

**TOWARDS CONSTRAINTS ON THE EPOCH OF
REIONIZATION:**

A PHENOMENOLOGICAL APPROACH

Matthew Malloy

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

2015

Supervisor of Dissertation

Adam Lidz

Professor, Physics and Astronomy

Graduate Group Chairperson

R.U. Stupid

Professor, Physics and Astronomy

Dissertation Committee:

Doc 1, Professor, Physics and Astronomy

Doc 2, Professor, Physics and Astronomy

Doc 3, Professor, Physics and Astronomy

Doc 4, Professor, Physics and Astronomy

Doc 5, Professor, Physics and Astronomy

Doc 6, Professor, Physics and Astronomy

**TOWARDS CONSTRAINTS ON THE EPOCH OF
REIONIZATION:
A PHENOMENOLOGICAL APPROACH**

COPYRIGHT ©

2015

Matthew Malloy

This work is licensed under the Creative Commons
Attribution-NonCommercial-ShareAlike 4.0 License.
To view a copy of this license, visit
<http://creativecommons.org/licenses/by-nc-sa/4.0/>

“Nobody exists on purpose, nobody belongs anywhere, everybody’s gonna die. Come watch TV.”

Acknowledgements

Funding for this project was provided by the great spaghetti monster. The great spaghetti monster proudly supports research through raining pasta on the hungry masses of graduate students.

ABSTRACT

TOWARDS CONSTRAINTS ON THE EPOCH OF REIONIZATION: A
PHENOMENOLOGICAL APPROACH

Matthew Malloy

Adam Lidz

Trust me . . . I did stuff . . . it was good. 350 words maximum here. . .

Contents

Title	i
Copyright	ii
Dedication	iii
Acknowledgements	iv
Abstract	v
Contents	vi
List of Tables	x
List of Figures	xi
Preface	xxvi
1 First Things First	1
1.1 Cosmic Context	1
1.2 The Shoulders of Giants	1
1.2.1 The Ly α Forest	2
1.2.1.1 Evolution of τ_{eff}	6

CONTENTS

1.2.1.2	Dark Pixel Covering Fraction	10
1.2.1.3	Damping Wing Redward of Ly α	12
1.2.1.4	IGM Temperature	16
1.2.2	The 21-cm Line	20
1.2.3	The Cosmic Microwave Background	20
1.2.4	Ly α Emitters	20
1.2.5	Luminosity Function Measurements	20
1.3	Where We Stand	20
2	How to Search for Islands of Neutral Hydrogen in the $z \sim 5.5$ IGM	25
2.1	Introduction	25
2.2	Viability of Transmission Through a Partially Neutral IGM	29
2.3	Simulations and Mock Spectra	32
2.4	Dark Gap Statistics	37
2.5	Stacking Toy Spectra	39
2.5.1	HI Damping Wing	39
2.5.2	Deuterium	41
2.6	Steps of Approach	44
2.7	Results	47
2.7.1	Detecting the Damping Wing	47
2.7.2	Deuterium Feature Results	51
2.7.3	Dark Gap Statistics	53
2.8	Forecasts	54
2.8.1	Deuterium	55
2.8.2	HI Damping Wing	56
2.9	Conclusion	57

CONTENTS

3 On Modelling and Measuring the Temperature of the $z \sim 5$ IGM	81
3.1 Introduction	81
3.2 Simulations	84
3.3 Reionization Histories	85
3.4 The Thermal State of the IGM	90
3.4.1 Modeling the Thermal State	93
3.4.2 Simulated Temperature Field	95
3.4.3 Variations around Fiducial Parameters	100
3.5 Measuring the Temperature of the $z \sim 5$ IGM	103
3.5.1 Hydrodynamic Simulations: Perfect Temperate-Density Relation Models	105
3.5.2 Degeneracy with the Mean Transmitted Flux	108
3.5.3 Wavelet Amplitude PDFs in Inhomogeneous Reionization Models . .	110
3.5.4 Forecasts	114
3.6 Conclusions	118
4 Identifying Ionized Regions in Noisy Redshifted 21-cm Observations	137
4.1 Introduction	137
4.2 Method	140
4.2.1 The 21 cm Signal	140
4.2.2 Semi-Numeric Simulations	141
4.2.3 Redshifted 21 cm Surveys and Thermal Noise	142
4.2.4 Foregrounds	144
4.3 Prospects for Imaging	145
4.3.1 The Wiener Filter	146
4.3.2 Application to a Simulated 21 cm Signal	149
4.4 Prospects for Identifying Ionized Regions	151
4.4.1 The Optimal Matched Filter	151

CONTENTS

4.4.2	Application to Isolated Spherical Ionized Regions with Noise	152
4.4.3	Application to a Simulated 21 cm Signal	155
4.4.4	Success of Detecting Ionized Regions	160
4.4.5	Range of Template Radius Considered	161
4.5	Variations on the Fiducial Model	163
4.5.1	Ionized Fraction	163
4.5.2	Timing of Reionization	165
4.5.3	Effects of Foreground Cleaning	166
4.5.4	128 Antenna Tile Configurations	167
4.6	Favorable Antenna Configurations for Bubble Detection	169
4.7	Comparisons to Previous Work	171
4.8	Conclusion	172
5	Conclusion	185
5.1	Wrapping up	185
A	Some Appendix	189
	Glossary	190
	References	191

List of Tables

List of Figures

1.1	Illustration of the basic physics behind the Ly α forest and how gas at different locations along the line of sight results in absorption lines at different wavelengths. (Image from http://www.astro.ucla.edu/)	3
1.2	Flux as a function of rest-frame wavelength for a quasar at $z = 0.158$ (top) and $z = 3.62$ (bottom). The denser IGM at higher z results in a dense “forest” of absorption lines blueward of the rest-frame Ly α line in the lower panel. (Image from http://www.astro.ucla.edu/)	4
1.3	Measurements of Γ_{HI} (top) and $\langle x_{\text{HI}} \rangle$ from (16) (bottom).	21
1.4	Current limits on $\langle x_{\text{HI}} \rangle$ derived from the dark-pixel covering fraction in (29). Lightly-shaded points are older limits obtained in (30).	22
1.5	Quasar ULAS J1120+0641 identified at redshift $z = 7.085$ along with several fits for the damping wing.	22
1.6	Spectrum of GRB140515A, a gamma-ray burst located at $z = 6.33$. The right-hand panel overlays damping wing models from a host absorber (blue), a pure IGM model with $\langle x_{\text{HI}} \rangle = 0.056$ (red), and a combination model (green). The authors argue that, while each curve provides an equally-good fit to the data, the sharp rise in transmission shown is inconsistent with a significantly-neutral IGM.	23

LIST OF FIGURES

- 1.7 This diagram represents the process of Doppler broadening. The HI atom is moving away with velocity v from incoming radiation with frequency ν . The observed frequency of the radiation in the atom's rest frame is $\nu(1 - \xi/c)$ where ξ is the component of the velocity parallel with the incident radiation. 23
- 1.8 This figure shows mock spectra, and corresponding simulated IGM properties, from (8) in the top four panels. The bottom panel shows the observed spectrum from SDSS J0818+1722, which (8) use in order to make temperature measurements inside the proximity zone. Dashed lines indicate regions where Voigt-profile fitting was performed and downward arrows indicate the detected centers of the Voigt profiles. 24
- 2.1 Example mock Ly α forest spectrum and corresponding neutral fraction. The top panel shows the Ly α transmission while the bottom panel is the neutral fraction along the line of sight, with ionized regions set to $x_{\text{HI}} \approx 0$ for illustration. The black curve in the top panel shows the transmission through the forest when absorption due to the hydrogen damping wing is neglected, while the red curve includes damping wing absorption. The comparison illustrates that damping wing absorption has a prominent impact, but it is also clear that the presence of the damping wing will be hard to discern by eye. The line of sight is extracted from a model with $\langle x_{\text{HI}} \rangle = 0.22$, but note that we have deliberately chosen a sightline with more neutral regions than typical. 69
- 2.2 Dark gap size distribution for the $\langle x_{\text{HI}} \rangle = 0.22$, $\langle F \rangle = 0.1$ model. The solid blue curve shows the total distribution of dark gaps from an ensemble of mock spectra, where the magenta (cyan) curve shows the same thing but for the dark gaps sourced by ionized (neutral) gas. Here, we have focused on dark gaps with $L > 0.75 \text{ Mpc}/h$. This clearly demonstrates that neutral hydrogen is the dominant source of *large* dark gaps in our mock spectra, provided there is an appreciable neutral fraction. 70

LIST OF FIGURES

LIST OF FIGURES

2.5	Presence of deuterium absorption revealed through stacking idealized Ly β spectra containing toy neutral regions. The red and black curves show the stacked Ly β transmission redward and blueward, respectively, of toy neutral regions of length $L = 5 \text{ Mpc}/h$ ($\approx 700\text{km/s}$) randomly inserted into many sightlines, with spectra generated assuming $\langle F_{\text{Ly}\alpha} \rangle = 0.1$. In each case, stacking begins at the underlying HI/HII boundary. We have also mimicked the effect of including foreground Ly α absorption by scaling the feature by the mean transmission in the foreground Ly α forest. This demonstrates that, at least in this idealized case, the presence of deuterium absorption can be easily seen out to $\sim 80\text{km/s}$ past the HI/HII boundaries.	73
2.6	Ly α stacking results for various neutral fractions. The top panel shows the mean (noiseless) stacked transmission outside of large absorption systems (solid) and small absorption systems (dashed) in the Ly α forest for neutral fractions $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (red), and 0 (magenta). The transmission here is estimated from a large ensemble of mock spectra to obtain a smooth estimate of the average transmission around saturated regions in each model. The bottom panel shows the statistical significance of the difference between the dashed and solid curves in the top panel assuming a sample of 20 spectra are used in the stacking process.	74
2.7	Ly α stacking results assuming $\langle F \rangle = 0.05$. The above panels are identical to those in Fig. 2.6 except that mock spectra have been generated assuming $\langle F \rangle = 0.05$	75
2.8	Results of Ly α stacking with HIRES-style spectra ($\langle F \rangle = 0.1$). The above panel is identical to the bottom panel in Fig. 2.6 except that the spectra have had the bin size and spectral resolution adjusted to match that of Keck-HIRES spectra. Additionally, we have added noise such that the spectra have a signal-to-noise value of 10 per pixel at the continuum.	76

