelength coverage allows the rejection of many potential low-redshift contaminants through the identification of other emission lines. However, the [O II] doublet can be difficult to resolve at low dispersion and, if redshifted to $z \sim 1$, there are often no other strong emission lines present in the optical spectrum. Thus final identification of an emission line as Lyman α may require follow-up spectroscopy at intermediate resolution.

Redshift identification is aided by two additional factors. Deep optical broadband imaging, available in at least one band for well-studied clusters, can be used, as in narrow-band searches, as a rejection filter: if a putative Lyman α system shows much observable flux shortward of Lyman α, then it is not likely a correct line identification, because the intrinsic UV spectrum combined with IGM absorption create a strong decrement across the Lyman α emission line (e.g., Songaila & Cowie, 2002). The second tool is available when two or more images (due to strong lensing) of the same high-redshift source are discovered. In this case the lensing model itself may place a reasonably strong constraint on the redshift of the system based on the observed image positions and flux ratios (Kneib et al., 1996).

The deepest survey for a given observing time would be to devote all of the time to a single slit position. However, we expect Lyman α sources to be clustered, resulting in a non-Poisson distribution. To estimate an accurate luminosity function we surveyed several independent volumes

Table 5.1 Clusters surveyed

Cluster	Redshift	RA	Dec	Lens Model Reference
Abell 68	0.255	00h36m59s	+09d09m	(1)
Abell 370	0.375	02h37m18s	-01d48m	(2)
Abell 773	0.217	09h14m30s	+51d55m	(1)
Abell 963	0.206	10h17m09s	+39d01m	(1)
Abell 1689	0.183	13h09m00s	-01d06m	(3)
ZwCl 1358.1+62.45	0.328	13h59m54.3s	+62d30m36s	(4)
Abell 2218	0.176	16h35m42s	+66d19m	(5,6)
Abell 2219	0.226	16h38m54s	+46d47m	(1)
Abell 2390	0.228	21h53m35s	+17d40m	(7)

References. — (1) Smith et al. 2003; (2) Bézecourt et al. 1999; (3) J. P. Kneib, unpublished; (4) Franx et al. 1997; (5) Kneib et al. 1996; (6) Ellis et al. 2001; (7) Pelló et al. 1999

(via surveying behind several lensing clusters) to ameliorate cosmic variance and recover the Poisson noise limit (see Section 5.8.2).

5.5.2 Survey parameters

Table 5.1 summarizes the 9 lensing clusters of our survey. We have constructed a detailed mass model for each, based on *HST* imaging and lensed arc redshifts.

Details of our spectroscopic observations are listed in Table 5.2. Clusters observed at multiple locations are further labeled by a number. We used the double-beam Low Resolution Imaging Spectrograph (LRIS, Oke et al., 1995) in long-slit mode on the Keck I 10-meter telescope at Mauna Kea to perform our survey. For the 2000 March observations we used a slit $0.7''$ wide and a spectroscopic range of $\lambda\lambda$ 6800–9500 Å (corresponding to Lyman α with $4.6< z <6.8$), using a 600-line grating blazed at λ7500 Å which gave a resolution of $\simeq$3.0 Å. In 2001 April we switched to a $1.0''$-wide slit and used a 600-line grating blazed at 1 μm over the same wavelength range as above, at a resolution of $\simeq$4.0 Å. Starting in 2001 April we also began using a 300 line grism blazed at 5000 Å and a dichroic at 6800 Å to simultaneously take spectra on the blue arm of the spectrograph, over $\lambda\lambda$ 4000–6700 Å (corresponding to Lyman α from $2.2< z <4.5$) at 3.5–4 Å resolution.

The length of the spectrographic slit was $175''$. We mapped an area on the sky via offsetting the telescope perpendicular to the long axis of the slit by a distance equal to the slit width. At each

Table 5.2 LRIS survey observations

Date	Cluster	Position Angle[a]	Integration time[b]	Photometric?
Mar 2000	Abell 773	-46.8	20	Yes
	Abell 1689 #1	84.1	23	Yes
Apr 2001	Abell 1689 #2	43	10	Yes
	Abell 2218 #1	-44	10	Yes
	ZwCl 1358	-15	12	Yes
Oct 2001	Abell 370 #1	-8	14	Yes
Apr 2002	Abell 963	3.6	14	No
	Abell 2218 #2	-49.2	14	No
May 2002	Abell 1689 #3	12.3	20	No
	Abell 2219	-69	14	No
Sep 2002	Abell 370 #2	1.7	14	No
	Abell 2390	-63	12	No
	Abell 68	-40	12	No

[a] in degrees North through East
[b] in ksec

slit position we made two exposures of 1000 sec to facilitate cosmic ray rejection. Each map was 5–10 adjacent slit positions at the same position angle, giving contiguous survey areas on the sky of 875–1750 arcsec^2.

The pointing of the slit on the sky was verified by registration of images from the LRIS slit-viewing guide camera to the *HST* images of the cluster (because the clusters are rich in bright galaxies, there were always many sources in the slit-viewing guide camera images). Our sequence of slit offsets typically agreed with a regular spacing of the slit width to a precision of $0.1''$ (10% of the $1''$ slit). More importantly, registration of the slit position on the *HST* image enabled us to look for a broadband counterpart at the location of emission lines detected in our spectra.

The areas mapped by the procedure above were chosen to take advantage of the strong magnification of background sources provided by the foreground cluster. The magnification map of the sky around the cluster is constrained by the distribution of visible light in the cluster and the measured velocity dispersion of some cluster members, but is crucially verified and refined by including information from the locations and redshifts of strongly lensed sources. These background galaxies, generally brighter and at lower redshift than the $z > 4.5$ sources we searched for, have been the target of previous observations (e.g., Kneib et al., 1996).

We used the cluster mass models to generate redshift-dependent maps of the magnification toward

$4 < z < 7$ sources (Kneib, 1993). The angular diameter distance between the cluster and those redshifts depends only weakly on redshift, so we were able to choose areas on the sky with high magnification over our entire redshift range of interest.

The geometry of the magnification map was generally characterized by two concentric ring-like curves of formally infinite magnification, called the inner and outer critical lines. These are related to the location in the image plane of the caustic of an elliptical potential (Blandford & Narayan, 1992), but modified by the deviations of the projected lensing potential from an ellipse. The areas of highest magnification are found next to the critical curves, so our survey maps generally follow the outer critical line. The outer critical line is more amenable to long-slit mapping because of its greater length on the sky and its less curved form. However, in one case (Abell 1689) we mapped sky near the inner critical line as well. Lensed sources close to the critical line are often multiply imaged, forming a merging pair on either side of the critical line. We considered this when we chose our map locations, but the irregular shape of the critical line, compared to the straight shape of our slit, limited the extent to which we could map exclusively one side of the critical line.

Figures 5.2 and 5.3 summarize the adopted strategy for each cluster in the context of the location of the critical line for a lensed source at $z = 5$ (dotted lines). In the most massive clusters with the best mass models, such as Abell 1689 and Abell 2218, we made multiple maps (see also Table 5.2). In these cases each survey region is labeled by a number corresponding to the observations listed in Table 5.2.

The total area on the sky covered by our survey was 4.2 arcmin^2. The effective areal coverage of the survey is smaller due to lensing by a spatially variable magnification factor (see Section 5.7.1.1).

5.5.3 Candidate selection and catalog

The 2-D spectroscopic data were reduced using standard techniques in the NOAO/IRAF software environment[2]. Cosmic rays were rejected from each pair of images at a given location with the `L.A.COSMIC` routine (van Dokkum, 2001) and sky emission was removed by subtracting block-filtered

[2]IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

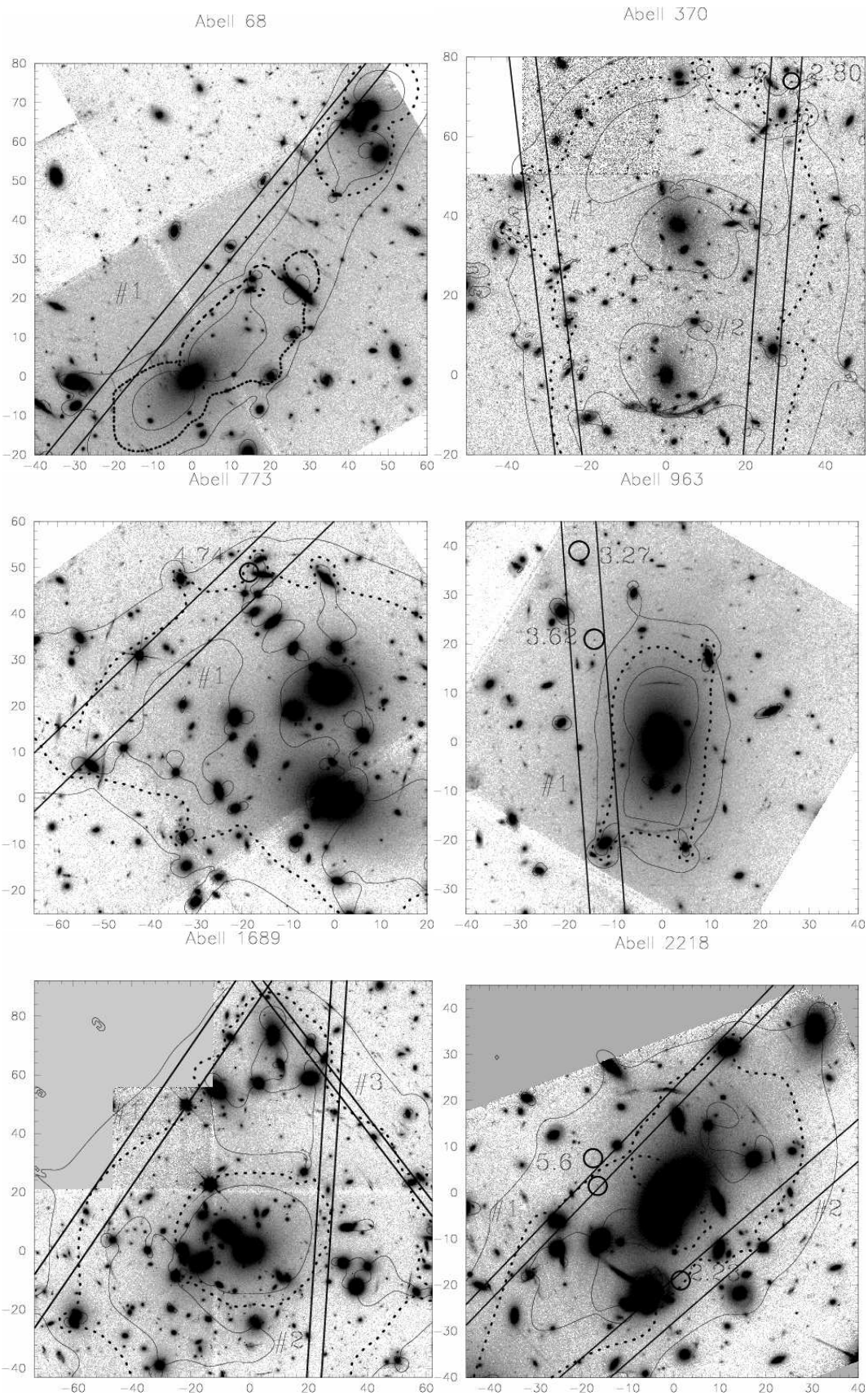


Figure 5.2 Survey clusters, with survey area and lensing critical curves. Superposed on the *HST*/WFPC2 image are the critical line for $z = 5$ (dotted line) and solid curves enclosing the area of sky magnified by a factor of > 10. The regions bounded by parallel straight lines are the area we surveyed; numeric labels correspond to observing run keys in Table 5.2. The axes are labeled in arcseconds. (upper left) Abell 68. (upper right) Abell 370. (middle left) Abell 773. (middle right) Abell 963. (lower left) Abell 1689. (lower right) Abell 2218.

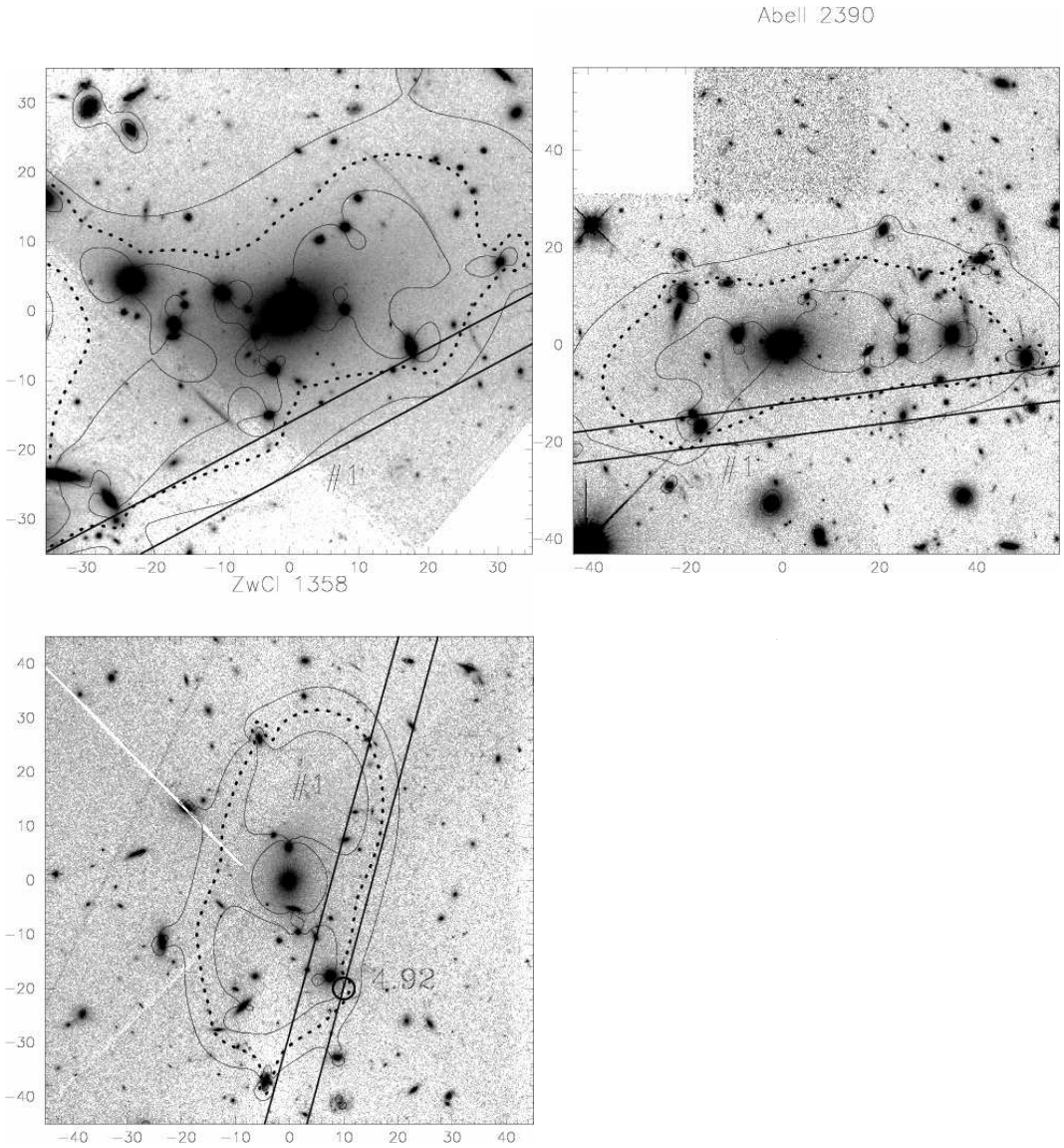


Figure 5.3 As in Fig. 5.2. (upper left) Abell 2219. (upper right) Abell 2390. (lower left) ZwCl 1358.

Table 5.3 Lyman α emission line candidates

Cluster ID	RA	Dec	λ^a	z^b	Flux^c	Comments
Abell 370.1.f	2:39:50.60	-1:33:45.0	4628	2.80		confirmed (LRIS)
Abell 370.1.g	2:39:51.80	-1:35:57.6	4630	2.80		confirmed (Ivison et al., 2002)
Abell 773.1.e	9:17:55.31	51:44:26.6	6978	4.74	1.1	likely
Abell 963.1.cd	10:17:05.10	39:03:30.5	5191	3.27		confirmed (ESI)
Abell 963.1.d	10:17:04.45	39:01:47.1	7025	4.77	0.69	likely
Abell 963.1.efg	10:17:04.77	39:03:11.0	5617	3.62		confirmed (ESI)
Abell 1689.2.f	13:11:25.38	-1:20:52.4	7141	4.82	3.0	confirmed (Frye et al., 2002)
Abell 1689.3.f	13:11:33.52	-1:19:45.2	8817	6.25	1.5	likely (ESI)
Abell 2218.1.a1	16:35:45.25	66:13:26.4	4216	2.47		likely
Abell 2218.1.a2	16:35:51.75	66:12:45.6	8001	5.58	4.4	confirmed (ESI, Ellis et al., 2001)
—	16:35:51.89	66:12:51.5				2nd image
Abell 2218.2.b	16:35:48.78	66:12:24.9	3928	2.23		likely
ZwCl 1358.1.ef	3:59:49.19	62:30:44.8	7205	4.92	10	confirmed (Franx et al., 1997)

[a] Wavelength of emission line in units of Å

[b] Source redshift assuming emission line is Lyman α

[c] Observed line flux (uncorrected for lensing, corrected for transparency) in units of 10^{-17} erg s^{-1} cm^{-2}

data. Sky subtraction was not photon-limited on the strong night sky lines due to the presence of fringing features; we account for this when determining our sensitivity in Section 5.7. We calibrated our absolute efficiency with observations of spectrophotometric stars (Massey & Gronwall, 1990).

The sky-subtracted 2-D spectral images were independently inspected by two of the authors (RSE and JR) and a catalog of 46 candidate Lyman α emission lines was compiled. Astrometric positions were determined for each and the *HST* images inspected for sources at the relevant location. In some cases, candidates were located beyond the boundary of the *HST* images and ground-based images were used.

Candidate Lyman α emission lines were characterized on the basis of several criteria. First, the full spectrum (generally 4000–9500 Å) was closely examined for other emission lines. On a second pass, candidates were ranked as marginal or promising depending on their apparent strength. Out of the initial list of 46, 7 sources are confirmed Lyman α lines at $2.8 < z < 5.7$, and 5 sources are promising candidates that we identify as likely to be Lyman α lines. Those 12 sources are listed in Table 5.3, with observed line fluxes for detections at $z > 4.5$ (on the red arm of the spectrograph).

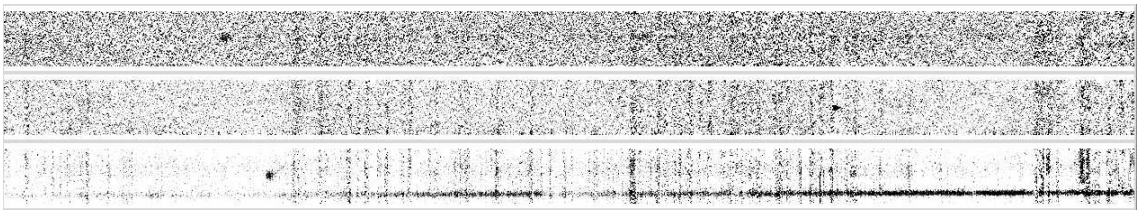


Figure 5.4 Two-dimensional spectra of the three confirmed $z > 4.5$ galaxies detected in the survey. The spectra, from top to bottom, are the $z = 4.89$ source in Abell 1689, the $z = 5.57$ source in Abell 2218, and the $z = 4.92$ source in ZwCl 1358. The wavelength coverage in all the spectra is 6800 to 8430 Å, increasing left to right.