LIST OF FIGURES

2.9 Deuterium Ly β stacking results for various neutral fractions. The top panel shows the mean ensemble-averaged noiseless stacked transmission moving blueward (solid) and redward (dashed) away from large absorption systems in the Ly β forest for neutral fractions $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and 0 (magenta). The bottom panel shows the excess blueward absorption in units of the standard deviation of the stacked redward transmission, assuming 20 spectra.	77
2.10 Results of Ly β stacking with HIRES-style spectra. The above panel is the same as in the bottom panel of Fig. 2.9, except that it is generated using HIRES-style spectra, with spectral resolution of FWHM = 6.7km/s and additive noise with signal to noise of 30 per 2.1 km/s pixel at the continuum.	78
2.11 Mock dark gap size distribution. This figure is identical to Fig. 2.3 except that it uses spectra with spectral resolution FWHM = 100km/s, bin size $\Delta v_{\text{bin}} = 50\text{km/s}$, and a signal-to-noise ratio of 10 at the continuum. This figure shows the expected histogram of dark gap sizes using 20 spectra with $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and 0 (magenta) at fixed $\langle F \rangle = 0.1$	78

LIST OF FIGURES

LIST OF FIGURES

- 3.1 Example reionization histories. The red triangles show the simulated volume-average ionization fraction in our semi-numeric High-z reionization model, the black squares are for the Mid-z reionization scenario, and the blue pentagons are for a low redshift (Low-z) reionization model. The black dashed line shows the reionization history computed by solving Eq. 3.1 with $\zeta = 46$, $M_{\min} = 10^9 M_\odot$ and $C = 3$. The semi-numeric efficiency parameters $\tilde{\zeta}(z)$ in the Mid-z case have been tuned to match this model.

3.2 Thermal state of gas elements with a given reionization redshift, as a function of that redshift. In each case, the gas elements are heated to a temperature of $T_r = 2 \times 10^4$ K during reionization, and the residual photo-heating after reionization is computed assuming that the (hardened) spectral index of the ionizing sources is $\alpha = 1.5$ near the HI photoionization edge. *Top panel:* The temperature at mean density (T_0) for gas elements at each of $z = 4.5, 5.0$ and 5.5 as a function of their reionization redshift. *Bottom panel:* This is similar to the top panel, except it shows the slope of the temperature-density relation ($\gamma - 1$) rather than T_0 . Note that although we assume that gas elements with a given reionization redshift all land on a well defined temperature-density relation, this will not generally be a good description once we account for the spread in reionization redshift across the universe.

LIST OF FIGURES

3.3 Reionization redshifts and temperatures at $z = 5.5$ in the low-z reionization model. <i>Top panel:</i> The reionization redshifts for a narrow slice ($0.25 \text{ Mpc}/h$ thick) through the simulation. Each slice is $130 \text{ Mpc}/h$ on a side. The red regions indicate locations with the highest reionization redshifts across the simulation slice, while the dark regions are the last to be reionized. <i>Bottom panel:</i> The temperature of the same slice as in the top panel. The red areas in this panel show the hottest locations in the slice, and correspond to the dark regions in the top panel that are reionized late. The dark blue regions in the temperature slice, on the other hand, are the coolest regions that reionized first. The color scales are chosen so that 99% of simulation cells in the slice shown here have redshifts and temperatures falling between the minimum and maximum values on the color bar.	124
3.4 Reionization redshifts and temperatures at $z = 5.5$ in the high-z reionization model. Identical to Fig. 3.3, except this figure shows the contrasting High-z model. Note that the color scale in this case also encompasses 99% of the reionization redshifts and temperatures in the simulation slice, but that these ranges are different than in the previous figure.	125
3.5 Temperature density relations at $z = 4.5$ and $z = 5.5$ in the Low-z reionization model. The blue points show the temperature and density of gas elements from the simulation at $z = 5.5$, while the black points are the same at $z = 4.5$. The red short dashed line shows the median simulated temperature as a function of density at $z = 5.5$. The green long dashed line is the same at $z = 4.5$	126
3.6 Temperature density relations at $z = 4.5$ and $z = 5.5$ in the High-z reionization model. Identical to Fig. 3.5, except the results here are for the High-z reionization model.	127

LIST OF FIGURES

- 3.7 Power spectrum of temperature fluctuations in various models. The curves show the power spectrum of $\delta_{T_0}(x) = (T_0(x) - \langle T_0 \rangle)/\langle T_0 \rangle$ from the simulated models. The blue dotted line, the black solid line, and the red short-dashed line are the power spectra at $z = 5.5$ in the Low-z, Mid-z, and High-z models respectively. The black long-dashed line shows the δ_{T_0} power spectrum at $z = 4.5$ in the Mid-z model to illustrate how the temperature fluctuations fade with time. 128
- 3.8 Thermal state at $z = 5.5$ for various reionization temperature and spectral shape models. This is similar to the $z = 5.5$ curves in Fig. 3.2, except here we vary the reionization temperature, T_r , and the spectral shape, α . Increasing T_r leads to a higher T_0 and a flatter γ for recently reionized gas parcels, while parcels that reionize at sufficiently high redshifts are insensitive to T_r . A harder ionizing spectrum after reionization (smaller α) leads mostly to a slightly larger value of the asymptotic temperature achieved at high z_r . The harder spectrum also slightly hastens the transition of γ to its asymptotic value. 129
- 3.9 Temperature density relation at $z = 5.5$ for various reionization temperatures in the High-z and Low-z models. The “X”s in the legend indicate the color of the points in the corresponding models, while the dashed lines in the same models have different colors to promote visibility. The models in the legend are listed from top to bottom: the highest points and line (indicating the median temperature at various densities) show the $T_r = 2 \times 10^4$ K, Low-z model; next is the $T_r = 1 \times 10^4$ K, Low-z model; then the $T_r = 3 \times 10^4$ K, High-z model; and finally the $T_r = 3 \times 10^4$ K, High-z model. 130

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

3.15 Heating/cooling rates at $z \sim 7$. <i>Left panel</i> : The (absolute value of) the rates for relevant processes in the IGM at $T = 10^4$ K as a function of density, assuming that hydrogen is highly ionized and that helium is mostly singly-ionized. <i>Right panel</i> : Similar to the left panel except the rates are shown as a function of temperature for gas at the comic mean density.	136
4.1 Fourier profile of the Wiener filter, $W(k)$. The filter is averaged over line-of-sight angle and the results are shown at $z_{\text{fid}} = 6.9$ for simulated models with $\langle x_i \rangle = 0.51$ (blue dotted), $\langle x_i \rangle = 0.68$ (cyan dot-dashed), $\langle x_i \rangle = 0.79$ (green dashed), and $\langle x_i \rangle = 0.89$ (red solid).	148
4.2 Application of the Wiener filter to simulated data. The results are for our fiducial model with $\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$. <i>Top-Left</i> : Spatial slice of the unfiltered and noise-less 21 cm brightness temperature contrast field (normalized by T_0). <i>Top-Right</i> : Simulated signal-to-noise field after applying the Wiener filter to a pure noise field. <i>Bottom-Left</i> : Simulated signal-to-noise field after applying the Wiener filter to the noisy signal. This can be compared with the uncorrupted input signal shown in the top-left panel and the noise realization in the top-right panel. <i>Bottom-Right</i> : Simulated signal-to-noise field after applying the Wiener filter to the noiseless signal. (The filtered noiseless signal shown here is normalized by the standard deviation of the noise to facilitate comparison with the other panels.) All panels show a square section of the MWA field of view transverse to the line of sight with sidelength $L = 1 h^{-1}$ Gpc. All slice thicknesses are $\sim 8 h^{-1}$ Mpc. Unless noted otherwise, the simulation slices in subsequent figures have these same dimensions.	174

LIST OF FIGURES

LIST OF FIGURES

4.6 Impact of foreground cleaning on the matched-filtered field. This is similar to Figure 4.3, except that the results here are for a matched filter with a template radius of $R_T = 35 h^{-1}$ Mpc	178
4.7 An example of a detected ionized region. <i>Top-left</i> : Signal-to-noise field after applying the matched filter to the noisy signal. The detected bubble is plotted on top of the corresponding region in the map. <i>Top-Right</i> : Zoomed-in view of the detected bubble in the matched-filtered map. <i>Bottom-Left</i> : Detected bubble superimposed on a zoomed-in view of the noise-less unfiltered 21 cm brightness temperature contrast map. <i>Bottom-Right</i> : A perpendicular zoomed-in view of the bubble depicted in the bottom-left panel. All matched-filtered maps use the template radius that minimizes the signal-to-noise at the center of the detected bubble. In the top-left case, the boxlength is $L = 1 h^{-1}$ Gpc, while in the zoomed-in slices it is $L \approx 500 h^{-1}$ Mpc.	179
4.8 An example of an ionized region that our algorithm detects as several neighboring bubbles. <i>Top-left</i> : Signal-to-noise field after applying the matched filter to the noisy signal. The main detected bubble is plotted on top of the corresponding region in the map. <i>Top-Right</i> : Zoomed-in view of the main detected bubble in the matched filtered map (solid curve) along with two other nearby detected bubbles (dashed curve). <i>Bottom-Left</i> : The detected bubble superimposed on the zoomed-in, noise-less, unfiltered 21 cm brightness temperature contrast map. Again, the additional nearby detected bubbles are shown (dashed curve). <i>Bottom-Right</i> : A perpendicular view of the bubble depicted in the bottom-left panel, with the nearby detected bubbles visible. All matched-filtered maps use the template radius that maximizes the signal to noise at the center of the main detected bubble. The box length in the top-left figure is $L = 1 h^{-1}$ Gpc, while in the zoomed-in panels, the box length is $L = 550 h^{-1}$ Mpc.	180

LIST OF FIGURES

- 4.9 A measure of the bubble detection success rate. The points (\times) show the volume-averaged ionized fraction of detected bubbles versus their detected radius. For comparison, the cyan shaded region shows the $1-\sigma$ spread in the ionized fraction of *randomly* placed bubbles of the same radii. The bubble depicted in Fig. 4.7 is marked with a large red square, while the three bubbles shown in Fig. 4.8 are marked with large green circles. 181
- 4.10 Size distributions of detected bubbles for varying (volume-averaged) ionization fractions. The histograms show the size distribution of (identified) ionized regions for simulation snapshots with volume-averaged ionized fractions of $\langle x_i \rangle = 0.51$ (top-left), 0.68 (top-right), 0.79 (bottom-left), and 0.89 (bottom-right). These figures demonstrate how the total number and size distribution of detected bubbles varies with ionized fraction. 182
- 4.11 Bubble detection with the MWA-128. This figure is similar to Figure 4.5, except it is for the MWA-128 configuration rather than for the MWA-500. . 183
- 4.12 Bubble detection with a LOFAR-style interferometer. This figure is similar to Figure 4.5, except it is for the LOFAR configuration rather than the MWA-500. Additionally, all boxes in this figure have a side length of $426 h^{-1}$ Mpc, corresponding to the field-of-view of the LOFAR-style interferometer at $z = 6.9$. 184

Preface

Perhaps you feel inclined to preface your work with poetic self-reflection about why you took up this project? Maybe a story that's kind of boring and ends with a joke that's not funny? Go for it!

Chapter 1

First Things First

1.1 Cosmic Context

Oh, you know, stars 'n' stuff.

1.2 The Shoulders of Giants

Before we begin, it is first worth appreciating the difficulty of what we are trying to do. Essentially, we care about measuring the properties of the intergalactic gas – not stars or galaxies – when the Universe was only \sim 1 billion years old, a seemingly impossible task. Fortunately, we are given the invaluable gift that light travels at a finite speed and, as such, if we look at distant objects, we see them as they were in the past. Therefore, if we look at the gas between galaxies \sim 13 billion light years away from us, we will see it as it was roughly 13 billion years ago, when the Universe was only 1 billion years old. This means that, in principle, this information of how the young IGM evolved is directly available to us. However, even taking this into account, the intergalactic gas we care about is not bright and it is located extremely^{extremely} far away, so how are we supposed to observe it? An exciting aspect of studying the Epoch of Reionization is that, when confronted with such

a seemingly impossible task, experts in the field have developed many different creative approaches toward constraining the properties of the young IGM in order to learn about the Epoch of Reionization. It is this impressive body of work that we aim to build upon. We discuss a selection of the existing and future methods for constraining the EoR in this chapter.

1.2.1 The Ly α Forest

Arguably the most powerful tool for constraining the high-redshift IGM to date has been the Ly α forest. The Ly α forest results, in part, from another invaluable gift to the field of cosmology: the redshifting of light with the expansion of the Universe. Basically, as the Universe expands, space itself expands and with it the wavelengths of photons travelling through it expand as well. If we know the expansion history of the Universe, and know the intrinsic color of a luminous object, then we can use its *observed* color to determine our distance to the object. Because of this relationship, distances to objects are often measured as a redshift, defined as the fractional increase in wavelength that a photon experiences when travelling from a given distance to us, denoted by z . In other “words”:

$$\lambda_{\text{observed}} = \lambda_{\text{emitted}}(1 + z_{\text{emitter}}). \quad (1.1)$$

The Ly α forest can be seen in the spectrum of an extremely bright background object, usually a quasar or a gamma-ray burst (GRB), after its light has been processed by the intervening gas. Since the intervening gas is primarily composed of hydrogen and since this hydrogen is generally in the ground state, any intervening neutral patches will absorb light from the background object at the Lyman-series wavelengths, with the strongest absorption occurring at the Ly α wavelength: $\lambda_\alpha \approx 1216\text{r A}$. If the Universe were not expanding, then all intervening neutral hydrogen would absorb light from the quasar at one wavelength: $\lambda_\alpha \approx 1216\text{r A}$, neglecting the other lines in the Lyman-series for the moment. However,

1.2 The Shoulders of Giants

due to the expansion of the Universe, photons emitted from the quasar *blueward* of the Ly α line will redshift as they travel towards us. If they encounter neutral hydrogen as they redshift through the Ly α line, then they will be absorbed and an absorption line will be seen in the spectrum of the background quasar at a wavelength *blueward* of Ly α . This process is sketched in Figure 1.1. Thus, the Ly α forest is the pattern of absorption lines seen blueward of the rest-frame Ly α line in quasar spectra due to intervening neutral gas.

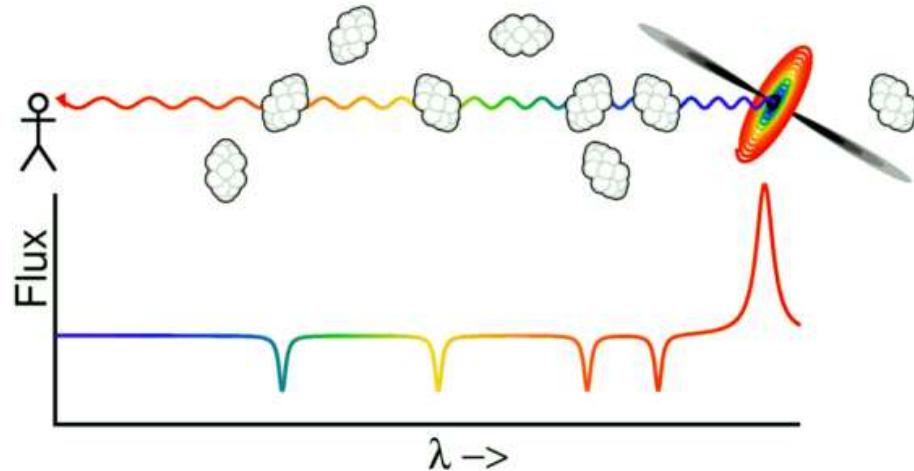


Figure 1.1: Illustration of the basic physics behind the Ly α forest and how gas at different locations along the line of sight results in absorption lines at different wavelengths. (Image from <http://www.astro.ucla.edu/>)

The same logical progression also applies to the other lines in the Lyman-series, with the difference that lines deeper in the series have a smaller cross section for absorption. Therefore, you could imagine observing a Ly β and Ly γ forest at smaller wavelengths. Another important difference here, though, is that photons emitted from a background source with energies larger than Ly β will redshift through the Ly β wavelength and *also* possibly through the Ly α wavelength before reaching us. The photon's physical location when it redshifts through those two wavelengths will be completely different and, therefore,

1.2 The Shoulders of Giants

when observing absorption lines in the Ly β forest, it can be difficult to tell if the photons were absorbed by distant gas undergoing a Ly β transmission or closer gas undergoing a Ly α transition. This problem is clearly exacerbated when considering higher-order lines since a larger number of distinct regions along the line of sight can contribute to the absorption.

We show two example quasar spectra in Figure 1.2. The spectrum in the top panel is for a quasar at relatively low redshift and shows very little absorption. Meanwhile, the quasar in the bottom panel shows little absorption for emitted wavelengths redward of Ly α but is heavily punctuated by absorption blueward of Ly α due to intervening neutral hydrogen. It is this pattern of absorption features with $\lambda_\beta \leq \lambda_{\text{emitted}} \leq \lambda_\alpha$ that is referred to as the Ly α forest.

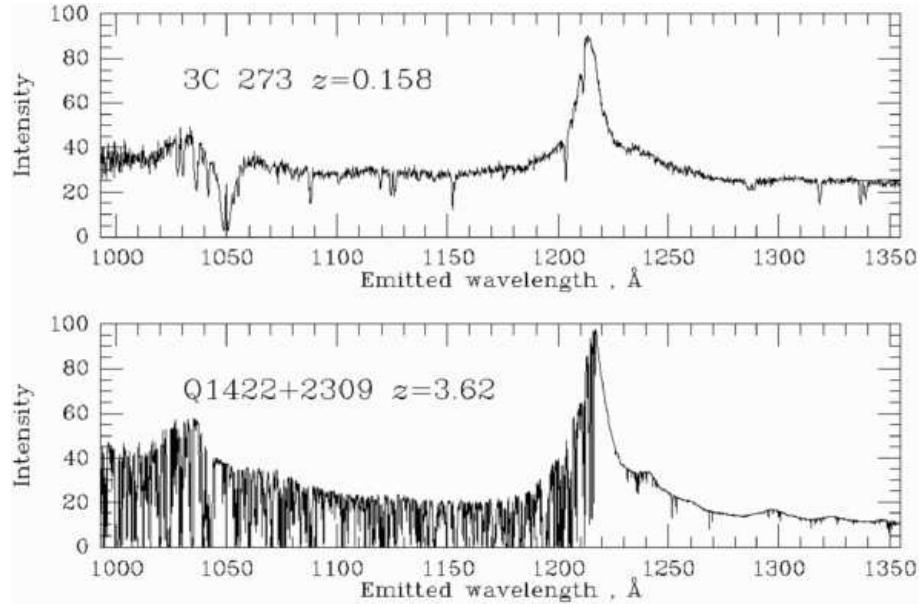


Figure 1.2: Flux as a function of rest-frame wavelength for a quasar at $z = 0.158$ (top) and $z = 3.62$ (bottom). The denser IGM at higher z results in a dense “forest” of absorption lines blueward of the rest-frame Ly α line in the lower panel. (Image from <http://www.astro.ucla.edu/>)

At this point, the Ly α forest should sound like a perfect tool: if we want to map

1.2 The Shoulders of Giants

the distribution of neutral hydrogen along the line of sight to a distant bright source, we can simply map each absorption line in the Ly α forest to a parcel of neutral hydrogen. However, the story becomes complicated here due to the extremely large cross section for Ly α absorption. Namely, the optical depth for Ly α absorption of a neutral hydrogen gas parcel obeys the relationship

$$\tau_\alpha \approx 3.3 \times 10^4 x_{\text{HI}}(1 + \delta) \left[\frac{1 + z}{6.5} \right]^{3/2}. \quad (1.2)$$