5.5.4 Intermediate-resolution spectroscopy

The [O II] 3727 Å doublet has a rest-frame separation of 3 Å. Thus at $z \sim 1$, when the doublet is redshifted into our most important spectral range, the observed doublet separation is ~ 6 Å. This is close to our LRIS spectral resolution of 4 Å, so to determine whether any of our Lyman α line candidates were unresolved [O II] doublets, we followed up 15 candidates with the Echellette Spectrograph and Imager (ESI, Sheinis et al., 2002). We took spectra using the echelle mode and a 0.75 $''$ slit, which delivered a spectral resolution of $R = 6000$. The exposure times varied depending on candidate strength. ESI spectroscopy confirmed three candidates as Lyman α emission lines (see Table 5.3), at $z = 3.27$, $z = 3.62$, and the $z = 5.57$ galaxy presented in Ellis et al. (2001). We identified a fourth object as a likely Lyman α source at 6.25 based on its ESI spectrum, which supported (though not conclusively) the identification of the emission line as Lyman α. Additionally several of the candidates observed were [O II] 3727 Å doublets.

5.6 Detections

5.6.1 $z > 4.5$

We detected three sources that are certainly at $z > 4.5$ (see Table. 5.3). All three were observed under photometric conditions. Figure 5.4 shows their two-dimensional spectra, and Fig. 5.5 shows a magnified view of the Lyman α emission line as well as *HST* images of the source locations.

We discovered a source toward Abell 2218 at $z = 5.6$ that we discussed in detail in Ellis et al.

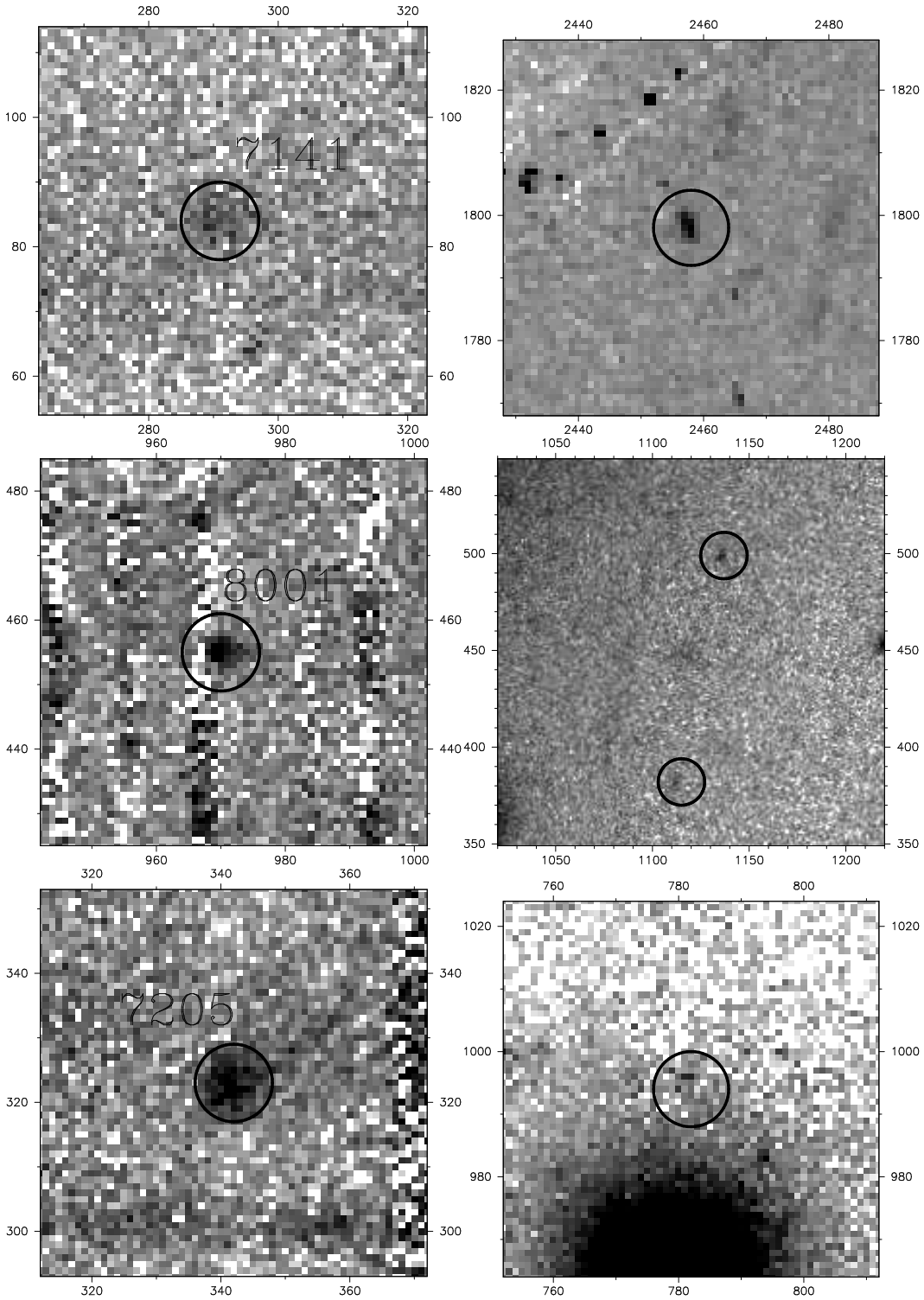


Figure 5.5 Postage stamps of spectra (left) and *HST* images (right). (upper) Abell 1689.2.f. (middle) Abell 2218.1.a2. (lower) ZwCl 1358.1.ef.

(2001). That source is strongly magnified (a factor of 33) and multiply imaged: we used the *HST* data to locate a second image outside of our survey region. The redshift identification was confirmed by an intermediate-resolution spectrum of both images, showing them to be identical, with P-Cygni line profiles characteristic of Lyman α, and certainly not the [O II] 3727 Å doublet. Our lensing model additionally constrained the redshift of the source to be consistent only with the identification of the line as Lyman α. The unlensed luminosity in the Lyman α line is $(7.8 \pm 0.8) \times 10^{41}$ erg s^{-1}.

We blindly recovered a $z = 4.89$ galaxy in Abell 1689 that was discovered originally by Frye, Broadhurst, & Benítez (2002). This object is not multiply imaged, but the Frye et al. (2002) spectrum shows a strong break across the line and metal absorption lines in the continuum redward of Lyman α, confirming the redshift. We estimate the magnification of this source at a factor of 7.2, within the range of 3-14 suggested by coarser modeling of Frye et al. (2002). The unlensed luminosity in the Lyman α line is $(7.4 \pm 0.7) \times 10^{42}$ erg s^{-1}.

In the field of ZwCl 1358, we discovered a source at $z = 4.92$, the same redshift as the strongly lensed arc system discovered by Franx et al. (1997). We tentatively identify this source as a component of the source responsible for the giant arc observed by Franx et al. (1997). This component of the source is magnified by a factor of 10, giving an unlensed luminosity in the Lyman α line of $(2.5 \pm 0.3) \times 10^{42}$ erg s^{-1}.

In addition to these confirmed sources, we discovered three more sources that we consider likely to be Lyman α emission lines at $z > 4.5$ (see Table. 5.3). We discovered a source in the field of Abell 773 that we consider to be a good candidate for Lyman α emission at $z = 4.74$. If this is the correct identification, the unlensed luminosity in the Lyman α line is $(2.8 \pm 0.6) \times 10^{41}$ erg s^{-1}, using a magnification factor of 9.5. Two more sources were discovered in non-photometric observations; though we lack absolute flux calibration for these sources, we apply systematic photometric corrections that we consider uncertain up to a factor of approximately 2. We identified a source in Abell 963 as a likely $z = 4.77$ source magnified by 2.2 times. The source would then have an unlensed Lyman α line luminosity of $(1.4 \pm 0.2) \times 10^{42}$ erg s^{-1} (statistical error only). Our final likely Lyman α candidate is a source that would be at $z = 6.25$, the highest redshift in our sample, in the

field of Abell 1689. We estimate the magnification of this source as a factor of 41, also the highest of all of our detections. The unlensed Lyman α line luminosity is $(3.3 \pm 0.8) \times 10^{41}$ erg s^{-1} (statistical error only). Two-dimensional spectra of the emission lines and images of the likely candidates are provided in Fig. 5.6.

5.6.2 $z < 4.5$

We detected with certainty four Lyman α emission line sources at $z < 4.5$. One was a blind recovery of a $z = 2.80$ source discovered by Ivison et al. (2002), and a second is another galaxy also at $z = 2.80$ in the same field. Two are new detections, at $z = 3.27$ and $z = 3.62$. Two more good candidate Lyman α emission line source identifications are pending. In addition two galaxies, both at $z \simeq 2.5$, were detected in Lyman α absorption. We measured secure redshifts for 51 other sources using other emission lines, primarily [O II], [O III], and Hα. These data are useful for further constraining the cluster mass models, among other things, and will be presented in a separate paper (Richard et al., in prep.).

5.7 Survey analysis

In this section we compute the number density of Lyman α emission-line galaxies in our survey as a function of Lyman α line luminosity, L. To accomplish this we first determine the effective volume of the survey as a function of the luminosity and redshift of a source.

5.7.1 Survey volume

A location in our survey volume is characterized both by location on the sky, $\mathbf{\Omega}$, and a redshift, z. The differential volume element located at position ($\mathbf{\Omega}$,z) in our survey is

$$\mathrm{d}V_\mathrm{c}(\mathbf{\Omega}, z) = \frac{1}{\mathcal{M}(\mathbf{\Omega}, z)} \left[\frac{\mathrm{d}l_\mathrm{c}(z)}{\mathrm{d}z}\mathrm{d}z\right] \times \left[D_\mathrm{c}^2(z)\mathrm{d}\mathbf{\Omega}\right]. \tag{5.1}$$

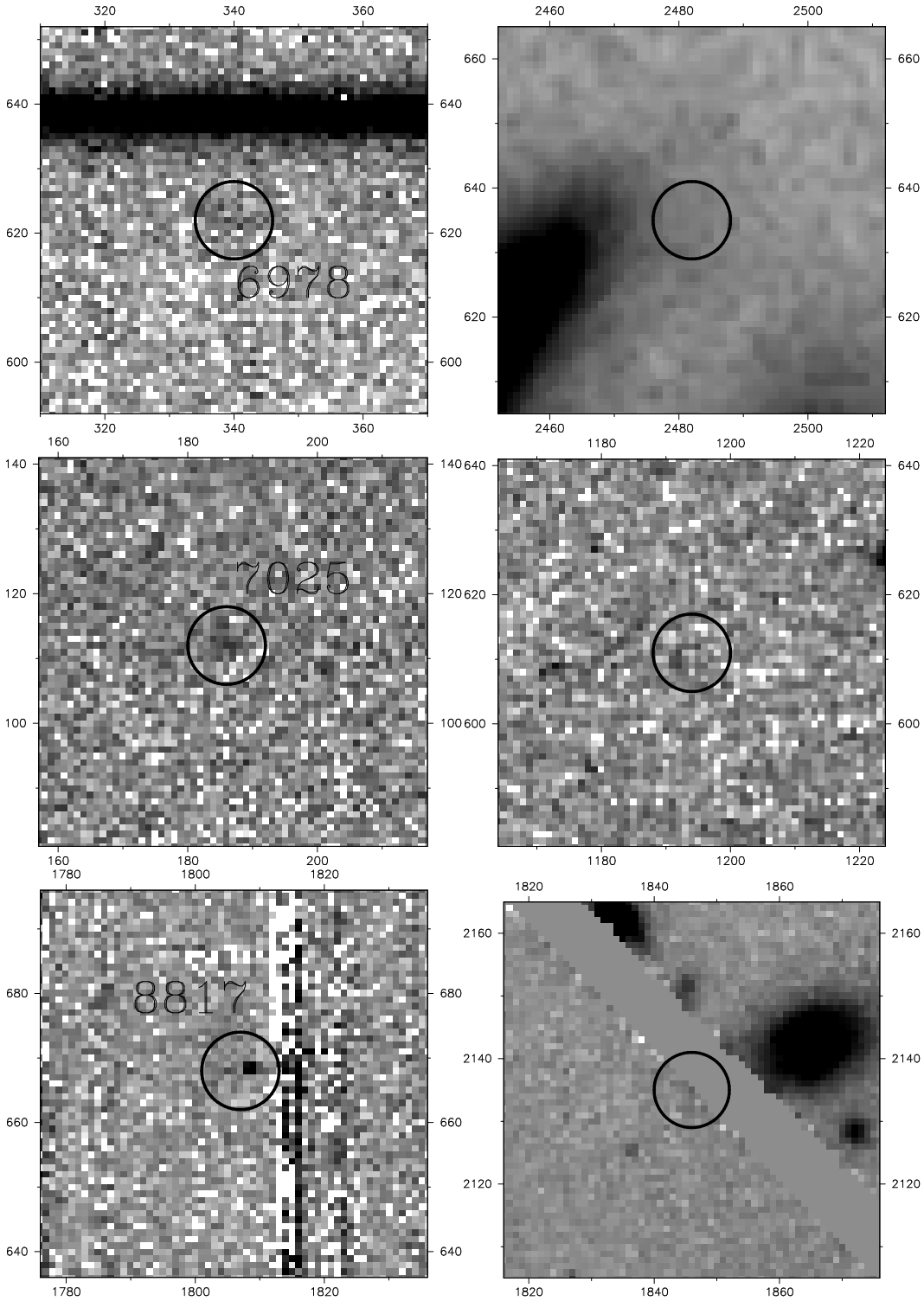


Figure 5.6 Postage stamps of spectra (left) and *HST* images (right). (upper) Abell 773.1.e. (middle) Abell 963.1.d. (lower) Abell 1689.3.f.

The first factor corrects for the lensing effect, which decreases the area surveyed. The second factor is the comoving length of the volume element along the line of sight, with

$$\frac{\mathrm{d}l_{\mathrm{c}}(z)}{\mathrm{d}z} = \frac{c}{H_0\left[\Omega_m(1+z)^3+\Omega_\Lambda\right]^{3/2}} \tag{5.2}$$

(we have assumed a flat universe). The third factor is the comoving transverse area of the volume element, with

$$D_{\mathrm{c}}(z) = \int_0^z \frac{\mathrm{d}l_{\mathrm{c}}(z')}{\mathrm{d}z'}\mathrm{d}z' \tag{5.3}$$

(subscript "c" denotes that the quantity is measured in comoving coordinates, which we use throughout).

Every volume element of our survey is characterized by a limiting Lyman α line luminosity, $L_{\mathrm{lim}}(\mathbf{\Omega}, z)$; a source with Lyman α line luminosity L will be detected in our survey provided it resides in a volume element with $L_{\mathrm{lim}}(\mathbf{\Omega}, z) \leq L$. The limiting luminosity of a volume element depends on the magnification (due to lensing by the foreground cluster), $\mathcal{M}(\mathbf{\Omega}, z)$, the limiting observed Lyman α line flux $f_{\mathrm{lim}}(z)$, and a slit transmission function, $T(\mathbf{\Omega})$,

$$L_{\mathrm{lim}}(\mathbf{\Omega}, z) = \frac{4\pi(1+z)^2 D_{\mathrm{c}}^2(z)}{T(\mathbf{\Omega})}\frac{f_{\mathrm{lim}}(z)}{\mathcal{M}(\mathbf{\Omega}, z)}. \tag{5.4}$$

5.7.1.1 Magnification, $\mathcal{M}(\mathbf{\Omega}, z)$

The magnification due to lensing by a given cluster is a function of position and redshift. The references for the cluster mass models are given in Table 5.1. These models were run through the `LENSTOOL` software (Kneib, 1993) to generate the magnification as a function of redshift at every position in the survey. In practice the area of the survey was divided into parcels of sky of length $0.8''$ and width equal to the slit width, and the magnification was calculated at the center of each parcel. The magnification at each position was sampled at nine redshifts, and the magnification at a particular redshift found by interpolation.

Figure 5.7 shows the magnification as a function of position along a slit observed in Abell 2218,

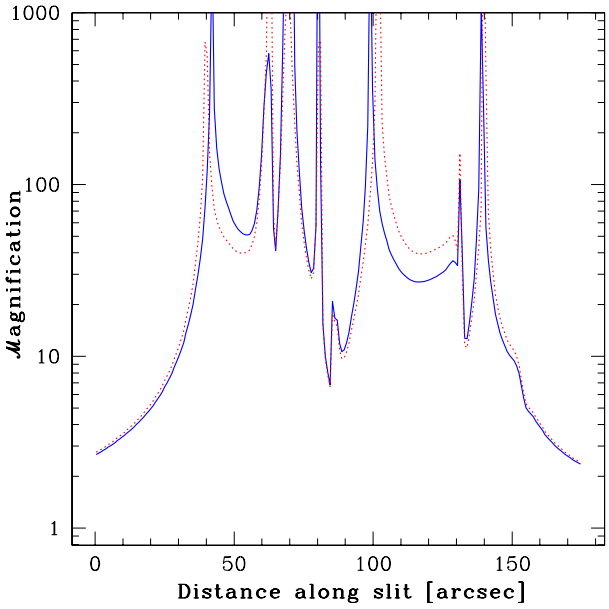


Figure 5.7 Magnification due to the cluster Abell 2218 of background sources, as a function of position and redshift. This figure shows the magnification factor at positions along a 175″ longslit at one of our survey positions in Abell 2218 (within pointing "#1" in Fig. 5.2). The solid curve is for sources at $z = 4.3$, and the dotted curve is for sources at $z = 6.8$.

for two different redshifts. The magnification at a given position is a weak function of redshift for magnification values up to ~ 100, because the angular diameter distance between the cluster and the source changes by less than 25% over the source redshift range $4.5 < z < 6.7$. Very near the critical lines magnification is a stronger function of redshift. Our survey maps sky by observing adjacent slit positions, so errors associated with interpolating the highest magnification $\mathcal{M}(\mathbf{\Omega}, z) > 100$ points should not be important in our estimate of the survey volume.

Figure 5.8 is a cumulative histogram of the magnification factor over the survey; the area surveyed as a function of magnification is very weak function of redshift, even at the highest magnifications. About half of the area we surveyed is magnified by at least a factor of 10, with the lower magnification values coming from area at the ends of the slits, because most strong-lensing clusters subtend a size less than the slit length (175″) on the sky (see Figs. 5.2 and 5.3).

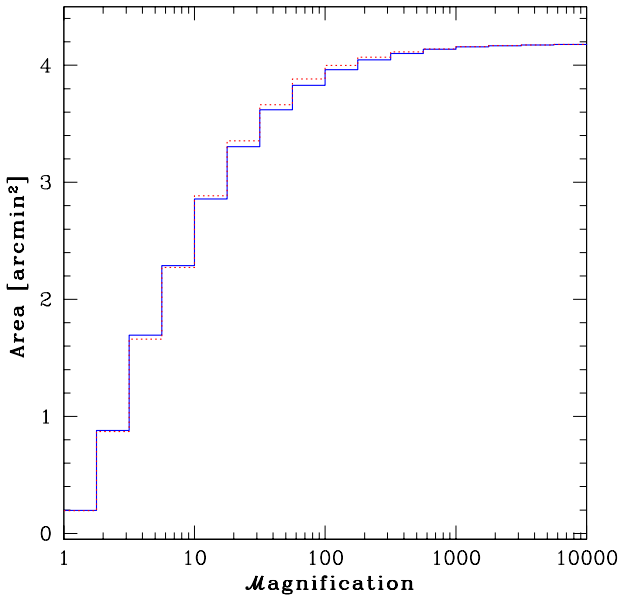


Figure 5.8 Cumulative histogram of the magnification factor over the entire survey area. The solid and dotted curves show the magnification histograms for sources at $z = 4.3$ and $z = 6.8$, respectively.

5.7.1.2 Limiting Lyman α line flux, $f_{\rm lim}(z)$

We define our limiting Lyman α line flux as the signal in an aperture of 1.3″ by 7.7 Å that exceeds 5 times the root-mean-square fluctuations (noise) in apertures of that size, i.e., a 5-σ limit. The spatial dimension of the aperture was chosen to be roughly matched to the seeing, and the spectral dimension was chosen match the expected line-width of high-redshift Lyman α emission from galaxies, $\sim$ 300 km s^{-1}. If a source is larger than our aperture, which is especially possible along the spatial direction if the source is strongly lensed, then we will not be as sensitive to that source as we estimate.