Using this expression, we can calculate the minimum neutral fraction needed for a gas parcel at mean density to allow 1% transmission at $z = 5.5$:

$$.01 = e^{-\tau_{\min}} \implies \tau_{\min} \approx 4.6 \quad (1.3)$$

$$4.6 \approx \tau_\alpha x_{\text{HI},\min} \quad (1.4)$$

$$\implies x_{\text{HI},\min} \approx 0.00014. \quad (1.5)$$

This reveals the fly in the ointment here: even a gas parcel that is 99.9% ionized will allow less than 1% transmission at the redshifts of interest for reionization. This allows absorption features to be caused by highly-ionized gas that happens to be over-dense. Therefore, we can not simply map absorption lines in the Ly α forest to regions of significantly-neutral hydrogen. In fact, the second example quasar we see in Figure 1.2 shows significant Ly α absorption and is located at $z = 3.62$, *much later than the end of reionization*. At this point, the reader may ask what utility does the Ly α forest have at all? Well, an enormous amount. To name a very few applications outside of reionization, since Ly α absorption can be caused by matter overdensities the Ly α forest can be used to measure the matter power spectrum in the IGM which in turn can be used to measure the baryon acoustic oscillations (BAO), provide lower limits on the mass of the dark matter (Viel et al. 50), and constrain dark energy. Additionally, absorption features due to damped Ly α absorbers

1.2 The Shoulders of Giants

(DLAs) can be used to measure the primordial deuterium abundance as a test of big bang nucleosynthesis.

But how to constrain the EoR?

1.2.1.1 Evolution of τ_{eff}

Perhaps the most common analysis performed on high-redshift quasar spectra in the context of constraining the EoR is measurements of the effective Gunn-Peterson optical depth, defined as

$$\langle F \rangle \equiv e^{-\tau_{\text{eff}}} \quad (1.6)$$

where $\langle F \rangle$ is the averaged transmission fraction over a redshift bin. Under the assumption of ionization equilibrium, where the rate that neutral hydrogen atoms are ionized is equal to the rate that they recombine, the effective optical depth encodes important information about the state of the IGM. In order to see this, we can take a few steps to express the optical depth in terms of the properties of the IGM.¹ First, the Gunn-Peterson optical depth can be expressed as

$$\tau_{\text{GP}} = \frac{\pi e^2}{m_e c} f_\alpha \lambda_\alpha \frac{n_{\text{HI}}}{H(z)}, \quad (1.7)$$

where $H(z)$ is the Hubble parameter at redshift z . All of these quantities are known with the exception of the number density of neutral hydrogen atoms, n_{HI} . To find this, we first utilize the statement of ionization equilibrium:

¹The following discussion will borrow heavily from (20) and (18).

1.2 The Shoulders of Giants

$$\Gamma_{\text{HI}} n_{\text{HI}} = R(T) n_e n_{\text{HII}} \quad (1.8)$$

$$n_{\text{HI}} = \frac{R(T) n_e n_{\text{HII}}}{\Gamma_{\text{HI}}} \quad (1.9)$$

$$x_{\text{HI}} = \frac{R(T) n_e}{\Gamma_{\text{HI}}} \quad (1.10)$$

where Γ_{HI} is the photoionization rate due to the ionizing sources, n_{HI} is the number density of neutral hydrogen atoms, n_e is the number density of free electrons, n_{HII} is the number density of ionized hydrogen atoms (protons), and x_{HI} is the hydrogen neutral fraction. Clearly, the left-hand side of Eq. 1.8 represents the rate of photoionizations per volume and the right hand side represents the rate of hydrogen recombinations per volume. Under the assumption of ionization equilibrium, it is true that $n_{\text{HII}} \approx n_{\text{HI}} + n_{\text{HII}} = n_{\text{H}}$ and

$$\bar{n}_{\text{H}} = \frac{\rho_c(z) \Omega_b(z) X_{\text{H}}}{m_p} = \frac{3H^2(z)}{8\pi G} \frac{\Omega_b(z) X_{\text{H}}}{m_p} \quad (1.11)$$

$$= \frac{3H_0^2 \Omega_{b,0}}{8\pi G} \frac{X_{\text{H}}}{m_p} (1+z)^3 \quad (1.12)$$

$$n_{\text{H}} = (1+\delta) \bar{n}_{\text{H}}. \quad (1.13)$$

In this expression, ρ_c is the critical density for a flat Universe, Ω_b is the baryon density in units of the critical density, X_{H} is the fraction of baryonic mass in the form of hydrogen, m_p is the mass of the proton which is effectively equal to the mass of the hydrogen atom, and $\delta = (\rho - \bar{\rho})/\bar{\rho}$ is the local overdensity in units of the cosmic mean. A subscript of “0” denotes that these are present-day values and \bar{n}_{H} denotes the average of n_{H} . Thus, as we expect, this expression is essentially equal to the mass density of hydrogen atoms in the Universe divided by the mass per atom.¹ The expression for the electron number density

¹It may be interesting to note that this value corresponds to 4 hydrogen atoms per cubic meter today and roughly ~ 1100 hydrogen atoms per cubic meter at $z = 5.5$.

1.2 The Shoulders of Giants

should be the same, since each ionized hydrogen atom releases one free electron. However, it is probably the case that helium is singly ionized along with hydrogen, so the number density will increase according to:

$$\bar{n}_e = \bar{n}_{\text{H}} + \bar{n}_{\text{He}} = \frac{3H^2}{8\pi G} \left(\frac{X_{\text{H}}}{m_p} + \frac{(1 - X_{\text{H}})}{4m_p} \right) \quad (1.14)$$

$$\approx 1.08\bar{n}_{\text{H}}. \quad (1.15)$$

However, for simplicity, let us approximate $n_{\text{tot}} \equiv n_e \approx n_{\text{H}}$. The quantity $R(T)$ in Eq. 1.8 is the recombination rate, which is equal to

$$R(T) \approx 4.2 \times 10^{-13} \left(\frac{T}{10^4 K} \right)^{-0.7} \text{cm}^3 \text{sec}^{-1}. \quad (1.16)$$

For $\delta \lesssim 5$, (24) showed that the temperature and density follow the relationship

$$T \approx T_0(1 + \delta)^{\gamma-1} \quad (1.17)$$

where $\delta \equiv (\rho - \bar{\rho})/\bar{\rho}$ is the baryon overdensity in units of the cosmic mean density, T_0 is the temperature of a parcel of gas at mean density, and γ is the slope of the temperature-density relation. For compactness, let's define $R_4 \equiv R(T = 10^4 K)$. At this point, we are ready to combine Eq. 1.17, 1.16, 1.13, 1.9, and 1.7 to get an expression for τ_{GP} :

$$\tau_{\text{GP}} = \frac{\pi e^2 f_{\alpha} \lambda_{\alpha}}{m_e c H(z)} \frac{R_4 (1 + \delta)^{-0.7(\gamma-1)} \bar{n}_{\text{tot}}^2(z) (1 + \delta)^2}{\Gamma_{\text{HI}}} \quad (1.18)$$

$$= \frac{\pi e^2 f_{\alpha} \lambda_{\alpha}}{m_e c H(z)} \frac{R_4 \bar{n}_{\text{tot}}^2(z)}{\Gamma_{\text{HI}}} (1 + \delta)^{2-0.7(\gamma-1)}. \quad (1.19)$$

Finally, we have an expression for the Ly α optical depth in terms of several properties of the IGM. The primary unknown in the above expression is the photoionization rate, which

1.2 The Shoulders of Giants

is a very complicated parameter which depends on the number, intensity, spectrum, and proximity of ionizing sources among other things. A common assumption in these types of analyses is that the photoionization rate is approximately spatially uniform. This seems like a plausible assumption in a post-reionization Universe, when any given gas parcel can receive ionizing radiation from hundreds of sources. On the other hand, early in reionization, the Universe is opaque to ionizing radiation and gas parcels tend to only see sources within their own ionized regions. Regardless, with this approximation, the observed mean transmission in a region of the spectrum is akin to an average of Eq. 1.19 over the density field:

$$\langle F \rangle = \int d\delta e^{-\tau(\delta)} P(\delta) \equiv e^{-\tau_{\text{eff}}}. \quad (1.20)$$

As such, an intriguing question is, if you have a model for the probability distribution of the underlying density field, what (uniform) value of Γ_{HI} will yield a value for $\langle F \rangle$ that is consistent with observations? This question has received a lot of attention (**cites**) and has led to many measurements of the photoionization rate. With estimates of the photoionization rate in hand, we can utilize Eq. 1.10 in order to obtain measurements of the IGM neutral fraction in each redshift bin in order to gauge the progress of the EoR. Results for measurements of Γ_{HI} and $\langle x_{\text{HI}} \rangle$ via this method, performed by (16), are shown in Figure 1.3. This figure demonstrates that, using τ_{eff} and the assumption of ionization equilibrium, estimates of the neutral fraction are exceedingly small for $z \lesssim 6$. This argument has played a large part in forming the common knowledge that reionization has ended by $z = 6$.

Despite the widespread analysis of τ_{eff} in constraining the end of reionization, the interpretation of τ_{eff} is quite complicated.¹ First, accepting the assumption of the IGM being in ionization equilibrium with a uniform Γ_{HI} is tantamount to assuming that reionization has ended. Specifically, as discussed in §1.1, reionization is likely a highly inhomogeneous

¹For a more thorough discussion of controversial aspects of these constraints, see the intro for (30).

process with ionized bubbles forming around the brightest sources, growing, and eventually overlapping. During the period prior to complete overlap, regions of neutral hydrogen will be shielded from the ionizing radiation while ionized bubbles will experience a very large Γ_{HI} . This is not reflected in Eq. 1.8 and so we expect conclusions derived from this method to be unreliable when we begin to push up against the end of reionization. Additionally, interpretation of the level of flux in quasar spectra in the context of reionization is further complicated by the fact that quasars likely reside in very atypical parts of the Universe. Specifically, (28) show that it is likely that many quasar lines of sight will pass through mostly ionized gas even up to neutral fractions of $\langle x_{\text{HI}} \rangle \sim 0.5$.

1.2.1.2 Dark Pixel Covering Fraction

As demonstrated in the previous section, interpreting measurements of the effective optical depth in the context of reionization is complicated and often relies on suspicious . However, an alternative approach is to consider what constraints can be made without resorting to model-dependent assumptions. In this regard, an important method in estimating $\langle x_{\text{HI}} \rangle$ from high-redshift quasar observations is the dark pixel covering fraction. This approach is rooted in the fact that neutral parcels of gas are certain to result in saturated absorption in quasar spectra due to their optical depths being $\tau_{\text{HI}} \gtrsim 10^4$ (Eq. 1.7). Therefore, a reliable upper bound on the neutral fraction at a given redshift can be estimated by the fraction of pixels in quasar spectra that are completely absorbed at that redshift.

An obvious drawback of this method is that, at $z \sim 6$, overdense yet ionized regions can very well result in saturated absorption as well and significantly increase the upper bound on the neutral fraction. One approach to combat this effect is to incorporate the Ly β forest into the analysis. The optical depth for Lyman-series transitions scales as $f\lambda$, where f is the oscillator strength of the transition and λ is the corresponding wavelength. Therefore, the analogous expression of Eq. 1.7 for Ly β is:

1.2 The Shoulders of Giants

$$\tau_\beta = \tau_\alpha \times \frac{f_\beta \lambda_\beta}{f_\alpha \lambda_\alpha} \approx 5.3 \times 10^3 x_{\text{HI}}(1 + \delta) \left[\frac{1+z}{6.5} \right]^{3/2}. \quad (1.21)$$

where $f_\alpha = 0.4162$, $\lambda_\alpha = 1216rA$, $f_\beta = 0.0791$, and $\lambda_\beta = 1026rA$. From this expression, we can see that a mean-density parcel of neutral gas should cause saturated absorption in both the Ly α and Ly β transitions. Meanwhile, ionized overdense regions are less likely to cause saturated absorption as their optical depth in Ly β is reduced by a factor of $f_\beta \lambda_\beta / f_\alpha \lambda_\alpha \approx 1/6$. Therefore, limits from the dark-pixel covering fraction may be improved by requiring simultaneous absorption in both Ly α and Ly β as part of the definition of a dark pixel. Additionally, the Ly β dark pixel covering fraction on its own is a viable tool for establishing an upper bound on the neutral fraction, although foreground Ly α absorption may undo some of the gains from the lower τ_β value.

This procedure faces several complications when actually carried out, however. First, quasar observations are subject to the noise from the night sky which effectively adds mean-zero random noise to the spectra. The addition of this random noise can result in spurious transmission in pixels that otherwise would have been completely absorbed. Therefore, to measure the dark pixel fraction in quasar spectra, one first needs to create a suitable definition of what classifies as a “dark” pixel. One approach here is to define dark pixels as having transmission below some threshold defined in terms of the noisy standard deviation, σ_N . This presents us with a tradeoff, however, since larger thresholds will reduce the number of neutral pixels we miss but also increase the number of ionized pixels that get incorporated into the dark pixel population.

A second complication is that, since pixels have a finite width, their transmission values effectively represent an average of the transmission over some region in the spectra. If the pixel width is large enough, then it is possible for a pixel to have non-zero transmission despite corresponding to a physical region that contains significantly-neutral gas. For example, if the physical region in space associated with the pixel is 80% composed of

completely-neutral gas and 20% composed of completely ionized gas which allows full transmission, then the transmission of that pixel will be 20% and will likely not qualify as a “dark pixel” despite containing neutral gas. Thus, even in a measurement as seemingly-simple as the dark-pixel covering fraction, these details must be kept in mind when interpreting results.

Regardless, (30) and (29) apply the dark-pixel covering fraction approach to 22 high-redshift quasar spectra to produce the constraints on $\langle x_{\text{HI}} \rangle$ shown in Figure 1.4. Dimly-colored points correspond to (30) while bold-colored points correspond to (29). The main conclusion here is that model-independent constraints have a hard time conclusively confirming the common knowledge that reionization has ended by $z \sim 6$.¹

1.2.1.3 Damping Wing Redward of Ly α

Much of the difficulties in using the Ly α forest to constrain the timing of the EoR can be boiled down to the following problem: interpreting Ly α absorption in high-redshift quasar spectra is difficult because both neutral and ionized gas can result in saturated absorption. Therefore, it is worth asking if there are any ways to break this degeneracy in Ly α absorption in order to determine which absorption is likely due to neutral hydrogen. One potential approach toward this goal, which has received much attention ((11), (10), (39), (9)), is looking for the hydrogen damping wing redward of the Ly α line.

To understand this approach, let us first understand what the hydrogen damping wing is. For many applications, it is suitable to consider an atom’s ability to absorb radiation as a series of delta functions in frequency: when incident radiation has a frequency exactly coinciding with the energy of the transition, then there is a non-zero probability for absorption and zero probability otherwise. In reality, the probability of absorbing a photon of

¹The authors here do go on to focus on a high-resolution subset of their quasar observations in order to argue in favor of an end to the EoR by $z \sim 6$. However, the precise interpretation of the spectra is complex, as discussed herein.

1.2 The Shoulders of Giants

a given frequency, e.g., the line profile, is a continuous distribution which, while small for frequencies $\nu \neq \nu_0$, is non-zero.

The intrinsic line profile for the Ly α transition in the hydrogen atom can be seen as arising from the time/energy uncertainty principle, $\Delta E \cdot \Delta t \gtrsim \hbar$. Specifically, the finite lifetime of the $n = 2$ excited state implies the existence of a range of energies that can excite, or result from, the transition. The distribution of this range of energies follows a Lorentzian distribution :

$$\phi(\nu) = \frac{1}{\pi} \frac{\Gamma/4\pi^2}{(\nu - \nu_0)^2 + (\Gamma/4\pi)^2} \quad (1.22)$$

with the corresponding absorption cross section

$$\sigma_\alpha(\nu) = \frac{\pi e^2}{m_e c} f_\alpha \phi(\nu). \quad (1.23)$$

The “damping wing” refers to the $\sigma \sim 1/(\nu - \nu_0)^2$ behavior far from line center. This can be used to break the degeneracy between HII absorption and HI absorption because the optical depth is so much smaller in the damping wing that, without significantly-neutral gas (optical depth scales with neutral fraction), the optical depth at such frequency separations will not be sufficient to cause absorption. Furthermore, the damping wing has a distinct shape which can be fit for in order to infer the properties of the neutral gas which sources it.

While the damping wing from an isolated neutral region in a sea of fully-ionized, $\tau = 0$ hydrogen would stand out like a sore thumb, in reality absorption from the surrounding dense, yet ionized, gas will punctuate the damping wing with additional absorption features and will make it harder to detect. This makes the prospect of looking for isolated damping wings in typical regions in quasar spectra unappealing. However, photons emitted slightly *redward* of Ly α cannot be absorbed by dense ionized gas since ionized gas has a negligible

1.2 The Shoulders of Giants

optical depth for $\nu \neq \nu_\alpha$. Neutral hydrogen, on the other hand, *will* allow absorption to take place redward of Ly α due to the significant optical depth in the damping wing. Because of this, searches for the damping wing slightly redward of the Ly α line will be able to avoid nuisance absorption from neighboring ionized gas.

We show the most well-known example of a potential damping-wing detection in Figure 1.5, taken from (39). This shows a region of the transmission spectrum for a quasar at redshift $z = 7.085$ (ULAS J1120+0641). The fractional transmission nearby the Ly α line exhibits a gradual recovery from almost complete absorption at $\lambda < \lambda_\alpha$ to almost complete transmission at $\lambda > \lambda_\alpha$, occurring over a wavelength interval consistent with a hydrogen damping wing. The curves in blue show models for damping wing absorption associated with an IGM with neutral fraction $\langle x_{\text{HI}} \rangle = 0.1$ (top), 0.5 (middle), and 1 (bottom) with a sharp ionization front at a distance of 2.2Mpc from the quasar. In green, a model for the absorption profile of a Damped Ly α Absorber (DLA) with column density $N_{\text{HI}} = 4 \times 10^{20} \text{ cm}^{-2}$ located 2.6 Mpc from the quasar is shown. Thus, the transmission profile appears consistent with both a significantly-neutral ($\langle x_{\text{HI}} \rangle > 0.1$) IGM or a proximate DLA. However, (47) perform a search for metal lines, which typically accompany DLA absorption, and find that the gas is extremely metal-poor. This bolsters the claim that the damping-wing absorption seen in this example is, in fact, due to diffuse neutral hydrogen in the IGM.

Other searches for damping-wing absorption redward of Ly α have been carried out on, for example, GRB 130606A (Chornock et al. 11) and GRB 140515A (Chornock et al. 10). These authors looked for the damping wing in the spectra of GRB afterglows at redshift $z = 5.913$ and $z = 6.33$, respectively. A non-detection in the spectra of the $z = 5.913$ GRB allowed the authors to place a 2σ limit on the nearby IGM neutral fraction of $\langle x_{\text{HI}} \rangle < 0.11$. Similarly, no strong evidence of a damping wing was found in the spectrum of GRB140515A, shown in Figure 1.6. The right-hand panel shows the transmission fraction nearby the Ly α transition, which is equally-well fit by pure host absorption (blue, $N_{\text{HI}} = 10^{18.62} \text{ cm}^{-2}$), pure IGM absorption from gas at $6.0 \leq z \leq 6.328$ with $\langle x_{\text{HI}} \rangle = 0.056$ (red), and a hybrid model

1.2 The Shoulders of Giants

with a host absorber lying within an ionized bubble with $R = 10$ comoving Mpc met by an IGM with $\langle x_{\text{HI}} \rangle = 0.12$ (green). As such, they argue against a significantly-neutral IGM at this redshift.