We assumed that the sky noise was constant over the length of the slit at fixed wavelength. This allowed us to include the non-Poisson contribution to the noise level from fringing features, which dominated the noise on strong sky lines. All three of our confirmed $z > 4.5$ detections were more than 5-σ detections, but we found that visual inspection generated candidates (some of which were subsequently confirmed as bona fide emission lines) with fluxes below the 5-σ limit; in particular, two of the likely candidates at $z > 4.5$ are just at the 5-σ limit. Thus a 5-σ limit should be appropriate for the calculation of our survey volume.

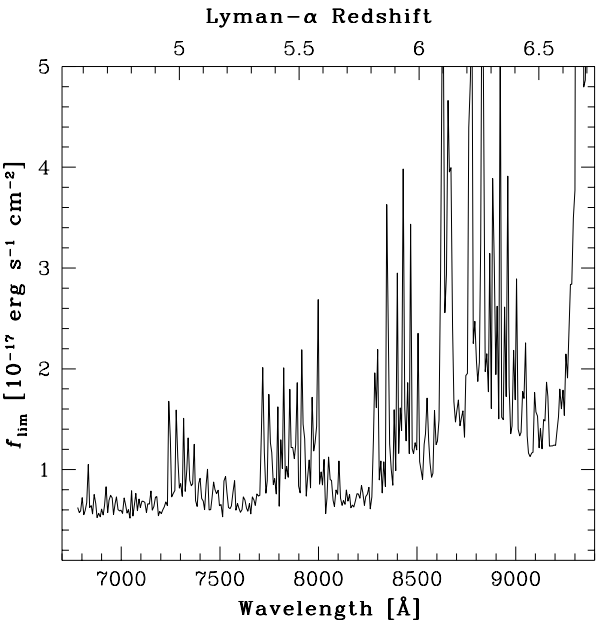


Figure 5.9 Limiting Lyman α line flux as a function of wavelength. The curve is the 5-σ flux limit to detect an emission line in a 1.3″ by 7.7 Å aperture at that wavelength. The top axis is labeled with the redshift corresponding to observed Lyman α falling at the wavelength on the bottom axis. The strong fluctuations in $f_{\rm lim}(z)$ are caused by atmospheric emission lines, and the rise at high redshift is due to decreased instrumental sensitivity. This plot is for a 2000 sec photometric observation.

The limiting line flux varies as a function of wavelength due to the wavelength dependence of the atmospheric absorption and the sensitivities of the telescope and instrument, but the largest dependence is due to the strong variation in atmospheric emission from OH airglow lines (except at $\lambda \gtrsim 9300$ Å, where the sharp drop in instrumental sensitivity dominates). Figure 5.9 shows $f_{\rm lim}(z)$ for the slit pointing illustrated in Fig. 5.7, a 2000 sec observation under photometric conditions. We compute $f_{\rm lim}(z)$ by simple conversion of the observed wavelength into the corresponding redshift for Lyman α to be observed at that wavelength.

Approximately half of our survey data were taken in non-photometric conditions. We account for this by dividing the limiting line flux measured from the observations by our best estimate of the sky transparency during the exposure. During some exposures we have sequences of guide-camera observations that were used to measure relative transparency between observations, and in some cases absolute transparency when photometric guide-camera images were available. In other cases we rely on observation log notes based on the count rate of the guide star as reported by the telescope

operator.

5.7.1.3 Slit transmission, $T(\mathbf{\Omega})$

The slit widths used in our survey, originally $0.7''$ and later $1''$, are comparable in size to the seeing disk. Consequently the transverse distance of a source from the center-line of the slit has a small impact on the source's observability: objects in the center of the slit are easier to detect than objects at the slit edge. Since the absolute calibration was performed with respect to standard stars in the center of the slit, we compute the fraction of light transmitted through the slit as a function of transverse position on the slit (ignoring objects outside of the slit, as they will in general fall on another slit), with respect to an object at the center of the slit,

$$T(\mathbf{\Omega}) = \frac{\mathrm{erf}\left[\frac{w+2x}{s}\beta\right] + \mathrm{erf}\left[\frac{w-2x}{s}\beta\right]}{2\ \mathrm{erf}\left[\frac{w}{s_0}\beta\right]}. \tag{5.5}$$

Here w is the slit width, s is the seeing full-width at half maximum (FWHM) during survey observations, s_0 is the seeing FWHM during standard star observations, $x(\mathbf{\Omega})$ is the transverse distance of the source from the center of the slit, and $\beta \equiv \sqrt{\ln(2)}$. The minimum value of $T(\mathbf{\Omega})$ in our survey is about 0.8, so it has a minor effect on the computation of $L_{\mathrm{lim}}(\mathbf{\Omega}, z)$.

5.7.2 Volume as a function of source redshift and luminosity

The total volume of our survey sensitive to a source of Lyman α line luminosity L is the integral over all volume elements in the survey with $L_{\mathrm{lim}}(\mathbf{\Omega}, z) \leq L$,

$$V_{\mathrm{c}}(L) = \int_{\mathbf{\Omega}} \int_{z} \mathrm{d}V_{\mathrm{c}}(\mathbf{\Omega}, z)\ H[L - L_{\mathrm{lim}}(\mathbf{\Omega}, z)], \tag{5.6}$$

where $H(y)$ is the step function defined with $H(y \geq 0) = 1$.

Figure 5.10 shows the redshift distribution of our survey volume as a function of $L_{\mathrm{lim}}(\mathbf{\Omega}, z)$. The general slight decrease toward high redshift is due to the evolution of the line element with redshift, and the modulation is due to the wavelength-dependent limiting line flux (see Fig. 5.9).

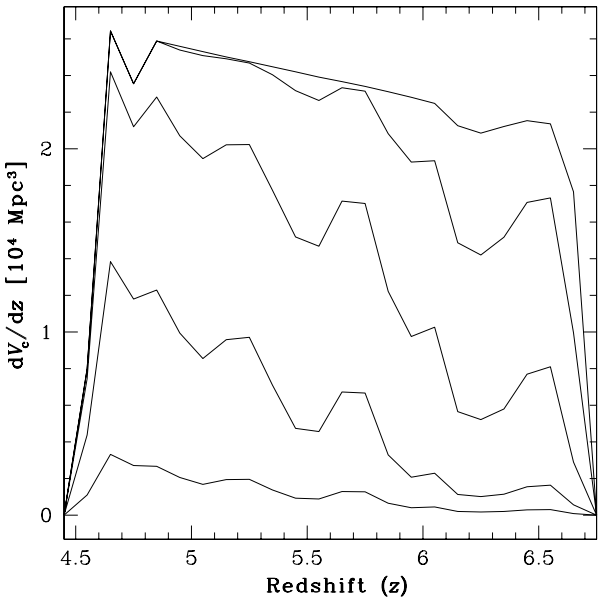


Figure 5.10 The redshift distribution of our survey volume. The curves, from top to bottom, represent the redshift distribution of subsurveys down to limiting luminosity $\log_{10} L = (43.5, 43, 42.5, 42, 41.5)$; for yet lower values of the limiting luminosity, the curves have the shape of the bottom curve, but are scaled down (see Fig. 5.11).

We examine our survey divided into two redshift bins, $4.6 < z < 5.6$, and $5.6 < z < 6.7$. There is no natural binning choice, but by breaking our survey at $z = 5.6$ we have almost equal survey volume (at the brightest luminosities) in each bin. However this places all three of our confirmed high-redshift detections into the $4.6 < z < 5.6$ bin, and none in the $5.6 < z < 6.7$ bin. If we had broken the bins at $z = 5.5$, the number density in the lower redshift bin would decrease, and the number density in the higher-redshift would increase, that is, the removal or inclusion of a source substantially outweighs the change in volume associated with changing the redshift binning.

In Fig. 5.11 we plot the survey volume sensitive to a source of luminosity L for each of our two redshift bins (represented by the two different symbols). At high luminosities there is no dependence of the survey volume on luminosity, because sources at such high luminosities are so bright that we would detect them at any magnification factor or redshift in our survey. At $L = 10^{42}$ erg s^{-1} the high-redshift bin has less volume because of the stronger sky lines at longer wavelengths (see Figs. 5.9 and 5.10) and larger luminosity distance compared to the low-redshift bin. At lower luminosities the survey volume for both bins falls off steadily and similarly. This is a because only strongly magnified

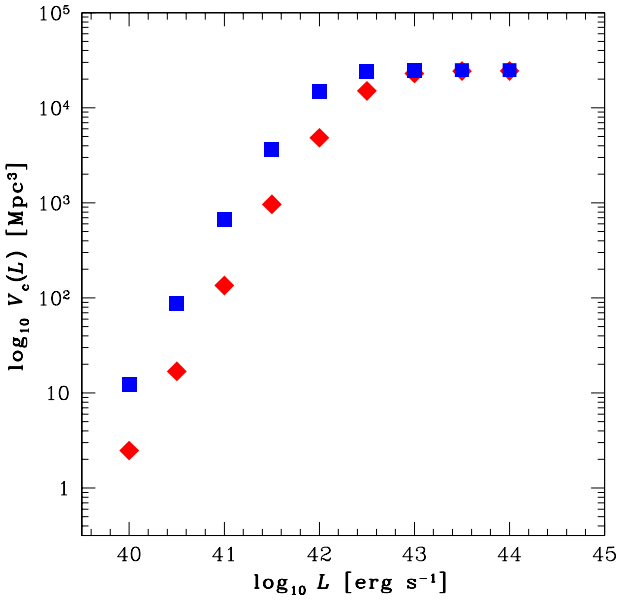


Figure 5.11 Survey volume sensitive to sources with Lyman α line luminosity L. The points show the volume of the survey within which a source of luminosity L would be detected. The survey has been divided into two redshift ranges: the volume with $4.6 < z < 5.6$ is shown with squares, and the volume with $5.6 < z < 6.7$ is shown with diamonds.

volume elements contribute to the survey volume, and the number of volume elements at a given magnification is not sensitive to redshift (Fig. 5.8).

5.7.3 Number density

Our survey detected three confirmed and three likely $z > 4.5$ sources, so to estimate a relatively robust number density parameter, and for comparison with other surveys, we compute a cumulative number density of sources. We construct the cumulative number density at each value of the Lyman α line luminosity L (in each redshift bin) by evaluating the survey volume at that luminosity (see above), and then counting the number of detected sources brighter than L in the survey volume.

Figure 5.12 shows $n(> L)$, the number density of sources with Lyman α line luminosities greater than L, for our two redshift bins, *considering only the three confirmed sources*. There are only upper limits at $L \geq 10^{43}$ erg s^{-1} because although all three detected sources are in the survey volume, none was that luminous. Our most luminous source is 7.4×10^{42} erg s^{-1}, so the first data point appears at $L = 10^{42.5}$ erg s^{-1} (in the low-redshift bin). All three of our confirmed detections contribute to the

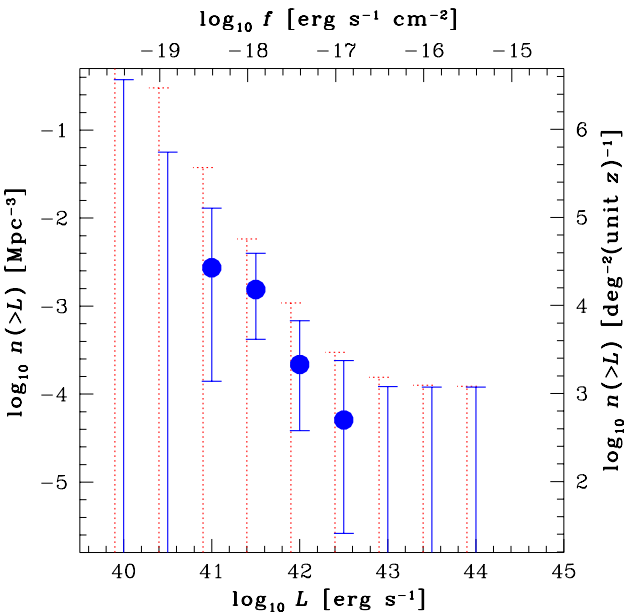


Figure 5.12 Number density of sources brighter than luminosity L, for the three confirmed $z > 4.5$ sources. Each luminosity is treated as an independent sub-survey down to luminosity limit L, and the cumulative number density is calculated from the number of sources in the sub-survey and the volume of the sub-survey. The survey was divided into two redshift bins: the solid lines and points are for $4.6 < z < 5.6$, and the dotted lines (off-set slightly for clarity) are for $5.6 < z < 6.7$. The error bars are 95% limits calculated using Poisson statistics. The top and right axis are labeled in units for comparison of the $4.6 < z < 5.6$ bin results with other work: the left and bottom axes were transformed assuming a redshift of $z = 5.0$.

$L = 10^{41.5}$ erg s^{-1} point because all are brighter than that limit, and each would have been detected even if its luminosity were only $10^{41.5}$ erg s^{-1}. In contrast at $L = 10^{41}$ erg s^{-1}, all three confirmed detections are still brighter than this luminosity, but only one is located inside the $L = 10^{41}$ erg s^{-1} survey volume. At yet fainter luminosities we are back to upper limits because none of the three confirmed detections would have been discovered if it was fainter than 10^{41} erg s^{-1}.

In our high-redshift bin we have no confirmed detections, and thus provide only upper limits at all luminosities. It is clear that though we can rule out a strong increase in the number counts of Lyman α emitters at $5.6 < z < 6.7$ compared to $4.6 < z < 5.6$, we cannot further constrain the number-count evolution. In particular, our results are consistent with no evolution or a decrease with increasing redshift in the Lyman α source counts as a function of redshift over $4.6 < z < 6.7$.

All upper limits and error bars in Fig. 5.12 are 95% confidence limits calculated using Poisson statistics. We have conceptually divided our survey into sub-surveys sensitive down to different

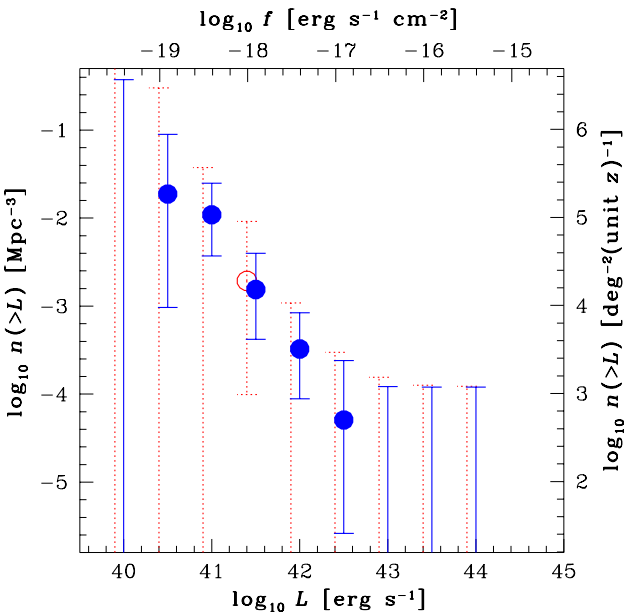


Figure 5.13 Number density of sources brighter than luminosity L, for the three confirmed plus three likely $z > 4.5$ sources. The open circle is the data point in the $5.6 < z < 6.7$ bin; otherwise, as in Fig. 5.12.

Lyman α line luminosities, but these sub-surveys are not independent (and in fact highly correlated). If, for example, a theoretical curve passed just through the upper error bars of two points, our data would indicate roughly a 95% inconsistency, not a 99.8% inconsistency.

The right and top axes of Fig. 5.12 are labeled with unit conversions of the left and bottom axes, assuming the data fall at $z = 5$. These serve to allow a rough reference of our results to be easily read off in the other units commonly used to describe the abundance of Lyman α emission line galaxies. The right and top axes are inapplicable to our high-redshift bin upper limits.

Figure 5.12 shows $n(> L)$, the number density of sources with Lyman α line luminosities greater than L, for our two redshift bins, *considering all six confirmed and likely sources.* Since we added detections while keeping the survey volume fixed, the number densities increased. The number densities of this sample are still marginally consistent with the 95% confidence limits from Fig. 5.12.

5.8 Comparison with other observations and with theory

5.8.1 Comparison with other observations

In Table 5.4 we present parameters inferred from our survey and existing $z \sim 5$ galaxy surveys. The first six entries in the table describe our survey, divided by redshift bin and sub-survey limiting Lyman α line luminosity. The remaining surveys above the horizontal rule are other Lyman α emission line surveys. Surveys below the horizontal rule are Lyman-break galaxy surveys, described later in this section.

There are two entries in the number of sources column for each row corresponding to our data. The first number is the total number of confirmed and likely Lyman α sources in that sub-survey, and the second number, in parentheses, in the number of those that are confirmed. We report the corresponding number densities analogously in the density column.

The limiting luminosity, volume, and number density of each of the previously published Lyman α emission line surveys appearing in Table 5.4 do not always appear in the corresponding reference. As necessary we have used the published information to calculate those values ourselves (for example, converting a limiting line flux and redshift into a limiting line luminosity). We expect that the final results published by the groups may differ somewhat. In particular, there may be a publication bias toward surveys with discoveries, so it is possible there is some bias in the data presented toward higher number density. We have included only systematic Lyman α emission surveys, because reconstructing the survey volumes of published serendipitous discoveries was not possible.

Fig.5.14 plots the data from the $4.6 < z < 5.6$ bin of our survey (as solid circles) with the data from the other Lyman α surveys listed in Table 5.4 (open squares). Note that these points are in general from different redshifts, and no redshift correction has been applied. The error bars shown are 95% confidence limits assuming Poisson errors.

A comparison of our data with published results shows that, by utilizing strong lensing, we have provided meaningful upper limits on the population of Lyman α emission line galaxies two orders of magnitude fainter than previous surveys, in addition to providing confirmed data one order of

Table 5.4 Galaxy surveys at $z \sim 5$

Redshift	$\log_{10}L$[a]	Number[b]	Volume[c]	Density[d]	Reference
4.6-5.6	40.5	1(0)	0.0053	$1.8 \times 10^{-2}(0)$	(1)
4.6-5.6	41	4(1)	0.037	$1.1 \times 10^{-2}(2.7 \times 10^{-3})$	(1)
4.6-5.6	41.5	3(3)	0.20	$1.5 \times 10^{-3}(1.5 \times 10^{-3})$	(1)
4.6-5.6	42	3(2)	0.92	$3.3 \times 10^{-2}(2.2 \times 10^{-4})$	(1)
4.6-5.6	42.5	1(1)	2.0	$5.1 \times 10^{-4}(5.1 \times 10^{-4})$	(1)
5.6-6.7	41.5	1(0)	0.052	$1.9 \times 10^{-3}(0)$	(1)
5-6	42.28	1	0.018	5.6×10^{-3}	(2)
5.7	42.61	1	0.55	1.8×10^{-4}	(3)
6.5	42.95	0	6.1	0	(4)
6.5	42.35	1	0.012	8.5×10^{-3}	(4)
4.9	42.46	87	9.2	9.5×10^{-4}	(5)
5.7	42.81	13	28	4.6×10^{-5}	(6)
6.5	42.89	16	20	8.0×10^{-6}	(7)
5.7	42.95	6	13	4.6×10^{-5}	(8)
~ 5.8	43.35	6	18	3.3×10^{-5}	(9)
~ 6.3	43.32	26	1.7	1.5×10^{-3}	(10)
~ 5	42.96	10	1.4	7.1×10^{-4}	(11)
~ 5.5	43.40	2	2.4	8.3×10^{-5}	(12)
~ 5.5	43.28	16	7.0	2.3×10^{-3}	(12)
~ 5.3	43.18	6	12	5.0×10^{-5}	(13)

[a] Survey limiting L in units of erg s^{-1}

[b] Number of sources detected in survey

[c] Volume of survey in units of 10^4 Mpc

[d] Number density of sources, in units of Mpc^{-3}

References. — (1) this paper; (2) Hu et al. 1998; (3) Hu et al. 1999; (4) Hu et al. 2002a; (5) Ouchi et al. 2003; (6) Rhoads & Malhotra 2001; Rhoads et al. 2003; (7) Kodaira et al. 2003; (8) Maier et al. 2003; (9) Stanway et al. 2003; (10) Yan et al. 2003; (11) Iwata et al. 2003; (12) Fontana et al. 2003; (13) Lehnert & Bremer 2003

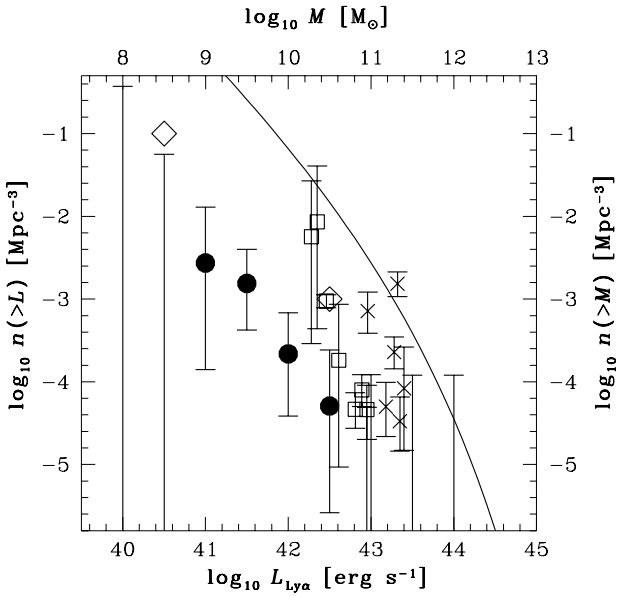


Figure 5.14 Number density of $z \sim 5$ sources brighter than L, for Lyman α surveys and $z \sim 5$ Lyman-break galaxy (LBG) surveys. The solid circles are our cumulative number densities of *confirmed* sources brighter than L for sub-surveys within the $4.6 < z < 5.6$ bin (described in Fig. 5.12). The open squares are the cumulative number densities of sources brighter than L inferred from other $z \sim 5$ Lyman α line surveys, and the crosses are data from $z \sim 5$ LBG surveys. The LBG surveys were converted to equivalent Lyman α line luminosities (see Section 5.8.1). Open diamonds are predictions from Haiman & Spaans (1999). The curve is the cumulative number density of halos above the mass given on the top axis; the vertical scale is the same. The data are described in Table 5.3.

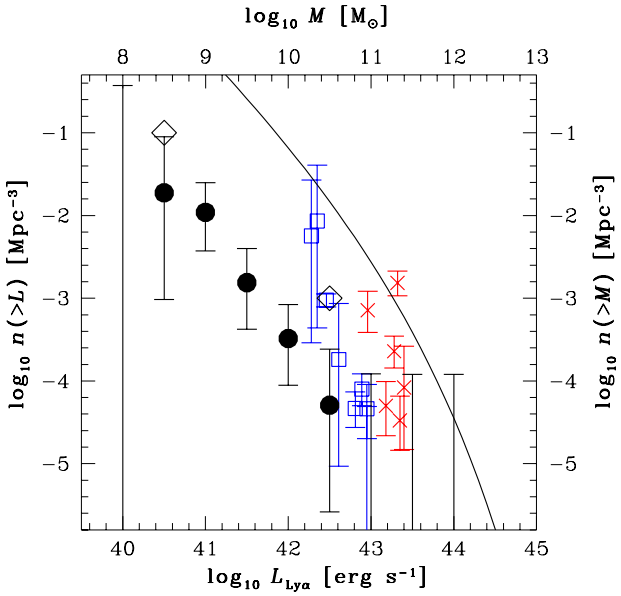


Figure 5.15 Number density of $z \sim 5$ sources brighter than L, for Lyman α surveys and $z \sim 5$ Lyman-break galaxy (LBG) surveys. The solid circles are our cumulative number densities of *confirmed and likely* sources brighter than L for sub-surveys within the $4.6 < z < 5.6$ bin (described in Fig. 5.13). The other symbols are the same as in Fig. 5.14.

magnitude fainter. With existing observational facilities, lensed surveys are the only way to probe to such depth.

At $L = 10^{42.5}$ erg s^{-1}, where our survey overlaps other Lyman α surveys, there is marginal consistency between our data and published results. Most of the other Lyman α surveys are narrow-band photometric surveys (in particular the three surveys with > 10 sources in Table 5.4), where the points plotted do not represent confirmed sources, but rather photometric candidates corrected for the spectroscopic success rate of a small sample. As noted previously this figure represents data reported inhomogeneously, so some of the scatter may be related to the different redshift ranges and Lyman α equivalent-width criteria of the surveys, as well as errors in contamination estimation and possible error in our interpretation of published information.

We detected no confirmed sources at $z > 5.6$. This is marginally inconsistent with existing data of source densities at $z = 5.7$ and $z = 6.5$. However, our lack of sources at $z > 5.6$ compared to other surveys is qualitatively consistent with our smaller number density of $4.6 < z < 5.6$ sources at $L = 10^{42.5}$ erg s^{-1}. The inclusion of our likely $z = 6.25$ candidate in our sample does not change

our number density limits where our survey overlaps other $z > 5.6$ Lyman α surveys.

Only for comparison we have plotted five results from four $z \sim 5$ Lyman-break galaxy (LBG) surveys (Stanway et al., 2003; Yan et al., 2003; Iwata et al., 2003; Fontana et al., 2003; Lehnert & Bremer, 2003). The parameters for these surveys are listed in Table 5.4, below the horizontal rule. Again we have converted published data into number density as necessary, and taken a further step to plot those points on a Lyman α line luminosity scale: the LBG survey limit was converted into a rest-frame UV continuum limit, then into a star-formation rate using the relation of Kennicutt (1998), then into a Lyman α line luminosity assuming 1 $M_\odot\,yr^{-1}$ of star formation produces 10^{42} erg s^{-1} in the Lyman α line (Kennicutt, 1998, after converting Hα luminosity into Lyman α luminosity). No unmitigated conclusions can be drawn from this comparison, though it is intriguing that the $z \sim 5$ LBG surveys may be discovering the same population as the Lyman α emission line galaxies, if the Lyman α line is typically 1/3 the value expected based on the UV continuum SFR. This is similar to the ratio observed in the $z = 5.7$ sample of Lyman α emitters of Ajiki et al. (2003) and in two galaxies at $z = 6.5$ by Hu et al. (2002b) and Kodaira et al. (2003). However, four of six confirmed $4.8 < z < 5.8$ galaxies selected by the Lyman-break technique by Lehnert & Bremer (2003) have Lyman α line fluxes less than 10% of the values naively predicted from their UV continuum SFRs.

5.8.2 Comparison with theoretical models

First we compare our results with the Lyman α emitter model of Haiman & Spaans (1999), who predicted the abundance of Lyman α emitters over a range of redshifts and luminosities. In Figs. 5.14 and 5.15 we plot the predictions of their fiducial model at $z = 5$ as open diamonds at $\log_{10} L = 40.5$ and 42.5. The shape of the luminosity function predicted by Haiman & Spaans (1999) is similar to our observed points, but their fiducial model predicts approximately an order of magnitude more sources than we find. Their models could be reconciled with our data by adopting mass-dependent values of the star-formation efficiency or covering fraction of dusty clouds inside the galaxies.

As a basis for comparing our results with a simple theoretical model, in Figs. 5.14 and 5.15 we re-plot the luminosity function from Fig 5.1 (converted into cumulative form), without any suppression

of the SFR in low-mass halos. This simple explanation of the Lyman α luminosity function relates the number density of galaxies to dark matter halos. We then converted baryons within those dark matter halos into stars, and stellar ionizing light into Lyman α photons. Unlike Haiman & Spaans (1999) we made no attempt to model the radiative transfer of the Lyman α photons.

Specifically, we assumed that 10% of the baryons in each halo were converted into stars every halo dynamical time (defined as the ratio of the halo virial radius to the halo circular velocity at the virial radius). The Hubble time at $z = 5$ is roughly 10 times longer than the halo dynamical time, thus it is possible for such halos to maintain steady star formation at this rate. Star-formation rate was converted into an ionizing-photon rate using a Salpeter IMF with 1/20 solar metallicity (Leitherer et al., 1999). We assumed that 10% of the ionizing photons escape the emitting galaxy, and that 2/3 of the remaining photons are converted into Lyman α emission.

The luminosity function curve of this simple model provides a poor fit to our data. In the context of the model, it is instructive to think of two modifications that would make the predicted luminosity function more closely match our data. The first is to decrease the efficiency factors used to convert halo mass to Lyman α luminosity in a given halo. Alternately, the model curve could be brought into agreement with our data if the efficiency factor for the production of Lyman α as a function of halo mass was correct for a fraction of halos, and the rest of the halos had no observable Lyman α emission at the time of observation.

There are three efficiencies that contribute to the overall conversion of halo mass into Lyman α luminosity, namely the fraction of baryons converted into stars per halo dynamical time, the stellar emissivity of ionizing photons, and the fraction of ionizing photons observed as Lyman α emission. Lowering the combined efficiency by 1.5–2 orders of magnitude would bring the model luminosity function into close agreement with our data. There is some difference in the shape of the curve compared to our data, but this difference is not significant.

In the case where only a fraction of halos contain Lyman α emitters, we would require about 1% of halos to contain emitters at any given time. This could be because, in contrast to the simple model we described, star formation is episodic in nature. In addition, there could be a timescale associated

with the escape of Lyman α photons, such that, for example, dust extinguishes Lyman α emission at the beginning of a starburst, but eventually the dust is expelled and the Lyman α emission line becomes visible (e.g., Shapley et al., 2003). If only some halos contain galaxies, for whatever reason, then this formalism of assuming only a fraction of halos contain Lyman α emitters can also be used, where the fraction now represents a filling factor, rather than a duty cycle.

The resolution of the discrepancy between our data and the model curve has important implications for the mass of the halos that contain the Lyman α emitters. If we over-estimated the Lyman α photon production efficiency in our model, then the association between halo mass and Lyman α emitter expressed in Figs. 5.14 and 5.15 is not correct: the halo mass of our population of emitters at $L \simeq 10^{41.5}$ erg s^{-1} should be $\sim 10^{11}$ M$_{\odot}$. This is the largest mass that could be inferred for this population, assuming a maximum of one Lyman α emitter per halo. From the arguments of Section 5.4, we may expect that halos with masses $\gtrsim 10^{10}$ M$_{\odot}$ to form stars roughly similarly to one another, i.e., though negative feedback may be important in regulating star formation, it is ineffective in halos this massive. Consequently, for this low-efficiency, high halo-mass solution to the discrepancy, we expect that our data should follow the shape of the dark-matter halo mass function, which they do.

In contrast, if we resolve the discrepancy between the model and our data by assuming that the efficiency we calculated is correct for a fraction of halos, and the rest are empty of Lyman α emission, then the mass association in Figs. 5.14 and 5.15 is correct. This implies a halo mass of only $\sim 10^{9.5}$ M$_{\odot}$ for our Lyman α emitters at $L \simeq 10^{41.5}$ erg s^{-1}. Depending on the characteristic mass scale where negative feedback becomes a dominant process, the Lyman α luminosity function may already deviate from the shape of the mass function at $\sim 10^{9.5}$ M$_{\odot}$ (see Fig. 5.1). Our data are slightly flatter than the mass function, and consistent with any of the luminosity functions plotted in Fig. 5.1.

Our theoretical interpretation so far has relied exclusively on our data. If we consider all of the available Lyman α data, there is evidence for a flatter Lyman α luminosity function than we derive from our data alone. The heterogeneous nature of the Lyman α survey data plotted in Figs. 5.14

and 5.15 makes it difficult to draw firm conclusions, but a combination of data at $L \gtrsim 10^{42.5}$ erg s^{-1} suggests that our data point at $L = 10^{41.5}$ erg s^{-1} may be 0.5–1 dex lower than an extrapolation of the Lyman α luminosity function from higher luminosity, assuming the luminosity function shape matches the mass function shape. Thus we conclude that our data, in combination with other Lyman α surveys, suggest that strong negative feedback is suppressing the star-formation rate, and thus Lyman α luminosity, in our sources.

Hamana et al. (2004) used clustering data to estimate the mass of the halos containing Lyman α emitters at $z = 4.9$. They concluded that the characteristic halo mass of those sources is 5×10^{12} M$_\odot$ (Shimasaku et al. 2003 find a halo mass of $\sim 10^{12}$ M$_\odot$ for similar $z = 4.9$ emitters on the basis of a large-scale structure feature in their survey). This conclusion would support the low-efficiency, high halo-mass solution to the difference between our model luminosity function and our data. However, the number density of $z = 4.9$ emitters is larger, by about a factor of five to ten, than the number density of 10^{12} M$_\odot$ halos (Hamana et al., 2004). This implies, contrary to our assumption above, that there is more than one Lyman α source per halo. The virial radius of a $z = 5$, 10^{12} M$_\odot$ halo is $8.5''$, so multiple sources inside a single halo may be observed as separate sources, though this should create a very distinct signature in the spatial distribution of sources (or extended nature, if the sources are unresolved) that has not been reported by other Lyman α emitter surveys.

While current information on the masses of Lyman α emitter halos is still limited, progress will continue to be made at $L \gtrsim 10^{42.5}$ erg s^{-1} by large Lyman α surveys. Unfortunately, surveys for low-luminosity Lyman α emitters will not provide sufficient survey area for clustering studies in the near future. Lensed surveys such as ours, in particular, do not lend themselves to easy clustering analysis, because the contiguous survey volume is very complex and limited in size by the mass of the lensing foreground cluster. As an aside we comment that Poisson errors dominate the uncertainty in the number densities plotted in Figs. 5.12 and 5.13 (and Figs. 5.14 and 5.15 for our survey), assuming the maximum mass for the halos containing our emitters (i.e., every halo contains a source; see above for a caveat), and using the clustering formalism of Mo & White (2002).

Until the advent of large-area, low-luminosity Lyman α surveys, the only constraint on the mass

of the halos containing the emitters, and thus the only path toward understanding the suppression of star-formation in low-mass halos, lies in detecting source populations with high number densities, such that the halo mass function, and the assumption that there is at most one source per halo, can be used to infer a maximum halo mass for the population of Lyman α emitters. This is strong motivation for future surveys to continue to use strong lensing to survey small volumes to incredible depths for faint, Lyman α emitting sources.

5.9 Summary

We performed a systematic survey for Lyman α emission at $2.2 < z < 6.7$ using strong lensing from intermediate-redshift clusters of galaxies to boost our survey sensitivity to unprecedented depths. We detected three confirmed and three likely Lyman α emitting galaxies at $4.7 < z < 6.25$, with Lyman α line luminosities of 2.8×10^{41} erg s^{-1} $< L < 7.4 \times 10^{42}$ erg s^{-1}. Our survey covered 4.2 arcmin^2 on the sky, with a maximum volume of 4×10^4 Mpc^3 over $4.6 < z < 6.7$. We find no evidence for redshift evolution of the number density of Lyman α emitting galaxies between $z \sim 5$ and $z \sim 6$, though our data are also consistent with a decrease in number density with increasing redshift.

We present the first meaningful constraints on the the luminosity function of Lyman α emitters at $4.6 < z < 5.6$ over the Lyman α luminosity range 10^{40} erg s^{-1} $< L < 10^{42}$ erg s^{-1}, corresponding to inferred star-formation rates of 0.01–1 M$_\odot$ yr^{-1}. From a consideration of the number density of dark-matter halos, we conclude that our population of sources at $L \sim 10^{41.5}$ erg s^{-1} resides in halos of mass $\lesssim 10^{11}$ M$_\odot$.

Our number density data are consistent with a Lyman α luminosity function with the same shape as the halo mass function, but a consideration of all available Lyman α survey data implies that we have observed a flattening of the Lyman α luminosity function with respect to the halo mass function. We may have detected evidence of the suppression of star-formation in low-mass halos at high redshift, as predicted by theoretical models of galaxy formation.

Acknowledgements

We thank Alice Shapley for many enlightening conversations. We also thank Pieter van Dokkum, Andrew Firth, and Tommaso Treu for help obtaining and reducing the observations.

We gratefully acknowledge the helpful staff at Keck Observatory, and the teams responsible for the creation and maintenance of the telescopes and instruments there.

The authors recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

MRS acknowledges the support of NASA GSRP grant NGT5-50339.

Chapter 6

The Formation of Metal-Poor Globular Cluster Systems[1]

6.1 Abstract

We present a model for the formation of metal-poor ("blue") globular cluster systems, in the context of hierarchical models of cosmological structure formation. We assume that the reionization epoch suppressed the formation of intermediate-metallicity globular clusters, resulting in the observed bimodal metallicity distributions of globular cluster systems. Our numerical simulations of the formation of the Milky Way metal-poor globular cluster system, using cosmological initial conditions of the galaxy formation process, reproduce the abundance and spatial extent of the observed Milky Way metal-poor globular cluster system. Models where metal-poor globular cluster formation happens early and is truncated at a fixed redshift generically predict an increase in the number of metal-poor globular clusters per unit total galaxy mass with increasing richness of galactic environment, as observed (e.g., Rosenblatt et al., 1988; West, 1993). Our numerical model confirms previous analytic arguments about the number and spatial distribution of globular clusters (Rosenblatt et al., 1988), and suggests that early, 'pre-galactic' formation scenarios for metal-poor globular cluster systems may provide a successful framework for interpreting current and future observations.

[1]The work described in this chapter will be submitted in 2004 to MNRAS for publication.

6.2 Introduction

The globular cluster system of the Milky Way is very well defined: there is little debate over the membership or properties of the system or even of individual globular clusters. Despite the fact that our globular cluster system is observationally characterized extremely well, theoretical modelling of the formation and evolution of globular cluster systems, even the Milky Way's, is an intransigent problem in galaxy formation theory.

There is no dearth of globular cluster system (GCS) formation models (see Ashman & Zepf 1998 for a review). However, few GCS formation models connected the process of GCS formation with current models of galaxy and structure formation (Beasley et al. 2002 is a notable exception). The rapid advance of GCS observations has led to a number of new challenges to formation models. Consequently, we are interested in using the modern paradigm of structure and galaxy formation to interpret and predict GCS observations.

Two major observational challenges motivate this work. First, the Milky Way globular clusters span a range of metallicity, $-2.3 \leq [\mathrm{Fe/H}] \leq 0.0$, but do not fill that space uniformly: the metallicity distribution appears bimodal, with peaks at $[\mathrm{Fe/H}] \simeq -1.6$ and -0.6 (Zinn, 1985; Armandroff & Zinn, 1988; Harris, 1996). Each metallicity population, called metal-poor and metal-rich, has a spatial distribution and kinematics distinct from the other. Not surprisingly, existing GCS formation models have often suggested that the metal-poor GCS[2] formed in an early epoch, and the metal-rich GCS formed later, but still early enough that metal-rich globular clusters (in the Milky Way, at least) have old ages (e.g. Ashman & Zepf, 1992). Bimodal GCS metallicity distributions must be a general prediction of formation models, because bimodality is observed in the GCS of all massive galaxies (e.g. Larsen et al., 2001).

However, an implementation of globular cluster formation within the context of a semi-analytic galaxy formation model failed to generate bimodal metallicity distributions of GCSs unless an *ad hoc* limit on the formation of metal-poor globular clusters was imposed (Beasley et al., 2002). This reflects the basic fact that the semi-analytic galaxy formation model used, from Cole et al. (2000),

[2]The metal-poor GCS is often called the "blue" GCS, because old, low-metallicity stellar systems have blue optical colors relative to old, metal-rich systems.

had no particular characteristic feature in the star-formation rate (SFR) as a function of time or metallicity (for the metallicity range of interest here). Consequently, there was nothing to "pick out" a metallicity of [Fe/H] $\simeq -1$ and stop globular cluster formation in those environments.

The second observational challenge is what we will call the metal-poor specific frequency problem. Specific frequency was originally defined as a ratio between the number of globular clusters in a galaxy and the V-band luminosity of that galaxy (Harris & van den Bergh, 1981). It is one of the only properties of GCSs that varies substantially from galaxy to galaxy, over a factor of 10. There is a clear correlation between density of environment and specific frequency, such that cluster galaxies have higher specific frequencies than field galaxies; cD galaxies have the highest specific frequencies of all. Harris (1991) reviewed these findings, and also reviewed their main implication for galaxy formation: if the ratio of globular cluster formation efficiency to star-formation efficiency is fixed in time, then the specific frequency of a galaxy-galaxy merger product will be the (mass-)weighted average of its progenitors. This was viewed as a challenge to the idea that elliptical galaxies formed by mergers of spirals (e.g., Schweizer, 1986), because spirals (being in the field) have lower specific frequencies than ellipticals; this gave rise to the description "specific-frequency problem." However, if globular cluster formation efficiency is relatively enhanced in mergers compared to the rest of the star-formation history of galaxies, then a merger product can have a specific frequency higher than its progenitors (see Ashman & Zepf 1998 for a review).