It is worth pointing out, however, that the method of searching for the damping wing redward of Ly α is not without drawbacks. First, detecting the damping wing redward of Ly α relies on your ability to understand what the quasar flux *would have* been in the absence of the absorbing gas nearby the Ly α line (this unabsorbed flux is referred to as the quasar *continuum* and predicting the unabsorbed flux for a given quasar is called *continuum fitting*). Predicting the Ly α line properties in quasars is notoriously complicated and so modelling the precise fractional transmission must be done with care. Second, searching for the damping wing redward of Ly α inherently involves measuring the gas properties nearby the quasar. However, quasars are extremely rare and special objects and it is not obvious that their surroundings are representative of the IGM on average. For example, (28) found that quasars are likely born into large galaxy-generated ionized regions, suggesting that interpreting the *lack* of a damping wing detection is not straightforward. Gamma-ray burst spectra are gaining attention in this regard (See Salvaterra 44 for a review) as they tend to occupy more typical regions of space and have an easier-to-model continuum flux. The drawbacks of GRBs, though, is that they are often accompanied by a host absorber whose damping-wing absorption must be separated from that of the IGM. Third, even when provided with a clean detection of the damping wing redward of Ly α , this will only tell you about one region of space and it will be difficult to use this single observation to extrapolate to the ionization state of the IGM as a whole. Later in this work, we propose a technique for searching for the hydrogen damping wing which, while faced with its own difficulties, is able to avoid the difficulties mentioned above.

1.2.1.4 IGM Temperature

A popular brainteaser exists where you are at the end of a hallway with three switches and out of view at the end of the hallway are three dark light bulbs with each bulb being controlled by one of the switches. You are allowed to flip the switches however you like and then walk to observe the light bulbs *only once*. From this single observation, you need to determine which light switch controls which bulb. This task is impossible except for the fact that a lit bulb will warm up. Therefore, you can flip two of the switches, wait, turn off one of the flipped switches, and then walk to the light bulbs where one will be lit, one will be dark and cold, and one will be dark and warm. From this information, you are able to determine which switch did what. In a contrived sense, this is not so different than with the EoR where utilizing temperature measurements can gain us additional insights. The process of reionization will heat freshly-ionized gas to extremely high temperatures. Subsequently, it will retain a thermal memory of the reionization process for some time. However, given enough time, the gas will cool to a point where the thermal memory is erased. In either case, it should be possible to measure the temperature of sufficiently high-redshift gas in order to put constraints on when it was ionized. In order to do this, we need two main ingredients: a method of estimating the temperature of high-redshift gas and an understanding of how the temperature of the gas evolves with time after being ionized.

One popular method for determining the temperature of the IGM utilizes the width of absorption lines in the Ly α forest. In §1.2.1.3, we described the line profile for Ly α absorption as obeying a Lorentzian distribution. While this is technically correct for any atom, in reality, the atoms themselves have random thermal motions according to their temperature and will therefore see incident radiation as being redshifted or blueshifted accordingly.¹ As such, a hydrogen atom travelling *toward* a photon with frequency just shy of the Ly α frequency will see the light redshifted and can increase the chance of absorption.

¹The following discussion of Doppler broadening and the derivation of the Voigt profile closely follows notes taken from Masao Sako’s “Radiative Transfer” class offered in the Spring of 2010.

1.2 The Shoulders of Giants

The effect of this is that the line profile for Ly α absorption gets smeared out, or, more precisely, gets convolved with Maxwell-Boltzmann distribution, which describes the distribution of velocities of particles in an ideal gas with a given temperature:

$$W(\xi) d\xi = \left(\frac{m_p}{2\pi k_B T} \right)^{1/2} e^{-m\xi^2/2k_B T} d\xi \quad (1.24)$$

$$= (\pi \xi_0^2)^{-1/2} e^{-\xi^2/\xi_0^2} d\xi \quad (1.25)$$

$$\xi_0 \equiv \sqrt{\frac{2k_B T}{m_p}}. \quad (1.26)$$

Incident radiation will appear red/blueshifted in the frame of the absorbing atom with the shift being proportional to the atom's velocity *parallel* to the incident radiation, as shown in Figure 1.7. Convolving our line profile with a Maxwell-Boltzmann distribution effectively involves replacing our expression in Eq. 1.23 with

factors of λ_α ??

$$\sigma(\nu) \rightarrow \frac{\pi e^2}{m_e c} f_\alpha \lambda_\alpha \int_{-\infty}^{\infty} d\xi \phi(\nu(1 - \xi/c)) (\pi \xi_0^2)^{-1/2} e^{-\xi^2/\xi_0^2} \quad (1.27)$$

$$= \frac{\pi e^2}{m_e c} \frac{f_\alpha}{\pi} (\pi \xi_0^2)^{-1/2} \int_{-\infty}^{\infty} d\xi \frac{(\Gamma/4\pi)e^{-\xi^2/\xi_0^2}}{(\nu - \nu_0(1 - \xi/c))^2 + (\Gamma/4\pi)^2}. \quad (1.28)$$

Here, we can make a couple substitutions and redefinitions:

$$\Delta v_D \equiv \nu_0 \frac{\xi_0}{c} \qquad \qquad \qquad y \equiv -\xi/\xi_0 \quad (1.29)$$

$$v \equiv \frac{\nu - \nu_0}{\Delta v_D} \qquad \qquad \qquad a \equiv \frac{\Gamma}{4\pi \Delta v_D}. \quad (1.30)$$

The quantity Δv_D is known as the “Doppler parameter” and is the red/blueshift in frequency space that the atom sees due to its thermal motion. The quantity v is just the distance from line center in velocity space in units of the Doppler parameter. The

1.2 The Shoulders of Giants

quantity $a \equiv (\Gamma/4\pi)/\Delta v_D$ represents the ratio of the natural line width to the Doppler width. Rewriting our expression in Eq. 1.28, we obtain

$$\sigma(\nu) = \frac{\pi e^2}{m_e c} f_\alpha \frac{1}{\sqrt{\pi} \Delta v_D} \frac{a}{\pi} \int_{-\infty}^{\infty} dy \frac{e^{-y^2}}{(v - y)^2 + a^2} \quad (1.31)$$

$$\equiv \frac{\sqrt{\pi} e^2}{m_e c} f_\alpha \frac{H(a, v)}{\Delta v_D} \quad (1.32)$$

where $H(a, v)$ is known as the Hjerting function or the Voigt function. So finally, we have obtained an expression for the Ly α line profile incorporating the natural line width and Doppler broadening. The profile is complex, but nonetheless, it contains information about the gas temperature and, given an observation of an absorption line, we should be able to perform fits to this profile in order to infer the temperature of the IGM.

In Figure 1.8, (8) look for Ly α absorption lines in the proximity zones of high-redshift quasars in order to fit for the associated Doppler parameter and make inferences about the temperature. The top panel shows the fractional transmission for a mock quasar spectrum. The dashed curves indicate regions where absorption-line fitting was carried out and the vertical arrows indicate where the line centers were found from the fits. The second and third panel show the underlying temperature and density field, respectively. The bottom panel shows the true spectrum in question along with the same information regarding the line profile fits. These authors were able to use these fits to make inferences about the temperature of the IGM nearby the quasar. From these temperature measurements, and through utilizing Eq. 1.33, these authors were able to place interesting limits on the ending of reionization ($z_H < 9$ assuming that the quasar reionizes its vicinity and Pop II stars drive reionization). This specific measurement is very difficult, however. The argument is essentially that the inferred temperature of the gas is too hot for reionization to have ended long before $z = 9$, otherwise the gas would have had more time to cool below the measured temperature. However, the regions nearby quasars should see a significantly-

1.2 The Shoulders of Giants

enhanced ionization field and all constraints made from measurements within this region hinge on ones ability to accurately account for such effects.

The aforementioned technique for measuring the temperature is useful, but has the drawback that it requires a certain level of transmission in the quasar spectra in order to confidently fit a single absorption-line profile. However, at $z > 5$, individual absorption features blend together and the forest becomes somewhat inverted where, instead of stretches of transmission being punctuated with absorption features, stretches of absorption are punctuated with transmission features. This renders the goal of fitting for individual Voigt profiles impossible in the IGM. This also explains why ?? analyze the nearby a high-redshift quasar, where the transmission is enhanced due to the strengthened ionization field.

we may want to look for a more detailed explanation of the $\frac{dT}{dt}$ expression.

This procedure has been applied to measure the temperature of the IGM at a variety of redshifts (**cites**). Alternatively, Lidz **cite** proposed using Morlet wavelet analysis as a measure of small-scale structure which is also indicative of the temperature. However, these analyses all depend on the Ly α forest having significant transmission. While this is true for $z < 5$, for $z > 5$ the picture is flipped long periods of absorption are punctuated by some transmission.

$$\frac{dT}{dt} = -2HT + \frac{2T}{3(1+\delta)} \frac{d\delta}{dt} + \frac{T}{\mu} \frac{d\mu}{dt} + \frac{2\mu m_p}{3\rho k_B} (\mathcal{H} - \Lambda) \quad (1.33)$$

One option here is to structure this section saying that people typically use something related to the Doppler widths in order to infer the temperature. However, at $z \gtrsim 5$, the forest becomes so absorbed that it inverted, in a sense: long periods of saturated absorption are punctuated with transmission spikes rather than the other way around. This makes Voigt-profile fitting impossible. This has kind of been a problem in measuring the temperature of the $z \gtrsim 5$ IGM, with some people using proximity zones of quasars. These

1.3 Where We Stand

are probably pretty special regions, though, hard to back out information on the IGM as a whole. However, Lidz is a boss and knows how to do it. (**this should be described be 0909.5210**)

1.2.2 The 21-cm Line

1.2.3 The Cosmic Microwave Background

1.2.4 Ly α Emitters

1.2.5 Luminosity Function Measurements

1.3 Where We Stand

1. Bouwens et al. (2015) might be good to include here. Strengthens argument that galaxies reionized the Universe and that quasars/AGN did not.

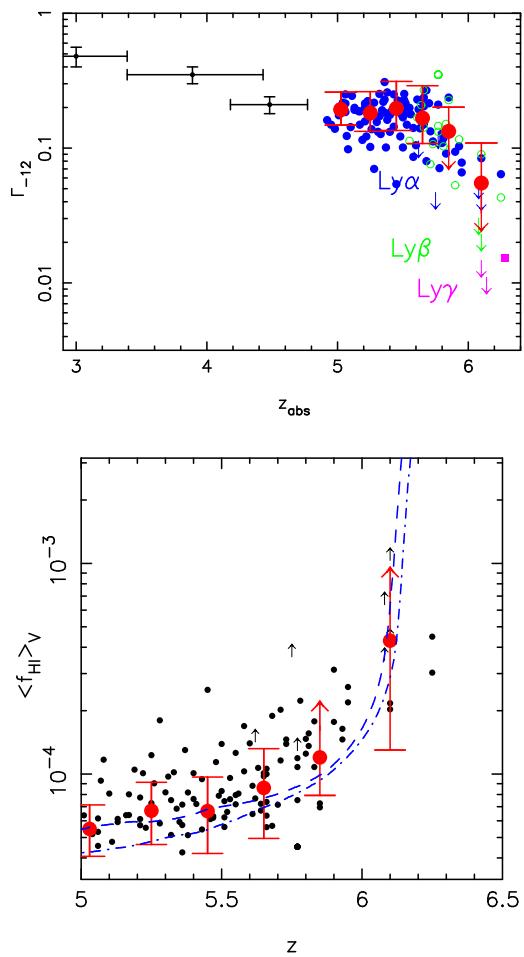


Figure 1.3: Measurements of Γ_{HI} (top) and $\langle x_{\text{HI}} \rangle$ from (16) (bottom).

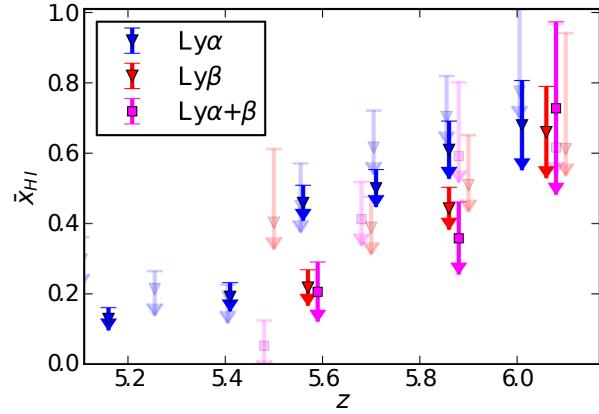


Figure 1.4: Current limits on $\langle \chi_{HI} \rangle$ derived from the dark-pixel covering fraction in (29). Lightly-shaded points are older limits obtained in (30).

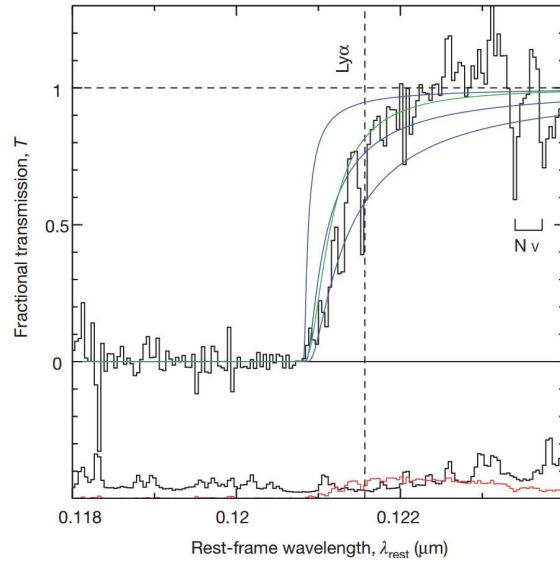


Figure 1.5: Quasar ULAS J1120+0641 identified at redshift $z = 7.085$ along with several fits for the damping wing.

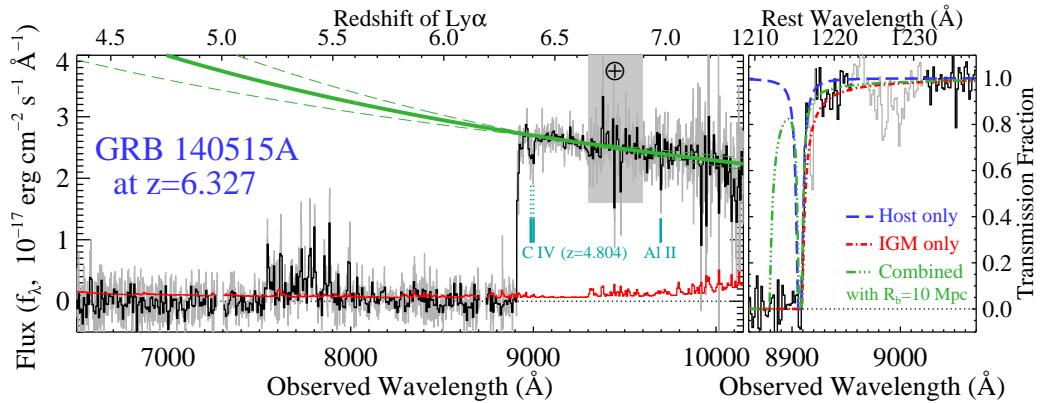


Figure 1.6: Spectrum of GRB140515A, a gamma-ray burst located at $z = 6.33$. The right-hand panel overlays damping wing models from a host absorber (blue), a pure IGM model with $\langle x_{\text{HI}} \rangle = 0.056$ (red), and a combination model (green). The authors argue that, while each curve provides an equally-good fit to the data, the sharp rise in transmission shown is inconsistent with a significantly-neutral IGM.

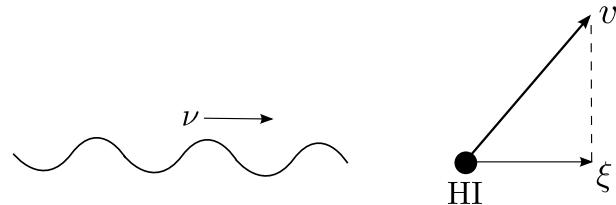


Figure 1.7: This diagram represents the process of Doppler broadening. The HI atom is moving away with velocity v from incoming radiation with frequency ν . The observed frequency of the radiation in the atom's rest frame is $\nu(1 - \xi/c)$ where ξ is the component of the velocity parallel with the incident radiation.

1.3 Where We Stand

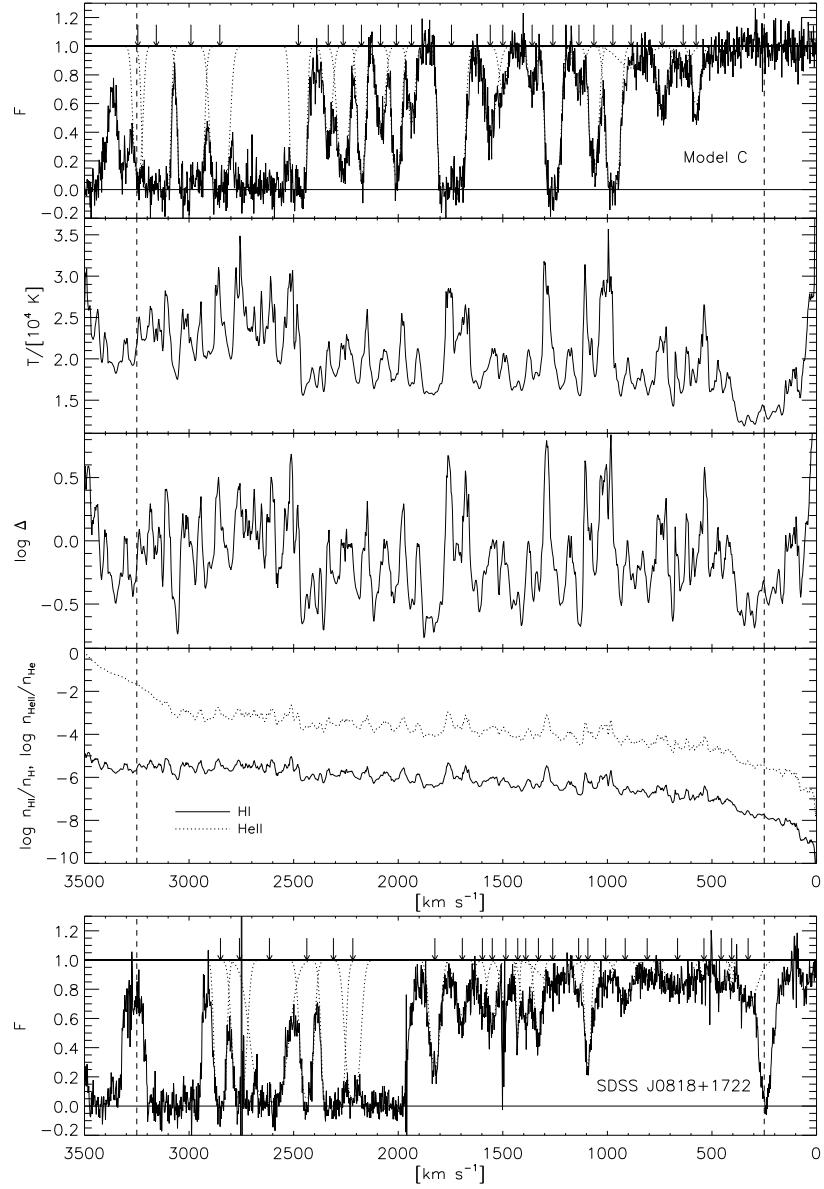


Figure 1.8: This figure shows mock spectra, and corresponding simulated IGM properties, from (8) in the top four panels. The bottom panel shows the observed spectrum from SDSS J0818+1722, which (8) use in order to make temperature measurements inside the proximity zone. Dashed lines indicate regions where Voigt-profile fitting was performed and downward arrows indicate the detected centers of the Voigt profiles.