Harris (2003) refined the "problem" by re-posing the specific frequency analysis restricting attention to metal-poor globular clusters. The specific frequency dependence on galaxy environment persists, that is, cluster galaxies have higher metal-poor specific frequencies than galaxies in the field. The stars in elliptical galaxies are generally metal-rich (e.g. Harris & Harris, 2002), implying that any globular clusters formed during the merger of, e.g., two spiral galaxies, would also be metal rich. This leads to an unresolved challenge: explain why cluster ellipticals formed more metal-poor globular clusters than, e.g., field ellipticals of the same mass.

There are two previous globular cluster formation calculations that demonstrated the promise of reproducing the observed trends in the specific frequency of GCSs within the modern structure-

formation paradigm. Rosenblatt et al. (1988) assumed metal-poor globular clusters (specifically those with spheroidal kinematics) formed in regions of 10^6 $M_\odot$ that collapsed before a fixed redshift, in the context of a "pre-galactic" model of globular clusters that identified globular clusters with collapsed structures of the globular-cluster mass scale (Peebles & Dicke, 1968). They concluded that if density fluctations of mass 10^6 $M_\odot$ with amplitude approximately 3 times larger than the r.m.s. value of density fluctations on that scale (i.e., 3σ perturbations) survived as observed globular clusters, then the resulting population may adequately match the observed properties metal-poor GCSs, including the trend in specific frequency with galaxy Hubble type, and the radial profile of the Milky Way GCS. Their inference of the radial profile of the Milky Way GCS was a qualitative one: they argued that since the pre-galactic structures are clustered, globular clusters formed in those sites should be more centrally concentrated in the assembled galaxy than the galaxy's dark matter halo. Despite some apparent successes of their model, Rosenblatt et al. (1988) concluded that their fundamental assumption of globular cluster formation at a fixed density threshold "lacks convincing physical justification."

West (1993) reconsidered a similar globular cluster formation scenario to Rosenblatt et al. (1988), and focused on explaining the relationship between globular cluster specific frequency and galactic environment, especially the observation that cD ellipticals at the centers of massive clusters of galaxies show the highest specific frequency of all known galaxies. For a CDM-like power spectrum and reasonable assumptions about the density fluctuations that grew into globular clusters, galaxies, and clusters of galaxies, West (1993) found that the observed trends in specific frequency with Hubble type and galactic environment could be reproduced.

In this paper we seek to extend models of GCS formation within the modern context of galaxy formation. Specifically, we postulate that metal-poor globular clusters formed before the assembly of their parent galaxy, in pre-galactic structures that collapsed before a fixed time, above a fixed density (corresponding to density fluctuations larger than a fixed initial amplitude, similar to, e.g., Rosenblatt et al., 1988; West, 1993). We physically motivate our choice of metal-poor globular cluster formation time by associating it with the redshift of reionization. We assume that, though metal-

rich globular cluster formation continued after reionization, the conditions for metal-poor globular formation were only satisfied before reionization. Though we do not present an explicit model of observed GCS metallicity bimodality in this paper, we describe the physical conditions under which the raising of the cosmological Jeans mass by reionization may restrict subsequent globular cluster formation to regions more enriched than metal-poor globular clusters.

We perform numerical calculations, based on realizations of cosmological initial conditions and collisionless evolution of structure formation through to the present, aimed at understanding the resulting properties of a metal-poor GCS formed under our assumptions. A simulation of a Milky Way-like galaxy shows that, consistent with the qualitative expectations of Rosenblatt et al. (1988), globular clusters formed in association with pre-galactic structure are more centrally concentrated within the final galaxy than its dark matter halo, although to reproduce the observed Milky Way metal-poor GCS spatial distribution, and other observations of globular clusters (such as their inferred lack of dark matter halos, Moore 1996), our mass scale and formation time choices differ from the choices of Rosenblatt et al. (1988) (we use larger mass scales at later times). Finally, we perform an analysis of the expected variation of metal-poor globular cluster specific frequency as a function of galaxy environment using statistics of the initial density field fluctations. This work will be complemented in the future by a full numerical exploration using a large set of simulations similar to our Milky Way simulation. Our preliminary conclusions support the finding of West (1993), that biased formation of globular clusters, coupled with biased galaxy formation, may explain the amplitude of the globular cluster specific frequency variation with galactic environment, using the currently favored cosmological model (described below, cf. West 1993, who considered an Einstein-de Sitter model).

This paper is organized as follows. In the next section we present our basic model. Section 6.4 describes the computational techniques we use to simulate the formation of the Milky Way metal-poor GCS. In Section 6.5 we show the results of our Milky Way simulation, and use those results to refine our model. We turn to the comparative properties of metal-poor GCSs of galaxies in different environments in Section 6.6. Section 6.7 summarizes.

Our model depends sensitively on our adopted cosmology. We use $(\Omega_{\rm m}, \Omega_\Lambda, \Omega_{\rm b}, \sigma_8)$=$(0.3, 0.7, 0.043, 0.9)$ and $h \equiv H_0/(100 \text{ km s}^{-1}\text{Mpc}^{-1}) = 0.7$; these values are consistent with the values derived in Spergel et al. (2003).

6.3 Model

First we present our motivation for introducing reionization as an important parameter in the formation history of GCSs. Within the framework we establish, we go on to describe the three free parameters of our model.

6.3.1 The importance of reionization

Our model explains the dichotomy between metal-poor and metal-rich globular cluster systems as a chronological sequence, where metal-poor globular clusters are older. We assume they formed in the early universe as gas-rich proto-galactic building blocks merged together to form larger structure: the merger rate was very high in the early universe. Specifically, we assume that globular clusters formed from cold gas within dark matter halos. We do not claim that globular clusters have dark matter halos of their own; this conflicts with observation of globular cluster tidal tails (Moore, 1996). But, just as a Milky Way globular cluster resides within a dark matter halo without having its own little halo, we postulate that globular cluster formation proceeds within dark matter halos without involving dark matter halos on the scale of globular clusters themselves.

At $z \sim 10$, the universe reionized: H II regions around early galaxies percolated throughout the universe, and the temperature of the intergalactic medium (IGM) jumped from tens of Kelvins to 10^4 K. The redshift of reionization is uncertain. Observations of the neutral component of the IGM suggested that reionization was finishing at $\simeq 6.5$ (Fan et al., 2002). However, the first *WMAP* satellite results indicated that the universe may have been reionized at $z \simeq 20$ (Kogut et al., 2003). Young globular clusters contributed both ionizing photons and metals to the IGM; under certain assumptions they may have even contributed substantially to the reionization process itself (Ricotti, 2002).

The IGM was heated when it reionized. There was an associated rise in the Jeans mass, and in the time-averaged Jeans mass, called the filtering mass, $M_{\rm F}$ (Gnedin, 2000). In the top panel of Fig. 6.1, the thick line illustrates the evolution of $M_{\rm F}$ with redshift, assuming reionization at $z \simeq 15$[3]. Dark matter halos more massive than $M_{\rm F}$ at any given redshift have sufficiently deep potential wells to draw in their associated baryons; halos less massive than $M_{\rm F}$ undergo dark matter collapse, but with little associated baryons. For comparison with $M_{\rm F}$, we characterize the halo abundance with the thin lines in the top panel of Fig. 6.1. The curves correspond, from left to right, to 5,4,3,2, and 1σ fluctuations of matter in the universe, as a function of redshift. For example, at $z = 9$, halos with total mass $M = 10^6$ $\rm M_\odot$ are only 1σ fluctuations, and thus are abundant. By comparison, at that same redshift $M = 10^{10}$ $\rm M_\odot$ halos are 5σ fluctuations and quite rare: about five orders of magnitude less abundant than 1σ fluctuations (Press & Schechter, 1974).

The scenario illustrated by Fig. 6.1 implies that galaxy formation proceeded unimpeded on mass scales of, e.g., 10^8 $\rm M_\odot$, until the filtering mass increased in association with reionization. After that point ($z \sim 9$ in the figure), $M = 10^8$ $\rm M_\odot$ halos no longer efficiently accreted baryons, and thus were relatively sterile to star formation. One way of examining this effect integrated over all mass scales is shown in the bottom panel of Fig. 6.1. The quantity F, shown by the dotted line, is the fraction of matter in the universe contained in halos with $M > M_{\rm F}$, at that redshift. If we had assumed that $M_{\rm F} = 0$ at all times, then F would be a smoothly increasing monotonic function of decreasing redshift, reflecting the progress of dark matter structure formation. However, by introducing a time-variable filtering mass, we generated an interval of redshift from $z \simeq 15$ to $z \simeq 7$ when very little new baryonic structure formed: the evolution of F stalls at a constant value. The time associated with that interval is $\sim 5 \times 10^8$ yr. The pause in the increase of F is emphasized by the solid curve, which shows $\dot{F}$, the rate of change of F per Gyr: there is a pre-reionization burst of excitement when the amount of collapsed baryonic matter increased by an order of magnitude in 10^8 yr, and a post-reionization recovery of of the growth of baryonic structure, with a hiatus in between. Barkana & Loeb (1999) used this observation to propose that the star-formation rate as a function of redshift

[3]This $M_{\rm F}(z)$ is scaled up in redshift from data presented by Benson et al. (2002b). As such, it is not an exact calculation, but illustrates the important qualitative behavior.

has a characteristic dip associated with reionization (but see also Benson et al. 2002b).

We propose that globular cluster formation traced baryonic structure formation. Benson et al. (2002b) showed in their galaxy formation calculation that negative feedback, due to, e.g., supernovae, already suppressed star formation in halos with $M < M_{\rm F}(z = 0) \simeq 10^{10}$ M$_\odot$, so that the integrated SFR of the universe did not show a suppression associated with reionization. However, though relatively large systems dominated the total SFR, these massive systems are also likely to be metal-rich due to the very fact that they are massive enough to form stars efficiently irrespective of feedback. If less-massive systems dominate the production of metal-poor globular clusters, then even though the SFR integrated over the universe may not be substantially suppressed by reionization, the formation of metal-poor globular clusters would be.

We assume that globular cluster formation resumed after reionization completed, as baryonic structure formation resumed, though note that metal-rich globular cluster formation may already occur before or during reionization. But by the end of reionization stellar evolution has had time to further enrich the gas that stars are forming out of, and thus though metal-rich globular clusters may still form, the era of metal-poor GCS formation is complete, at least in structure that will evolve into $z = 0$ galaxies. Some post-reionization low-metallicity gas may actively form stars, and even metal-poor globular clusters, in certain rare environments, but this is not relevant to the GCS formation of massive galaxies. Note that even at $z = 5$, the universe is only 1 Gyr old, so that metal-rich globular clusters forming then may still be nearly as old as the age of the universe, consistent with observations of Milky Way globular clusters (e.g., Krauss & Chaboyer, 2003).

This is the end of our discussion of GCS metallicity. We emphasize that we do not calculate any metallicities, and thus do not explain the value of the metallicity gap, its width, or the metallicity spread of the metal-poor population. Moreover, we will not discuss the formation of the metal-rich GCS further. We believe that metal-rich globular clusters form in detail similarly to metal-poor globular clusters, explaining the structural similarity among globular clusters of all metallicities. However, in our model metal-rich globular clusters form, in general, later than metal-poor globular clusters, and under circumstances that we do not attempt to address. The modelling of Beasley

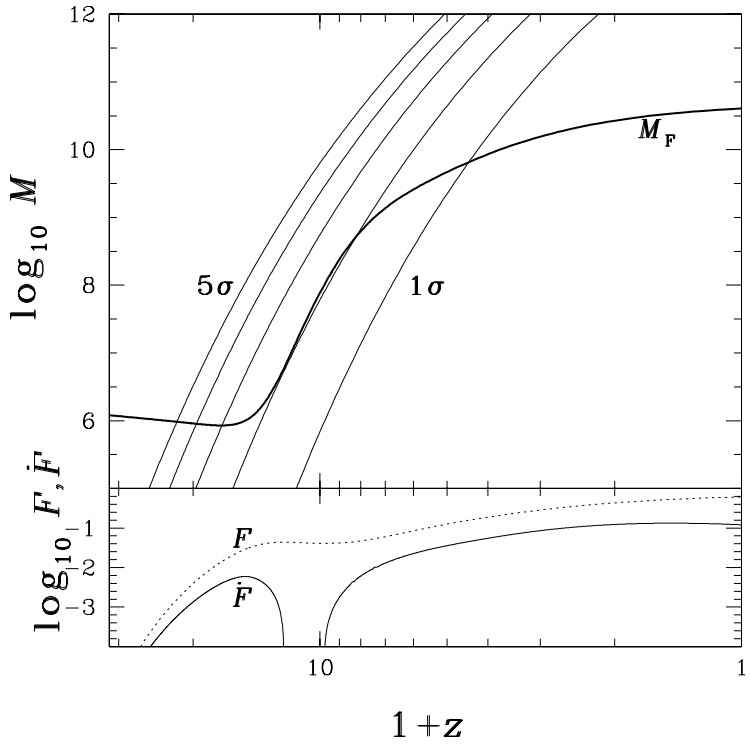


Figure 6.1 Baryonic structure formation and reionization. (**top**) This panel shows both the characteristic mass scales of dark matter structure formation, and also the characteristic mass scale of baryonic collapse. The filtering mass (*thick line*) schematically represents the mass scale below which baryons cannot collapse and cool (Gnedin, 2000); above this scale the dark matter gravitational potential is strong enough to overcome the baryonic gas pressure. The other curves (*thin lines*) show the progress of dark matter structure formation. The curves represent 1,2,3,4,5σ fluctuations of matter in the universe, as labelled. (**bottom**) This panel shows the fraction (*dotted curve*, F) of matter in the universe collapsed in structures with mass greater than the filtering mass, and its rate of change per Gyr (*solid curve*, $\dot{F}$) as a function of redshift.

et al. (2002) shows great promise in explaining the formation of metal-rich GCSs.

6.3.2 Model parameters

Under the assumption that metal-poor GCSs formed before reionization, we have (in the simplest case) three parameters to describe our model. The first is the end of the metal-poor GCS formation epoch, the redshift of reionization, $z_{\rm reion}$[4]. Though it is possible for reionization to occur at different times in different locations, especially given the uncertainty regarding the details of reionization in light of the *WMAP* results, we proceed under the simplifying assumption that one redshift characterizes the end of the metal-poor GCS formation everywhere.

Our second parameter is the total halo mass below which globular clusters cannot form, $M_{\rm cut}$. Globular cluster formation in detail is an uncertain process (e.g., Ashman & Zepf, 2001, for a recent review of some of the complexities). It may be that the conditions required for globular cluster formation do not occur in 10^6 $M_\odot$ halos, the smallest that can collapse baryons before reionization, but only in halos bigger than 10^8 $M_\odot$. We parameterize this ignorance with $M_{\rm cut}$. Future work on the formation of individual globular clusters within a cosmological framework, such as Kravtsov & Gnedin (2003), may eventually provide a physically motivated value for $M_{\rm cut}$.

Our third and final free parameter is the efficiency of globular cluster formation, η (see McLaughlin, 1999, for a general discussion of this topic). A key constraint on this parameter is that it must not be larger than unity: we cannot make more globular cluster stars than we have gas for. Beyond this constraint, we treat η as a free parameter to be adjusted to fit data; for simplicity and definiteness, we assume a universal value of η applies throughout the universe (before reionization; our η may be unrelated to the efficiency of formation of metal-rich globular clusters).

[4]Strictly, the end of the metal-poor GCS formation epoch occurs after reionization, because the filtering mass lags the Jeans mass. We will still name this time $z_{\rm reion}$ because of the association with reionization.

6.4 Computational methods

In this section we present the computational implementation of our model, specifically for the modelling of the formation of the Milky Way metal-poor GCS. First, we describe our method for generating the cosmological initial conditions of the formation of the Milky Way; these initial conditions are used to identify the formation sites of globular clusters. Second, we describe the N-body collisionless gravitational simulations used to evolve the initial conditions to $z = 0$, and to compute the final spatial distribution of the Milky Way metal-poor GCS.

6.4.1 Initial conditions

Our goal is to generate initial conditions that are consistent with our best knowledge of the power spectrum of matter fluctuations (Spergel et al., 2003), while at the same time we want to impose the constraint that we form a galaxy of a specified mass and collapse redshift within a relatively small simulation box (due to computational limits). One approach would be to simulate a large volume of the universe to identify a suitable halo, then excise that part of the simulation box and resimulate the halo of interest at higher resolution. Because we were interested in exploring a variety of final halo masses and collapse histories (Section 6.6), resimulation is a computationally costly method. Instead, we have made use of the technique of constrained realizations of cosmological initial conditions. We used the spherical overdensity formalism (e.g., Peebles, 1993) to make an educated guess at the length scale and perturbation strength that corresponded to the galaxy formation history we are interested in. Roughly speaking, the length scale sets the mass of the final galaxy, and the overdensity sets the collapse history (larger overdensities collapse earlier).

Our implemented technique allowed us to specify an overdensity averaged over a Gaussian sphere, while still allowing a full realization of power on all scales represented in our simulation (Hoffman & Ribak, 1991). We realized our constrained initial conditions using the publicly-available software package `COSMICS/GRAFIC` by E. Bertschinger[5] (van de Weygaert & Bertschinger, 1996). We reiterate that our initial conditions were not simple spherical overdensities, but rather realistic realizations

[5]available at http://arcturus.mit.edu/cosmics/

with power on all scales in the simulation box, constrained so that a halo close to the desired mass collapsed at close to the desired redshift. Using the simulations described in the following subsection, we evolved these conditions to $z = 0$ and used the results to calibrate our choices of perturbation strength and size.

We also used the spherical overdensity formalism to locate sites of globular cluster formation. In spherical overdensity theory, a sufficiently positive value for the cosmic overdensity on any scale corresponds to a collapse time for that overdensity. Consequently, if we are interested in identifying the locations of halos of mass $M = 10^8$ $M_\odot$ that collapse at $z = 6$, we can compute the overdensity around every point in our simulation volume. If the overdensity linearly extrapolated to $z = 0$ exceeds a critical threshold that depends on collapse redshift (i.e., 6, in this example), then we assume that location in space was the site of a collapsed halo at $z = 6$.

A more reliable way of identifying collapsed structure at a (high) redshift of interest would be to find halos directly from the N-body simulations described next. However, the spherical overdensity technique can be applied down to the smallest scale represented in the simulation volume, whereas identifying halos reliably requires many tens of particles per halo. We used simulation cubes that measured 5 Mpc on a side, with 64^3 total particles. This gave a particle mass of 2×10^7 $M_\odot$. Even with the spherical overdensity method of identifying halo collapse times and masses we only resolved down to the particle mass. Limiting ourselves to $M > 10^9$ $M_\odot$, a realistic limit if we insisted on using a halo finder to identify high-redshift structure, would mean that we could not explore values of $M_{\rm cut}$ below 10^9 $M_\odot$. Larger simulations in the future will allow direct application of halo identification at high redshifts on all interesting mass scales.

In Fig. 6.2 we show a visualization of the initial conditions of our Milky Way simulation. The simulation began at $z = 52$, and the particles were initially located on a regular grid before being displaced according to the Zeldovich approximation using the realized density field. In the figure we have plotted as small points initial positions for locations with overdensities on $M = 10^8$ $M_\odot$ scales that exceed the critical overdensity for collapse at $z = 15$. The large squares are the initial positions for locations with overdensities on $M = 10^9$ $M_\odot$ scales that exceed the critical overdensity

for collapse at $z = 8$.

From Fig. 6.2 it is clear that the high overdensities with substantial mass in them are clustered toward the center of the box. This is a consequence of our constraint that a Milky Way-sized galaxy form near the center of the box. However, it is also clear that, while centrally concentrated, the identified peaks are clustered to each other more strongly than they are clustered to the box center. Even in projection, the various clumps of peaks have separations of approximately 2 Mpc (comoving). The mass contained within a spherical region that bounds the identified peaks is approximately 10^{12} $M_\odot$, the total mass of our final Milky Way.

To return to our globular cluster formation model, the two different families of identified peaks represent two different choices of $z_{\rm reion}$ and $M_{\rm cut}$. These particular parameter choices will be justified in the next section. Now, though, we describe the collisionless evolution of these peaks within the formation of our simulated Milky Way, and examine the final configuration.

6.4.2 N-body evolution

It is not computationally demanding to generate cosmological initial conditions of structure formation, even the constrained initial conditions described above. Evolving the growth of structure, in particular the simulated formation of the Milky Way at reasonably high resolution, is a computationally demanding task. However, our goal is to accurately estimate the final spatial distribution of the metal-poor GCS (having identified globular clusters with simulation particles in the previous section), and this can only be accomplished by following the non-linear growth of our simulated Milky Way (but see Rosenblatt et al., 1988, for an analytic estimate).

We evolved our constrained initial conditions through non-linear structure formation using the publicly available N-body gravitational solver `GADGET`[6], a tree-based code (Springel et al., 2001). Though `GADGET` is capable of simulating the evolution of gas using smoothed-particle hydrodynamics, we perform strictly gravitational computations. There are many user-controlled parameters associated with `GADGET`; we adopt values[7] based on investigations of Power et al. (2003), who exten-

[6] available at http://www.mpa-garching.mpg.de/gadget/

[7] The exact set of parameters used is available from the author upon request.

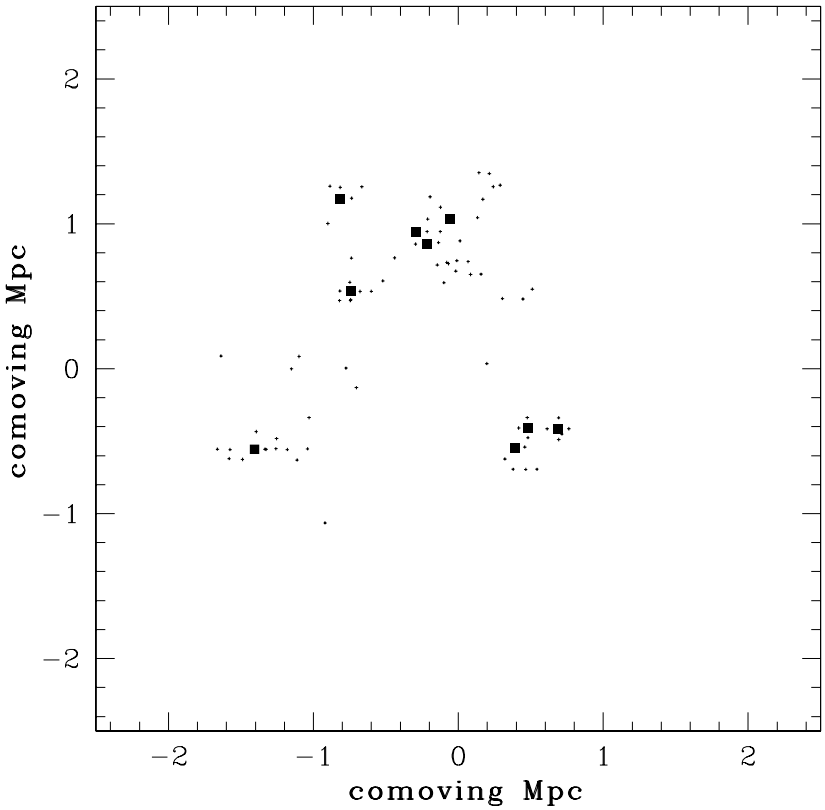


Figure 6.2 Initial positions of sites of metal-poor globular cluster formation. The positions are shown at $z = 52$, for $M = 10^8$ $M_\odot$ peaks that exceed the critical overdensity for collapse at $z = 15$ (small points), and for $M = 10^9$ $M_\odot$ peaks that that exceed the critical overdensity for collapse at $z = 8$. The simulation box size is 5 Mpc on a side; this figure shows the entire simulation box.

sively explored the dependence of the radial density distribution of the inner parts of a dark matter halo on the GADGET input parameters. We comment specifically on the convergence of our simulation in the inner regions when we present our Milky Way simulation results in the next section.

6.5 Milky Way results

In this section we discuss the results of our Milky Way simulation. For a given realization of the constrained initial conditions described in Section 6.4.1, our three free model parameters determine the location and mass of the original Milky Way metal-poor GCS. Then our evolution of these initial conditions to $z = 0$ shows the final spatial extent of the Milky metal-poor GCS, which we compare to both the distribution of the observed Milky metal-poor GCS and the spatial extent of our dark matter halo.

In Fig. 6.2 we showed the spatial distribution of globular cluster formation sites, assuming either $z_{\rm reion} = 15$, $M_{\rm cut} = 10^8$ M$_\odot$, or $z_{\rm reion} = 8$, $M_{\rm cut} = 10^9$ M$_\odot$. These assumptions resulted in the identification of 78 or 9 globular cluster formation sites, respectively. The mean mass of Milky Way globular clusters is now 2.4×10^5 M$_\odot$ (McLaughlin, 1999), and there are 86 globular clusters with [Fe/H] < -1.25 (Harris, 1996), which we identify as the metal-poor population. The total current mass of the Milky Way metal-poor GCS is then 2.1×10^7 M$_\odot$. However, the mass of the metal-poor GCS at formation was larger: both stellar evolution and dynamical disruption decrease the total mass of the GCS in time. Based on the results of Fall & Zhang (2001), we estimate that the mass of the Milky Way metal-poor GCS was approximately 10^8 M$_\odot$ at formation.

The total masses in globular cluster formation sites identified with our $M_{\rm cut} = 10^8$ M$_\odot$ and $M_{\rm cut} = 10^9$ M$_\odot$ parameter choices are 7.8×10^9 M$_\odot$ and 9×10^9 M$_\odot$, respectively. The associated baryonic masses are 1.1×10^9 M$_\odot$ and 1.3×10^9 M$_\odot$. Thus our final free parameter, the efficiency of formation of globular clusters from collapsed baryons at $z_{\rm reion}$, is fixed to match the inferred mass of the original Milky Way metal-poor GCS: $\eta \simeq 0.1$ for either parameter choice. Note that for $M_{\rm cut} = 10^8$ M$_\odot$ we need to form approximately one globular cluster per formation site, whereas for $M_{\rm cut} = 10^9$ M$_\odot$ we need to form ~ 10 globular clusters per formation site.

These results make it clear that, e.g., setting our free parameters to $z_{\rm reion} = 15$, $M_{\rm cut} = 10^9$ M$_\odot$ would produce less baryonic mass to form the Milky Way metal-poor GCS. Thus our requirement that $\eta \leq 1$ combined with the inferred original mass of the Milky Way metal-poor GCS already constrains our choice of our other two free parameters. Future theoretical calculations of η could further restrict the available parameter space of $z_{\rm reion}$ and $M_{\rm cut}$.

We turn now to our predicted spatial distribution of globular clusters in the Milky Way at $z = 0$. Figure 6.3 shows the final spatial distribution of globular cluster formation sites, using the same symbols as Fig. 6.2. The scale of the box is similar to twice the virial radius of the Milky Way (see below). It is immediately clear that globular cluster formation sites are centrally concentrated with respect to the entire dark matter halo. We show this quantitatively in Fig. 6.4, which presents the cumulative radial distribution of the observed Milky Way metal-poor GCS, as well as two predictions from our simulation, and also the cumulative Milky Way halo mass as a function of radius.

The solid curve is the observed radial (galactocentric) distribution of the Milky Way metal-poor GCS (Harris, 1996). Almost all Milky Way globular clusters are contained within a radius of 40 kpc from the Galactic center. The short-dashed and long-dashed curves show the final radial distribution of our simulated Milky Way metal-poor GCS, for $M_{\rm cut} = 10^8$ M$_\odot$ and $M_{\rm cut} = 10^9$ M$_\odot$, respectively, where we have assumed 10 globular clusters formed in each formation site for $M_{\rm cut} = 10^9$ M$_\odot$ (see above). For both model parameter choices, the 40 kpc extent of the Milky Way metal-poor GCS is very well matched. In striking contrast, 75 per cent of the virial mass of the Milky Way halo is located beyond 40 kpc, illustrated by the dotted curve read off against the right axis (note the logarithmic scale). We have demonstrated that a pre-galactic globular cluster formation model may easily match the radial extent of the Milky Way metal-poor GCS (we accomplish this with two different model parameter choices), and that the radial distribution of the simulated GCS does not resemble the radial distribution of the dark matter halo at large radii, because it is much more centrally concentrated.

Within 10 kpc our simulation suffers from numerical effects that prevent us from making a reliable comparison of our radial profiles to observations (future simulations will address this). In addition

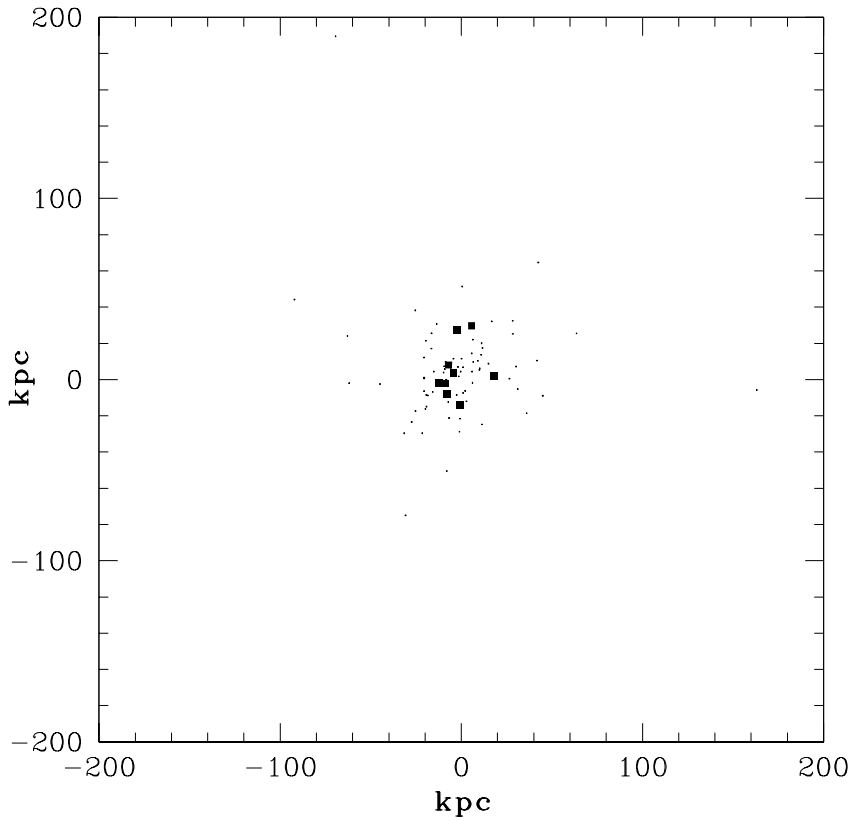


Figure 6.3 Final positions of sites of metal-poor globular cluster formation. The positions are shown at $z = 0$, for $M = 10^8$ $M_\odot$ peaks that exceed the critical overdensity for collapse at $z = 15$ (small points), and for $M = 10^9$ $M_\odot$ peaks that that exceed the critical overdensity for collapse at $z = 8$ (as in Fig. 6.2). The simulation box size is 5 Mpc on a side, and the region shown roughly corresponds to the size of the Milky Way dark matter halo.

to numerical effects, many physical effects may influence the radial profile of globular clusters with 10 kpc of the galactic center. Adiabatic contraction of the baryonic galaxy will pull the inner globular clusters in a little with it. Dynamical friction will also pull inner globular clusters further toward the center. Globular cluster destruction may be more efficient for the inner globular clusters than outer globular clusters, which could change the shape of the radial distribution. We emphasize that none of these considerations, if included in our model, would increase the radial extent of the simulated GCS. Thus the conclusion that our simulations can match the radial extent of the Milky Way metal-poor GCS is robust.

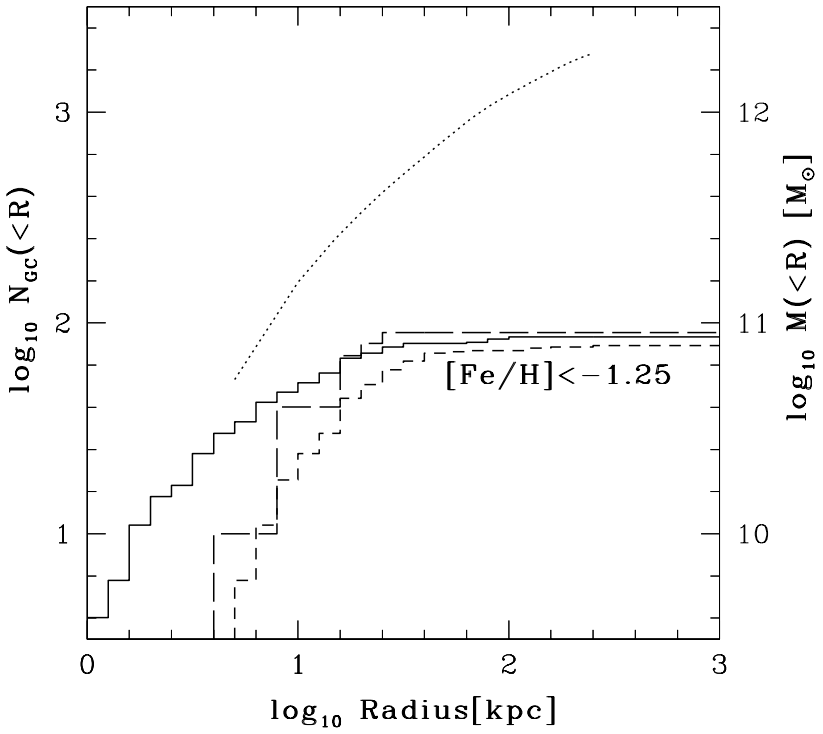


Figure 6.4 Radial distribution of observed and simulated Milky Way globular clusters, and Milky Way dark matter halo. The solid curve is the observed cumulative radial (galactocentric) distribution of the Milky Way metal-poor GCS. The short-dashed and long-dashed curves show the final radial distribution of our simulated Milky Way metal-poor GCS, for $M_{\rm cut} = 10^8$ M$_\odot$ and $M_{\rm cut} = 10^9$ M$_\odot$, respectively, where we have assumed 10 globular clusters formed in each formation site for $M_{\rm cut} = 10^9$ M$_\odot$ (see Section 6.5). The dotted curve, to be read against the right axis, is the enclosed mass as a function of radius of the Milky Way halo in our simulation.

6.6 Environment

We now examine the predictions of our metal-poor GCS formation model in the context of the specific frequency of metal-poor globular clusters as a function of galactic environment. As discussed in the introduction, it is expected that a pre-galactic model such as this one should naturally reproduce the observed trend of larger numbers of metal-poor globular clusters per unit total mass for galaxies in richer environments (West, 1993).

We note that traditionally specific frequency was defined with respect to V-band luminosity (Harris & van den Bergh, 1981, intended for use with early-type galaxies without active star formation), and later an analogous specific frequency with respect to stellar mass was introduced (Zepf & Ashman, 1993, to be applied to more general galactic conditions, such as spirals and star-forming mergers). However, our model is only suitable for calculating the number of metal-poor globular clusters per unit total (baryonic plus dark) matter of a galaxy. Star-formation and population-synthesis modelling could be used to quantitatively relate our specific frequency quantity to other definitions; we will assume that stellar mass scales with total mass, and thus our specific frequency is proportion to other measures.

6.6.1 Schematic argument for increasing specific frequency with increasing richness of galactic environment

First we present a schematic argument demonstrating why, statistically, galaxies in denser environments have higher metal-poor specific frequencies than similar galaxies in the field. Figure 6.5 shows cartoon depictions of two galaxies. In this figure initial overdensity is shown on the vertical axis. Because higher overdensities collapse earlier, overdensity is directly related to collapse redshift: points closer to the top collapse before lower points. The horizontal axis represents a spatial scale. The mass of a collapsed halo at redshift $z_{\rm collapse}$ is proportional to the cross-section of a horizontal line drawn at $z_{\rm collapse}$ and the structure drawing.

In this figure galaxy formation is simplified into a two-step process. First low-mass, high-redshift

proto-galactic halos collapse, then later a roughly monolithic inside-out collapse of the galaxy incorporates the pre-galactic structure into a massive galaxy. Real conditions are of course more complex than this, and modelled numerically in the next subsection. This figure illustrates the formation of two galaxies with identical final mass, and strictly similar dark-matter collapse histories (though one collapses earlier, the process of collapse is just rescaled in time between both). However, one galaxy is located in a cluster of galaxies, and the other in the field. The cluster galaxy sits on top of a pedestal of overdensity associated with a much larger scale, the cluster. So though the galactic (and sub-galactic) perturbations for each galaxy are the same, the peak-on-peak effect of the cluster overdensity causes the cluster galaxy to collapse before the field galaxy.

If we examine the collapsed structure of both galaxies at a fixed redshift, $z_{\rm reion}$, we find that the cluster galaxy has 15 pre-galactic halos collapsed, compared to only 7 for the field galaxy. If metal-poor globular cluster formation were proportional to the amount of collapsed structure before reionization, as we have assumed, then we would expect an enhancement (a factor of 2 in this cartoon) in the specific frequency of the cluster galaxy with respect to the field galaxy (the galaxies have the same mass by construction). Figure 6.5 shows that the variation of specific frequency with environment depends on the choice of $z_{\rm reion}$ (it also depends on $M_{\rm cut}$, though this is not illustrated here).

The concentrated nature of the galaxy-scale perturbation shows that even though the pre-galactic structure is not concentrated toward the future center of the galaxy, the distribution of globular cluster formation sites is concentrated toward the center of the future galaxy, relative to the total mass of the galaxy (the width of the galaxy pentagon).

We note that it is not environment per se, but rather galaxy collapse history, that is responsible for the variation of specific frequency with environment. Correlations naturally present in the initial conditions of structure formation, illustrated with the "peak-patch" split we used to describe Fig. 6.5, dictate that for otherwise similar galaxies, galaxies in richer environments statistically form earlier. A field elliptical with a cluster-like specific frequency would therefore not be surprising, though such galaxies should be the exception rather than the rule. Future analytic calculations and numerical

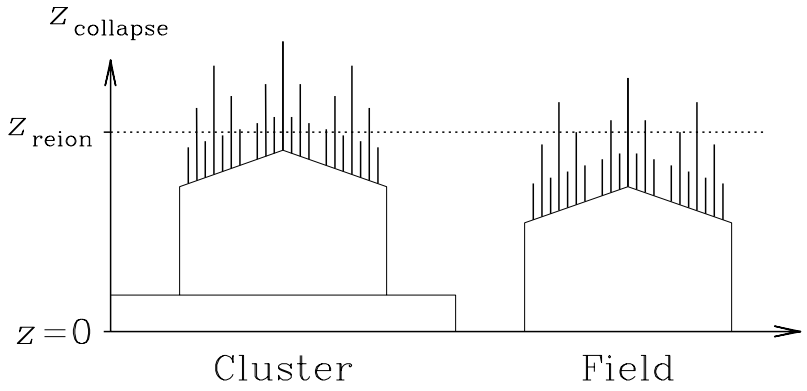


Figure 6.5 Schematic illustration of dependence of specific frequency on galactic environment. The ordinate represents the collapse redshift of points on the plot: points at larger values of $z_{\rm collapse}$ collapse earlier. The abscissa is a spatial scale: the collapsed mass at redshift z is proportional to the cross-section of a line drawn at $z_{\rm collapse} = z$ and the figures. The thin vertical "forests" of lines represent low-mass, pre-galactic structure; the pentagons under the forest represent galaxy-scale structure; and the broad pedestal above "Cluster" represents cluster of galaxies scale structure. Thus, as cosmic time moves forward in this figure, first low-mass peaks collapse, then galaxies, then clusters in some locations. In our model metal-poor globular clusters form in structure collapsed before reionization. Under this scheme the vertical peaks that cross above the dashed line labelled $z_{\rm reion}$ are our globular cluster formation sites. This figure illustrates that, for otherwise similar galaxies, galaxies in rich environments are expected to have more metal-poor globular clusters per unit galaxy total mass than field galaxies have.

simulations could quantify the likelihood of this phenomenon.

6.6.2 Semi-quantitative analysis

Figure 6.6 shows a more quantitative estimate of the specific frequency of metal-poor globular clusters as a function of galaxy collapse history, which is statistically related to galaxy environment. The points show the number of metal-poor globular clusters per unit *predicted* total halo mass at $z = 0$, as a function of the overdensity of the galaxy-scale density perturbation. The specific frequency is arbitrarily normalized (see our discussion at the beginning of this section). The overdensity is defined as the linearly-extrapolated $z = 0$ overdensity averaged over the size of the overdensity constraint imposed when we generated our initial conditions. Larger overdensity corresponds to earlier collapse, and thus richer galactic environments on average.

The points here were computed using our $M_{\rm cut} = 10^8$ M$_\odot$, $z_{\rm reion} = 15$ parameter set; our

predicted specific frequencies are sensitive to these choices (as discussed in the previous subsection). The two sets of points (connected by lines) represent two different predicted galactic mass scales. The lower set of points were calculated to be representative of a spiral like the Milky Way, and the upper points should resemble a massive elliptical. It is a general prediction of the model that more massive halos should have higher specific frequencies than less massive halos, at least for massive galaxies (dwarf galaxies are a more subtle issue, since for low-mass halos assumption of a constant stellar mass to halo mass is likely a very bad assumption).

Unfortunately we do not yet have available the extensive numerical simulations required to calibrate our predicted total halo mass (based on the spherical overdensity criteria applied to the constrained initial conditions) to the final virialized mass at $z = 0$. Our computed specific frequency values do show the expected increase with earlier collapse history, and thus richness of galactic environment. We can easily generate variations in specific frequency up to two orders of magnitude, if we allow collapse of the galaxy-scale perturbation to occur as early as $z = 5.7$ (corresponding to an overdensity of 8.8). If the galaxy-scale perturbation collapses closer to $z = 2$, probably a more representative value for a cluster central elliptical galaxy, then we estimate a specific frequency enhancement of about 10 compared to the field, consistent with observations (e.g. Harris, 1991).

These considerations also demonstrate that a comparison of the metal-poor GCS of the Local Group to an equally massive elliptical should show that the Local Group has a lower specific frequency than the elliptical, because obviously the elliptical collapsed before the Local Group (which has not yet collapsed). This is exactly what an observational study discovered (Forbes et al., 2000).

6.7 Summary

The simple scenario for metal-poor GCS formation presented here has so far shown promise as a model. We have introduced cosmic reionization as a new element of GCS formation theory, a physically motivated mechanism to suppress intermediate-metallicity globular cluster formation and generate the observed bimodal globular cluster metallicity distributions. This hypothesis of fixed GCS formation time creates a simple framework within which we can calculate many properties of

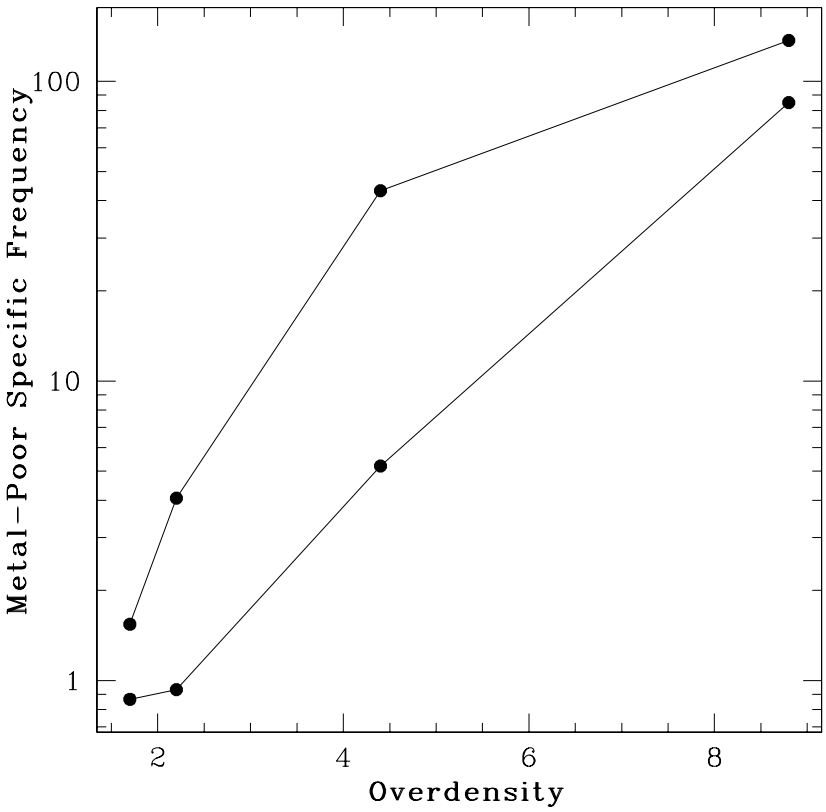


Figure 6.6 The specific frequency of metal-poor globular clusters. This figure shows the number of metal-poor globular clusters per unit *predicted* total halo mass at $z = 0$, as a function of the overdensity of the galaxy-scale density perturbation. The specific frequency is arbitrarily normalized. The abscissa shows a measure of how early a given galaxy collapses: galaxies with higher overdensity values collapse earlier; lines connect points of similar mass but different collapse histories. Early collapse is likely to correlate with denser galaxy environments. See Section 6.6.2 for a more complete description.

metal-poor GCSs, as exploited by earlier work (e.g., Rosenblatt et al., 1988; West, 1993).

Our model has three free parameters, which we constrained by requiring our model to reproduce the total mass and extent of the Milky Way metal-poor GCS. Once the parameters are fixed, we can generate the radial distribution of the metal-poor GCS and the specific frequency as a function of mass and collapse history, which is closely related to environment. We demonstrated that our model (and, by extension, other pre-galactic models of globular cluster formation) is capable of reproducing the observed radial extent of the Milky Way metal-poor globular cluster system. We can also match the observed trend of increasing metal-poor globular cluster specific frequency with increasing richness of galactic environment. This is consequence not of environment itself, but rather the correlation between rich environments and early (in cosmic time) galactic collapse histories.

Our model may be applied in the future to many more interesting problems in globular cluster system formation. For example, the kinematics of the metal-poor GCS are also predicted. As another specific example, in Section 6.5 we showed two solutions within our parameter space that matched the observed mass and spatial extent of the Milky Way metal-poor GCS. One of these solutions formed approximately one globular cluster per formation site, and the other solution formed approximately 10 globular clusters per formation site. In the latter case, dynamical evolution implies that we would not expect to observe all 10 globular clusters right next to each other today. However, we would expect them to trace out a common orbital path. Two such streams of globular clusters have recently been discovered (Bellazzini et al., 2003). Future high-precision astrometry will map the phase-space distribution of the Milky Way GCS, allowing further quantitative measurement of the fraction of globular clusters in dynamically associated streams. This model may also make other predictions, such as the dependence of the radial distribution and kinematics as a function of galaxy mass and environment, specific predictions about extragalactic metal-poor globular clusters, and specific frequencies of globular clusters in dwarf galaxies.

However, our model also has some weaknesses. We do not simulate the process of globular cluster formation: we assign a metallicity of "metal-poor" by assumption. We calculate specific frequency always with respect to total baryonic plus dark matter mass of a galaxy and its halo; this is not in

general a direct observable. It will be difficult for our model to make predictions for the exact shape of the radial distribution of metal-poor GCSs at small galactic radii until a treatment of globular cluster destruction is incorporated (and a similar caveat applies to studying kinematics).

On balance, we have used numerical techniques to confirm earlier analytic conclusions that a "pre-galactic" model of globular cluster formation may do an excellent job reproducing many observed properties of the Milky Way metal-poor GCS and trends observed between metal-poor GCSs of galaxies in different environments. Future work will continue to compare the predictions of our model with the wealth of unexplained observational data, as well as make predictions for new observations.

Acknowledgments

We thank Mike Fall, Peter Goldreich, Mike Rich, Mike Beasley, and especially Kathy Rhode for many helpful conversations. We also thank the organizers and participants of the ESO Workshop on Extragalactic Globular Cluster Systems, where many of these ideas were formulated. We acknowledge the support of NASA GSRP grant NGT5-50339.

Bibliography

Abel, T., Anninos, P., Norman, M. L., & Zhang, Y. 1998, Ap.J., 508, 518

Abel, T., Bryan, G. L., & Norman, M. L. 2000, Ap.J., 540, 39

Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, Ap.J., 584, 45

Agol, E. & Kamionkowski, M. 2002, M.N.R.A.S., 334, 553

Ajiki, M., Taniguchi, Y., Fujita, S. S., Shioya, Y., Nagao, T., Murayama, T., Yamada, S., Umeda, K., & Komiyama, Y. 2003, A.J., 126, 2091

Alcock, C., Allsman, R. A., Alves, D. R., Axelrod, T. S., Becker, A. C., Bennett, D. P., Cook, K. H., Dalal, N., Drake, A. J., Freeman, K. C., Geha, M., Griest, K., Lehner, M. J., Marshall, S. L., Minniti, D., Nelson, C. A., Peterson, B. A., Popowski, P., Pratt, M. R., Quinn, P. J., Stubbs, C. W., Sutherland, W., Tomaney, A. B., Vandehei, T., & Welch, D. L. 2001, Ap.J., 550, L169

Anderson, S. F., Fan, X., Richards, G. T., Schneider, D. P., Strauss, M. A., Vanden Berk, D. E., Gunn, J. E., Knapp, G. R., Schlegel, D., Voges, W., Yanny, B., Bahcall, N. A., Bernardi, M., Brinkmann, J., Brunner, R., Csabái, I., Doi, M., Fukugita, M., Hennessy, G. S., Ivezić, Ž., Kunszt, P. Z., Lamb, D. Q., Loveday, J., Lupton, R. H., McKay, T. A., Munn, J. A., Nichol, R. C., Szokoly, G. P., & York, D. G. 2001, A.J., 122, 503

Armandroff, T. E. & Zinn, R. 1988, A.J., 96, 92

Ashman, K. M. & Zepf, S. E. 1992, Ap.J., 384, 50

—. 1998, Globular cluster systems (Globular cluster systems / Keith M. Ashman, Stephen

E. Zepf. Cambridge, U. K. ; New York : Cambridge University Press, 1998. (Cambridge astrophysics series ; 30))

—. 2001, A.J., 122, 1888

Bézecourt, J., Kneib, J. P., Soucail, G., & Ebbels, T. M. D. 1999, A.&A., 347, 21

Barger, A. J., Cowie, L. L., Capak, P., Alexander, D. M., Bauer, F. E., Brandt, W. N., Garmire, G. P., & Hornschemeier, A. E. 2003, Ap.J., 584, L61

Barkana, R. 2004, M.N.R.A.S., 347, 59

Barkana, R. & Loeb, A. 1999, Ap.J., 523, 54

—. 2000, Ap.J., 539, 20

—. 2001, Phys. Rep., 349, 125

—. 2003, Nature, 421, 341

—. 2004, Ap.J., 601, 64

Beasley, M. A., Baugh, C. M., Forbes, D. A., Sharples, R. M., & Frenk, C. S. 2002, M.N.R.A.S., 333, 383

Bechtold, J. 1994, Ap.J.S., 91, 1

Bechtold, J. 2003, in Galaxies at high redshift. XI Canary Islands Winter School of Astrophysics, Santa Cruz de Tenerife, Spain, November 15 - 26, 1999, edited by I. Pérez-Fournon, M. Balcells, F. Moreno-Insertis and F. Sánchez. Cambridge, UK: Cambridge University Press, p. 131 - 184, 131

Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., & York, D. G. 2001, A.J., 122, 2850

Bellazzini, M., Ferraro, F. R., & Ibata, R. 2003, A.J., 125, 188

Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002a, M.N.R.A.S., 333, 177

Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002b, M.N.R.A.S., 333, 156

Bernstein, R. A., Freedman, W. L., & Madore, B. F. 2002, Ap.J., 571, 107

Bertschinger, E. 1998, A.R.A.&A., 36, 599

Birnboim, Y. & Dekel, A. 2003, M.N.R.A.S., 345, 349

Blandford, R. D. & Narayan, R. 1992, A.R.A.&A., 30, 311

Bond, J. R., Carr, B. J., & Hogan, C. J. 1986, Ap.J., 306, 428

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

Bromm, V., Coppi, P. S., & Larson, R. B. 1999, Ap.J., 527, L5

—. 2002, Ap.J., 564, 23

Bromm, V., Ferrara, A., Coppi, P. S., & Larson, R. B. 2001a, M.N.R.A.S., 328, 969

Bromm, V., Kudritzki, R. P., & Loeb, A. 2001b, Ap.J., 552, 464

Burles, S., Nollett, K. M., & Turner, M. S. 2001, Ap.J., 552, L1

Callaway, J. 1985, Phys. Rev. A, 32, 775

Callaway, J., Unnikrishnan, K., & Oza, D. H. 1987, Phys. Rev. A, 36, 2576

Cambrésy, L., Reach, W. T., Beichman, C. A., & Jarrett, T. H. 2001, Ap.J., 555, 563

Carr, B. 1994, A.R.A.&A., 32, 531

Carr, B. J., Bond, J. R., & Arnett, W. D. 1984, Ap.J., 277, 445

Cen, R. 1992, Ap.J.S., 78, 341

—. 2003a, Ap.J., 591, L5

—. 2003b, Ap.J., 591, 12

Chapman, S. C., Blain, A. W., Ivison, R. J., & Smail, I. R. 2003, Nature, 422, 695

Charlot, S. & Fall, S. M. 2000, Ap.J., 539, 718

Ciardi, B., Ferrara, A., Marri, S., & Raimondo, G. 2001, M.N.R.A.S., 324, 381

Ciardi, B., Ferrara, A., & White, S. D. M. 2003, M.N.R.A.S., 344, L7

Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, M.N.R.A.S., 319, 168

Cooray, A., Bock, J. J., Keating, B., Lange, A. E., & Matsumoto, T. 2004, Ap.J., 606, 611

Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1

Couchman, H. M. P. & Rees, M. J. 1986, M.N.R.A.S., 221, 53

Dekel, A. & Silk, J. 1986, Ap.J., 303, 39

Dey, A., Spinrad, H., Stern, D., Graham, J. R., & Chaffee, F. H. 1998, Ap.J., 498, L93

Djorgovski, S. G., Castro, S., Stern, D., & Mahabal, A. A. 2001, Ap.J., 560, L5

Dobrzycki, A., Bechtold, J., Scott, J., & Morita, M. 2002, Ap.J., 571, 654

Dwek, E., Arendt, R. G., Hauser, M. G., Fixsen, D., Kelsall, T., Leisawitz, D., Pei, Y. C., Wright, E. L., Mather, J. C., Moseley, S. H., Odegard, N., Shafer, R., Silverberg, R. F., & Weiland, J. L. 1998, Ap.J., 508, 106

Eisenstein, D. J. & Hu, W. 1999, Ap.J., 511, 5

Elbaz, D., Cesarsky, C. J., Chanial, P., Aussel, H., Franceschini, A., Fadda, D., & Chary, R. R. 2002, A.&A., 384, 848

Ellis, R., Santos, M. R., Kneib, J., & Kuijken, K. 2001, Ap.J., 560, L119

Fall, S. M. & Zhang, Q. 2001, Ap.J., 561, 751

Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., & Rix, H. 2002, A.J., 123, 1247

Fan, X., Strauss, M. A., Richards, G. T., Newman, J. A., Becker, R. H., Schneider, D. P., Gunn, J. E., Davis, M., White, R. L., Lupton, R. H., Anderson, J. E., Annis, J., Bahcall, N. A., Brunner, R. J., Csabai, I., Doi, M., Fukugita, M., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Knapp, G. R., McKay, T. A., Munn, J. A., Pier, J. R., Szalay, A. S., & York, D. G. 2001a, A.J., 121, 31

Fan, X., Strauss, M. A., Schneider, D. P., Becker, R. H., White, R. L., Haiman, Z., Gregg, M., Pentericci, L., Grebel, E. K., Narayanan, V. K., Loh, Y., Richards, G. T., Gunn, J. E., Lupton, R. H., Knapp, G. R., Ivezić, Ž., Brandt, W. N., Collinge, M., Hao, L., Harbeck, D., Prada, F., Schaye, J., Strateva, I., Zakamska, N., Anderson, S., Brinkmann, J., Bahcall, N. A., Lamb, D. Q., Okamura, S., Szalay, A., & York, D. G. 2003, A.J., 125, 1649

Fan, X., Strauss, M. A., Schneider, D. P., Gunn, J. E., Lupton, R. H., Anderson, S. F., Voges, W., Margon, B., Annis, J., Bahcall, N. A., Brinkmann, J., Brunner, R. J., Carr, M. A., Csabai, I., Doi, M., Frieman, J. A., Fukugita, M., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Newberg, H. J., Pauls, A. G., Pier, J. R., Rechenmacher, R., Richards, G. T., Rockosi, C. M., Stoughton, C., Szalay, A. S., Thakar, A. R., Tucker, D. L., Waddell, P., & York, D. G. 2000, A.J., 119, 1

Fan, X., Strauss, M. A., Schneider, D. P., Gunn, J. E., Lupton, R. H., Becker, R. H., Davis, M., Newman, J. A., Richards, G. T., White, R. L., Anderson, J. E., Annis, J., Bahcall, N. A., Brunner, R. J., Csabai, I., Hennessy, G. S., Hindsley, R. B., Fukugita, M., Kunszt, P. Z., Ivezić, Ž., Knapp, G. R., McKay, T. A., Munn, J. A., Pier, J. R., Szalay, A. S., & York, D. G. 2001b, A.J., 121, 54

Fan, X., Strauss, M. A., Schneider, D. P., Gunn, J. E., Lupton, R. H., Yanny, B., Anderson, S. F., Anderson, J. E., Annis, J., Bahcall, N. A., Bakken, J. A., Bastian, S., Berman, E., Boroski, W. N., Briegel, C., Briggs, J. W., Brinkmann, J., Carr, M. A., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Elms, B. R., Evans, M. L., Federwitz, G. R., Frieman, J. A., Fukugita, M., Gurbani, V. K., Harris, F. H., Heckman, T. M., Hennessy,

G. S., Hindsley, R. B., Holmgren, D. J., Hull, C., Ichikawa, S., Ichikawa, T., Ivezić , Ž., Kent, S., Knapp, G. R., Kron, R. G., Lamb, D. Q., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loveday, J., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McKay, T. A., Munn, J. A., Nash, T., Newberg, H. J., Nichol, R. C., Nicinski, T., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Petravick, D., Pier, J. R., Pordes, R., Prosapio, A., Rechenmacher, R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Sandford, D., Sergey, G., Sekiguchi, M., Shimasaku, K., Siegmund, W. A., Smith, J. A., Stoughton, C., Szalay, A. S., Szokoly, G. P., Tucker, D. L., Vogeley, M. S., Waddell, P., Wang, S., Weinberg, D. H., Yasuda, N., & York, D. G. 1999, A.J., 118, 1

Fardal, M. A., Katz, N., Gardner, J. P., Hernquist, L., Weinberg, D. H., & Davé, R. 2001, Ap.J., 562, 605

Ferland, G. J. 1980, P.A.S.P., 92, 596

Fontana, A., Poli, F., Menci, N., Nonino, M., Giallongo, E., Cristiani, S., & D'Odorico, S. 2003, Ap.J., 587, 544

Forbes, D. A., Masters, K. L., Minniti, D., & Barmby, P. 2000, A.&A., 358, 471

Franx, M., Illingworth, G. D., Kelson, D. D., van Dokkum, P. G., & Tran, K. 1997, Ap.J., 486, L75

Frye, B., Broadhurst, T., & Benítez, N. 2002, Ap.J., 568, 558

Fryer, C. L., Woosley, S. E., & Heger, A. 2001, Ap.J., 550, 372

Gallego, J., Zamorano, J., Aragon-Salamanca, A., & Rego, M. 1995, Ap.J., 455, L1

Gnedin, N. Y. 2000, Ap.J., 535, 530

Haardt, F. & Madau, P. 1996, Ap.J., 461, 20

Haiman, Z. 2002, Ap.J., 576, L1

Haiman, Z. & Holder, G. P. 2003, Ap.J., 595, 1

Haiman, Z. & Loeb, A. 1997, Ap.J., 483, 21

Haiman, Z. & Spaans, M. 1999, Ap.J., 518, 138

Haiman, Z., Spaans, M., & Quataert, E. 2000, Ap.J., 537, L5

Hamana, T., Ouchi, M., Shimasaku, K., Kayo, I., & Suto, Y. 2004, M.N.R.A.S., 347, 813

Harris, W. E. 1991, A.R.A.&A., 29, 543

—. 1996, A.J., 112, 1487

Harris, W. E. 2003, in Extragalactic Globular Cluster Systems, Proceedings of the ESO Workshop held in Garching, Germany, 27-30 August 2002, 317

Harris, W. E. & Harris, G. L. H. 2002, A.J., 123, 3108

Harris, W. E. & van den Bergh, S. 1981, A.J., 86, 1627

Hauser, M. G. & Dwek, E. 2001, A.R.A.&A., 39, 249

Heger, A. & Woosley, S. E. 2002, Ap.J., 567, 532

Hewitt, A. & Burbidge, G. 1987, Ap.J.S., 63, 1

Hoffman, Y. & Ribak, E. 1991, Ap.J., 380, L5

Hu, E. M., Cowie, L. L., & McMahon, R. G. 1998, Ap.J., 502, L99

Hu, E. M., Cowie, L. L., McMahon, R. G., Capak, P., Iwamuro, F., Kneib, J.-P., Maihara, T., & Motohara, K. 2002a, Ap.J., 568, L75

—. 2002b, Ap.J., 576, L99

Hu, E. M., McMahon, R. G., & Cowie, L. L. 1999, Ap.J., 522, L9

Ivison, R. J., Greve, T. R., Smail, I., Dunlop, J. S., Roche, N. D., Scott, S. E., Page, M. J., Stevens, J. A., Almaini, O., Blain, A. W., Willott, C. J., Fox, M. J., Gilbank, D. G., Serjeant, S., & Hughes, D. H. 2002, M.N.R.A.S., 337, 1

Iwata, I., Ohta, K., Tamura, N., Ando, M., Wada, S., Watanabe, C., Akiyama, M., & Aoki, K. 2003, P.A.S.J., 55, 415

Jakobsen, P. 1998, A.&A., 335, 876

Karzas, W. J. & Latter, R. 1961, Ap.J.S., 6, 167

Kennicutt, R. C. 1998, A.R.A.&A., 36, 189

Kim, T., Hu, E. M., Cowie, L. L., & Songaila, A. 1997, A.J., 114, 1

Kneib, J.-P. 1993, Ph.D. Thesis

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, Ap.J., 471, 643

Kodaira, K., Taniguchi, Y., Kashikawa, N., Kaifu, N., Ando, H., Karoji, H., Ajiki, M., Akiyama, M., Aoki, K., Doi, M., Fujita, S. S., Furusawa, H., Hayashino, T., Imanishi, M., Iwamuro, F., Iye, M., Kawabata, K. S., Kobayashi, N., Kodama, T., Komiyama, Y., Kosugi, G., Matsuda, Y., Miyazaki, S., Mizumoto, Y., Motohara, K., Murayama, T., Nagao, T., Nariai, K., Ohta, K., Ohyama, Y., Okamura, S., Ouchi, M., Sasaki, T., Sekiguchi, K., Shimasaku, K., Shioya, Y., Takata, T., Tamura, H., Terada, H., Umemura, M., Usuda, T., Yagi, M., Yamada, T., Yasuda, N., & Yoshida, M. 2003, P.A.S.J., 55, L17

Kogut, A., Spergel, D. N., Barnes, C., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Wollack, E., & Wright, E. L. 2003, Ap.J.S., 148, 161

Krauss, L. M. & Chaboyer, B. 2003, Science, 299, 65

Kravtsov, A. V. & Gnedin, O. Y. 2003, Ap.J., submitted, astro-ph/0305199

Kriss, G. A., Shull, J. M., Oegerle, W., Zheng, W., Davidsen, A. F., Songaila, A., Tumlinson, J., Cowie, L. L., Deharveng, J.-M., Friedman, S. D., Giroux, M. L., Green, R. F., Hutchings, J. B., Jenkins, E. B., Kruk, J. W., Moos, H. W., Morton, D. C., Sembach, K. R., & Tripp, T. M. 2001, Science, 293, 1112

Kunth, D., Mas-Hesse, J. M., Terlevich, E., Terlevich, R., Lequeux, J., & Fall, S. M. 1998, A.&A., 334, 11

Lacey, C. & Cole, S. 1993, M.N.R.A.S., 262, 627

Larsen, S. S., Brodie, J. P., Huchra, J. P., Forbes, D. A., & Grillmair, C. J. 2001, A.J., 121, 2974

Lehnert, M. D. & Bremer, M. 2003, Ap.J., 593, 630

Leitherer, C., Schaerer, D., Goldader, J. D., Delgado, R. M. G., Robert, C., Kune, D. F., de Mello, D. F., Devost, D., & Heckman, T. M. 1999, Ap.J.S., 123, 3

Lilly, S., Tran, K., Brodwin, M., Crampton, D., Juneau, S., & McCracken, H. 2003, Ap.J., submitted, astro-ph/0304376

Liske, J., Webb, J. K., Williger, G. M., Fernández-Soto, A., & Carswell, R. F. 2000, M.N.R.A.S., 311, 657

Loeb, A. & Barkana, R. 2001, A.R.A.&A., 39, 19

Loeb, A. & Rybicki, G. B. 1999, Ap.J., 524, 527

Madau, P. & Rees, M. J. 2001, Ap.J., 551, L27

Magliocchetti, M., Salvaterra, R., & Ferrara, A. 2003, M.N.R.A.S., 342, L25

Maier, C., Meisenheimer, K., Thommes, E., Hippelein, H., Röser, H. J., Fried, J., von Kuhlmann, B., Phleps, S., & Wolf, C. 2003, A.&A., 402, 79

Massey, P. & Gronwall, C. 1990, Ap.J., 358, 344

Mather, J. C. & Stockman, H. S. 2000, in Proc. SPIE Vol. 4013, UV, Optical, and IR Space Telescopes and Instruments, James B. Breckinridge; Peter Jakobsen; Eds., 2–16

McCarthy, J. K., Cohen, J. G., Butcher, B., Cromer, J., Croner, E., Douglas, W. R., Goeden, R. M., Grewal, T., Lu, B., Petrie, H. L., Weng, T., Weber, B., Koch, D. G., & Rodgers, J. M. 1998, in Proc. SPIE Vol. 3355, Optical Astronomical Instrumentation, Sandro D'Odorico; Ed., 81–92

McLaughlin, D. E. 1999, A.J., 117, 2398

Miralda-Escude, J. 1998, Ap.J., 501, 15

Miralda-Escude, J. & Ostriker, J. P. 1992, Ap.J., 392, 15

Miralda-Escude, J. & Rees, M. J. 1998, Ap.J., 497, 21

Mo, H. J., Mao, S., & White, S. D. M. 1998, M.N.R.A.S., 295, 319

Mo, H. J. & White, S. D. M. 2002, M.N.R.A.S., 336, 112

Moller, P. & Jakobsen, P. 1990, A.&A., 228, 299

Moore, B. 1996, Ap.J., 461, L13

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999, Ap.J., 524, L19

Neufeld, D. A. 1991, Ap.J., 370, L85

Oh, S. P. 2001, Ap.J., 553, 499

Oh, S. P., Nollett, K. M., Madau, P., & Wasserburg, G. J. 2001, Ap.J., 562, L1

Oke, J. B., Cohen, J. G., Carr, M., Cromer, J., Dingizian, A., Harris, F. H., Labrecque, S., Lucinio, R., Schaal, W., Epps, H., & Miller, J. 1995, P.A.S.P., 107, 375

Osterbrock, D. E. 1989, Astrophysics of gaseous nebulae and active galactic nuclei (Research supported by the University of California, John Simon Guggenheim Memorial Foundation, University of Minnesota, et al. Mill Valley, CA, University Science Books, 1989, 422 p.)

Ouchi, M., Shimasaku, K., Furusawa, H., Miyazaki, M., Doi, M., Hamabe, M., Hayashino, T., Kimura, M., Kodaira, K., Komiyama, Y., Matsuda, Y., Miyazaki, S., Nakata, F., Okamura, S., Sekiguchi, M., Shioya, Y., Tamura, H., Taniguchi, Y., Yagi, M., & Yasuda, N. 2003, Ap.J., 582, 60

Padmanabhan, T. 1993, Structure formation in the universe (Cambridge, UK: Cambridge University Press, —c1993)

Partridge, R. B. & Peebles, P. J. E. 1967, Ap.J., 147, 868

Peebles, P. J. E. 1993, Principles of physical cosmology (Princeton Series in Physics, Princeton, NJ: Princeton University Press, —c1993)

Peebles, P. J. E. & Dicke, R. H. 1968, Ap.J., 154, 891

Pelló, R., Kneib, J. P., Le Borgne, J. F., Bézecourt, J., Ebbels, T. M., Tijera, I., Bruzual, G., Miralles, J. M., Smail, I., Soucail, G., & Bridges, T. J. 1999, A.&A., 346, 359

Pettini, M., Madau, P., Bolte, M., Prochaska, J. X., Ellison, S. L., & Fan, X. 2003, Ap.J., 594, 695

Power, C., Navarro, J. F., Jenkins, A., Frenk, C. S., White, S. D. M., Springel, V., Stadel, J., & Quinn, T. 2003, M.N.R.A.S., 338, 14

Press, W. H. & Schechter, P. 1974, Ap.J., 187, 425

Pritchet, C. J. 1994, P.A.S.P., 106, 1052

Prochaska, J. X., Wolfe, A. M., Tytler, D., Burles, S., Cooke, J., Gawiser, E., Kirkman, D., O'Meara, J. M., & Storrie-Lombardi, L. 2001, Ap.J.S., 137, 21

Renault, C., Barrau, A., Lagache, G., & Puget, J.-L. 2001, A.&A., 371, 771

Reuland, M., van Breugel, W., Röttgering, H., de Vries, W., Stanford, S. A., Dey, A., Lacy, M., Bland-Hawthorn, J., Dopita, M., & Miley, G. 2003, Ap.J., 592, 755

Rhoads, J. E., Dey, A., Malhotra, S., Stern, D., Spinrad, H., Jannuzi, B. T., Dawson, S., Brown, M. J. I., & Landes, E. 2003, A.J., 125, 1006

Rhoads, J. E. & Malhotra, S. 2001, Ap.J., 563, L5

Rhoads, J. E., Malhotra, S., Dey, A., Stern, D., Spinrad, H., & Jannuzi, B. T. 2000, Ap.J., 545, L85

Ricotti, M. 2002, M.N.R.A.S., 336, L33

Rosenblatt, E. I., Faber, S. M., & Blumenthal, G. R. 1988, Ap.J., 330, 191

Salvaterra, R. & Ferrara, A. 2003, M.N.R.A.S., 339, 973

Santos, M. R. 2004a, M.N.R.A.S., 349, 1137

—. 2004b, in preparation

Santos, M. R., Bromm, V., & Kamionkowski, M. 2002, M.N.R.A.S., 336, 1082

Santos, M. R., Ellis, R. S., Kneib, J., Richard, J., & Kuijken, K. 2004, Ap.J., 606, 683

Santos, M. R. & Loeb, A. 2003, MNRAS, submitted, astro-ph/0304130

Sargent, W. L. W., Steidel, C. C., & Boksenberg, A. 1989, Ap.J.S., 69, 703

Scannapieco, E. & Barkana, R. 2002, Ap.J., 571, 585

Scannapieco, E. & Broadhurst, T. 2001, Ap.J., 549, 28

Scannapieco, E., Ferrara, A., & Broadhurst, T. 2000, Ap.J., 536, L11

Schaerer, D. 2002, A.&A., 382, 28

—. 2003, A.&A., 397, 527

Schaye, J. 2001, Ap.J., 559, 507

Schechter, P. 1976, Ap.J., 203, 297

Scheffler, H. & Elsaesser, H. 1987, Physics of the galaxy and interstellar matter (Berlin and New York, Springer-Verlag, 1987, 503 p. Translation.)

Schneider, D. P., Richards, G. T., Fan, X., Hall, P. B., Strauss, M. A., Vanden Berk, D. E., Gunn, J. E., Newberg, H. J., Reichard, T. A., Stoughton, C., Voges, W., Yanny, B., Anderson, S. F., Annis, J., Bahcall, N. A., Bauer, A., Bernardi, M., Blanton, M. R., Boroski, W. N., Brinkmann, J., Briggs, J. W., Brunner, R., Burles, S., Carey, L., Castander, F. J., Connolly, A. J., Csabai, I., Doi, M., Friedman, S., Frieman, J. A., Fukugita, M., Heckman, T. M., Hennessy, G. S., Hindsley,

R. B., Hogg, D. W., Ivezić, Ž., Kent, S., Knapp, G. R., Kunzst, P. Z., Lamb, D. Q., Leger, R. F., Long, D. C., Loveday, J., Lupton, R. H., Margon, B., Meiksin, A., Merelli, A., Munn, J. A., Newcomb, M., Nichol, R. C., Owen, R., Pier, J. R., Pope, A., Rockosi, C. M., Saxe, D. H., Schlegel, D., Siegmund, W. A., Smee, S., Snir, Y., SubbaRao, M., Szalay, A. S., Thakar, A. R., Uomoto, A., Waddell, P., & York, D. G. 2002a, A.J., 123, 567

Schneider, R., Ferrara, A., Natarajan, P., & Omukai, K. 2002b, Ap.J., 571, 30

Schweizer, F. 1986, Science, 231, 227

Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P. 2000, Ap.J.S., 130, 67

Scott, J., Bechtold, J., Morita, M., Dobrzycki, A., & Kulkarni, V. P. 2002, Ap.J., 571, 665

Seaton, M. J. 1959, M.N.R.A.S., 119, 81

Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, Ap.J., 588, 65

Sheinis, A. I., Bolte, M., Epps, H. W., Kibrick, R. I., Miller, J. S., Radovan, M. V., Bigelow, B. C., & Sutin, B. M. 2002, P.A.S.P., 114, 851

Sheth, R. K. & Tormen, G. 2002, M.N.R.A.S., 329, 61

Shimasaku, K., Ouchi, M., Okamura, S., Kashikawa, N., Doi, M., Furusawa, H., Hamabe, M., Hayashino, T., Kawabata, K., Kimura, M., Kodaira, K., Komiyama, Y., Matsuda, Y., Miyazaki, M., Miyazaki, S., Nakata, F., Ohta, K., Ohyama, Y., Sekiguchi, M., Shioya, Y., Tamura, H., Taniguchi, Y., Yagi, M., Yamada, T., & Yasuda, N. 2003, Ap.J., 586, L111

Smith, G. P., Edge, A. C., Eke, V. R., Nichol, R. C., Smail, I., & Kneib, J. 2003, Ap.J., 590, L79

Sokasian, A., Abel, T., & Hernquist, L. 2003, M.N.R.A.S., 340, 473

Somerville, R. S. 2002, Ap.J., 572, L23

Songaila, A. 2001, Ap.J., 561, L153

Songaila, A. & Cowie, L. L. 2002, A.J., 123, 2183

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2003, Ap.J.S., 148, 175

Springel, V., Yoshida, N., & White, S. D. M. 2001, New Astronomy, 6, 79

Stanway, E. R., Bunker, A. J., & McMahon, R. G. 2003, M.N.R.A.S., 342, 439

Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, Ap.J., 592, 728

Stengler-Larrea, E. A., Boksenberg, A., Steidel, C. C., Sargent, W. L. W., Bahcall, J. N., Bergeron, J., Hartig, G. F., Jannuzi, B. T., Kirhakos, S., Savage, B. D., Schneider, D. P., Turnshek, D. A., & Weymann, R. J. 1995, Ap.J., 444, 64

Stern, D. & Spinrad, H. 1999, P.A.S.P., 111, 1475

Storey, P. J. & Hummer, D. G. 1995, M.N.R.A.S., 272, 41

Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J., & Hazard, C. 1994, Ap.J., 427, L13

Storrie-Lombardi, L. J. & Wolfe, A. M. 2000, Ap.J., 543, 552

Taniguchi, Y., Shioya, Y., Ajiki, M., Fujita, S. S., Nagao, T., & Murayama, T. 2003, Journal of Korean Astronomical Society, 36, 123

Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T., & Palla, F. 1997, Ap.J., 474, 1

Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, Ap.J., 565, 773

Theuns, T., Bernardi, M., Frieman, J., Hewett, P., Schaye, J., Sheth, R. K., & Subbarao, M. 2002, Ap.J., 574, L111

Totani, T., Yoshii, Y., Iwamuro, F., Maihara, T., & Motohara, K. 2001, Ap.J., 550, L137

van de Weygaert, R. & Bertschinger, E. 1996, M.N.R.A.S., 281, 84

van Dokkum, P. G. 2001, P.A.S.P., 113, 1420

Vanden Berk, D. E., Richards, G. T., Bauer, A., Strauss, M. A., Schneider, D. P., Heckman, T. M., York, D. G., Hall, P. B., Fan, X., Knapp, G. R., Anderson, S. F., Annis, J., Bahcall, N. A., Bernardi, M., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Castander, F. J., Connolly, A. J., Crocker, J. H., Csabai, I., Doi, M., Finkbeiner, D., Friedman, S., Frieman, J. A., Fukugita, M., Gunn, J. E., Hennessy, G. S., Ivezić, Ž., Kent, S., Kunszt, P. Z., Lamb, D. Q., Leger, R. F., Long, D. C., Loveday, J., Lupton, R. H., Meiksin, A., Merelli, A., Munn, J. A., Newberg, H. J., Newcomb, M., Nichol, R. C., Owen, R., Pier, J. R., Pope, A., Rockosi, C. M., Schlegel, D. J., Siegmund, W. A., Smee, S., Snir, Y., Stoughton, C., Stubbs, C., SubbaRao, M., Szalay, A. S., Szokoly, G. P., Tremonti, C., Uomoto, A., Waddell, P., Yanny, B., & Zheng, W. 2001, A.J., 122, 549

Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, Ap.J., 465, 487

West, M. J. 1993, M.N.R.A.S., 265, 755

Weymann, R. J., Jannuzi, B. T., Lu, L., Bahcall, J. N., Bergeron, J., Boksenberg, A., Hartig, G. F., Kirhakos, S., Sargent, W. L. W., Savage, B. D., Schneider, D. P., Turnshek, D. A., & Wolfe, A. M. 1998a, Ap.J., 506, 1

Weymann, R. J., Stern, D., Bunker, A., Spinrad, H., Chaffee, F. H., Thompson, R. I., & Storrie-Lombardi, L. J. 1998b, Ap.J., 505, L95

White, R. L., Becker, R. H., Fan, X., & Strauss, M. A. 2003, A.J., 126, 1

White, S. D. M. 1999, Ap&SS, 267, 355

Wright, E. L. & Johnson, B. D. 2001, Ap.J., submitted, astro-ph/0107205

Wyithe, J. S. B. & Loeb, A. 2003a, Ap.J., 586, 693

—. 2003b, Ap.J., 588, L69

Yan, H., Windhorst, R. A., & Cohen, S. H. 2003, Ap.J., 585, L93

Zeldovich, Y. B. 1970, A.&A., 5, 84

Zepf, S. E. & Ashman, K. M. 1993, M.N.R.A.S., 264, 611

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, Ap.J., 475, 469

Zinn, R. 1985, Ap.J., 293, 424

Zuo, L. & Phinney, E. S. 1993, Ap.J., 418, 28