Chapter 2

How to Search for Islands of Neutral Hydrogen in the $z \sim 5.5$ IGM

2.1 Introduction

It has been nearly half a century since Gunn and Peterson (23) pointed out that the lack of prominent absorption troughs, blueward of the Ly- α line in quasar spectra, implies that intergalactic hydrogen is highly ionized. Only in the year 2001 were complete “Gunn-Peterson” absorption troughs finally revealed in the Ly- α forest of high redshift ($z \gtrsim 6$) quasars discovered using the Sloan Digital Sky Survey (SDSS) (5, 15, 19). Although these prominent absorption troughs were discovered more than a decade ago, the precise interpretation of the observations, and their implications for the reionization history of the universe, remain unclear. One difficulty here relates to the large optical depth to Ly- α absorption: near $z \sim 6$, the optical depth is $\tau_\alpha \sim 4 \times 10^5$ in a fully neutral IGM at the cosmic mean density (23). Based on this, it is common to infer that the IGM must be highly ionized below $z \lesssim 6$, at which point quasar spectra do show some transmission through the Ly- α

2.1 Introduction

line. In addition, it is clearly hard to discern whether the gas above $z \gtrsim 6$ – that does show complete absorption in the Ly- α line – is mostly neutral or is only neutral at the level of about one part in ten-thousand or so (e.g. Fan et al. 17); in either case, the Ly- α line will be completely absorbed.

However, if reionization is sufficiently inhomogeneous and ends late, there may be some transmission through the Ly- α forest *before reionization completes* (28, 34). Theoretical models of reionization show that the IGM during reionization resembles a two-phase medium, containing a mixture of highly ionized bubbles along with mostly neutral regions. The ionized bubbles grow and merge, eventually filling essentially the entire volume of the IGM with ionized gas; the redshift at which this process completes is highly uncertain and still awaits definitive empirical constraint. In principle, the ionized bubbles may allow transmission through the Ly- α forest even when some of the IGM volume is still in fact filled by neutral regions, i.e., before reionization completes. This calls into question the conventional wisdom described above – that the presence of transmission through the $z \lesssim 6$ forest necessarily implies reionization completed by $z = 6$ (28, 34) – strictly speaking, this conclusion follows only in the unrealistic case of a homogeneously-ionized IGM.

Indeed, some portions of the $z \sim 5 - 6$ Ly- α forest are completely absorbed, while other portions of the forest at these redshifts show transmission through the Ly- α line. Quantitatively, if one counts only the fraction of pixels with some transmission through the forest as “certain to be ionized”, the volume-averaged neutral hydrogen fraction need only be smaller than $\langle x_{\text{HI}} \rangle < 0.2$ at $5 \leq z \leq 5.5$, and smaller than $\langle x_{\text{HI}} \rangle < 0.5$ at $z = 6$ (30). These constraints are conservative since even mostly-ionized gas will give rise to some completely absorbed regions at these redshifts, but it is nevertheless interesting to ask whether some of the absorbed regions could in fact come from remaining “islands” of mostly neutral hydrogen gas in the IGM. The dark pixel fraction constraints of McGreer et al. (30) certainly leave plenty of parameter space open for reionization completing at $z \leq 6$.

2.1 Introduction

In fact, there are hints – albeit indirect ones – that significant amounts of neutral gas may remain in the IGM at these late times and so we believe that investigating this possibility amounts to *more* than closing a remaining “loophole” in the analysis of the $z \lesssim 6$ Ly- α forest. For example, recent measurements of the rest-frame ultraviolet galaxy luminosity function suggest a relatively low ionizing emissivity at $z \gtrsim 5 - 6$, even for seemingly generous assumptions about the escape fraction of ionizing photons ($f_{\text{esc}} \sim 0.2$) and allowing significant extrapolations down the faint end of the luminosity function; e.g. the preferred model of Robertson et al. (43) (that matches these observations) has $\langle x_{\text{HI}} \rangle = 0.1$ at $z = 6$. In addition, the fraction of Lyman-break galaxies with detectable Ly- α emission lines shows evidence for a rapid drop between $z \sim 6 - 7$ which may require a significant neutral fraction at $z \sim 7$ (e.g., Pentericci et al. 40, Schenker et al. 45, although see Bolton and Haehnelt 7, Taylor and Lidz 48). The inferred $z \sim 7$ neutral fraction here would be easier to accommodate if there is still some neutral gas at $z \leq 6$. Furthermore, Becker et al. (4) recently discovered an impressive $\sim 110 \text{ Mpc}/h$ dark region in the $z \sim 5.7$ Ly- α forest. This may result from an upward opacity fluctuation – driven by a fluctuating ultraviolet radiation field in a mostly ionized IGM – but this striking observation invites contemplating the more radical possibility that diffuse neutral regions remain in the IGM at this late time. Finally, Mesinger and Haiman (35) and Schroeder et al. (46) argue that the proximity zones of quasars at $z \geq 6$ show evidence for damping wing absorption and a significant neutral fraction, further motivating the search for neutral gas at slightly later times.

Perhaps more importantly, we can design robust observational tests for the presence of neutral islands in the $z \sim 5.5$ IGM, and either definitively detect neutral hydrogen at these redshifts, or significantly improve on the existing upper limits from McGreer et al. (30). Towards this end, we study three possible tests for identifying neutral islands in the $z \sim 5 - 6$ IGM, each of which can be applied using existing Ly- α forest spectra. The presence of some transmission through the Ly- α forest at $z \leq 6$ allows us to consider tests that can

2.1 Introduction

not be applied at still higher redshift where the forest is completely absorbed (asides for in the “proximity zones” close to the quasar itself). We develop these tests using mock quasar spectra extracted from the numerical reionization simulations of McQuinn et al. (31). The first test we consider has been studied before (e.g. Fan et al. 17, Mesinger 33, McGreer et al. 30), but is the most model dependent: the abundance and size distribution of “dark gaps”, i.e., regions of saturated absorption in the Ly- α forest. Here we focus on the plausible impact of inhomogeneous reionization on the dark gap statistics. The second test utilizes the fact that the natural line width of the Ly- α line gives rise to extended damping wing absorption, in the case that highly neutral gas is present in the IGM (36). As a result, the transmission recovers more slowly around significantly neutral absorbed regions than around absorbed yet ionized regions. We find that this signature can be detected in partly neutral models by examining the stacked profile around extended absorbed regions. Note that, in contrast to previous work, here we propose to search for the damping wing signature in typical regions of the IGM, as opposed to in the proximity zones of quasars (Mesinger and Haiman 35, Schroeder et al. 46), or redward of Ly- α at the source redshift. Our third test involves the stacked profile of extended absorbed regions in the Ly- β forest. If these regions are significantly neutral, there should be a feature from absorption in the deuterium Ly- β line just blueward (but not redward) of absorbed regions.

The outline of this paper is as follows. In §2.2, we briefly discuss which range of (volume-averaged) neutral fractions are physically plausible at $z = 5.5$. In §2.3 we describe the simulations used and the process for generating mock spectra. We discuss how the dark gap size distribution may be used to constrain the neutral fraction in §2.4. In §2.5, we describe how quasar spectra may be stacked in order to reveal the presence of deuterium and HI damping wing absorption in an idealized scenario and discuss adapting this approach for more realistic spectra in §2.6. We then apply this approach to mock quasar spectra in §2.7, discuss the constraining power of the stacking approaches in §2.8, and conclude in §4.8. Throughout, we consider a Λ CDM cosmology parametrized by $n_s = 1$, $\sigma_8 = 0.8$,

2.2 Viability of Transmission Through a Partially Neutral IGM

$\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.046$, and $h = 0.7$, (all symbols have their usual meanings), broadly consistent with recent Planck constraints from (1).

2.2 Viability of Transmission Through a Partially Neutral IGM

Ideally, this study would make use of mock Ly- α forest spectra extracted from fully self-consistent simulations of reionization, in which the efficiency of the ionizing sources and other relevant parameters are tuned so that reionization completes at $z \leq 6$. Unfortunately, large-scale reionization simulations that simultaneously resolve the properties of the gas distribution, as well as the sources and sinks of ionizing photons, while capturing large enough volumes to include a representative sample of the ionized regions, are still quite challenging. Here, we instead explore more approximate, yet more flexible, models. As we describe in more detail in the next section, we make use of the reionization simulations of McQuinn et al. (31) to describe the size and spatial distribution of the ionized and neutral regions during reionization. Inside of the ionized regions, we rescale the simulated photoionization rates, adjusting the intensity of the UV radiation field to match the observed mean transmitted flux through the Ly- α forest. For simplicity, we assume that the intensity of the UV radiation field in the ionized regions is uniform and comment on the possible impact of this approximation where relevant.

Before proceeding further, however, it is worth considering which (volume-averaged) neutral fractions are physically plausible near $z \sim 5.5$. In order to get transmission through the $z \sim 5.5$ Ly- α forest, at least some of the hydrogen needs to be highly ionized. This requires the mean free path of ionizing photons to be relatively large, although we should keep in mind that the attenuation length will vary spatially during and after reionization, and so this quantity needs to be large only across some stretches of the IGM. This in turn demands some minimum separation between the neutral islands, because otherwise the

2.2 Viability of Transmission Through a Partially Neutral IGM

neutral islands themselves will limit the mean free path and prevent a sufficiently intense UV radiation field from building up between the islands. Hence, it may be inconsistent to have remaining neutral islands in the IGM, yet still have some transmission through the Ly- α forest. Here we briefly quantify this reasoning; we will be content with only a rough estimate, as our focus here is more on designing empirical tests. Further theoretical exploration here might be valuable, however, perhaps along the lines of (53).

Quantitatively, previous studies infer that a photoionization rate on the order of $\Gamma_{\text{HI}} \sim 5 \times 10^{-13} \text{s}^{-1}$ is required to match the mean transmitted flux in the $z \sim 5.5$ Ly- α forest (e.g., Kuhlen and Faucher-Giguère 25, Bolton and Haehnelt 6).¹ If we demand that the photoionization rate between the neutral islands needs to be in this ballpark to allow transmission through the forest, we can translate this into a required minimum average separation between the neutral islands, given an assumed ionizing emissivity. The average ionizing emissivity is likely on the order of $\epsilon_{\text{HI}} \sim 3$ photons per atom per Gyr (e.g. Bolton and Haehnelt 6). This is close to the value required simply to balance recombinations and maintain the ionization of the IGM at the redshifts of interest. This emissivity is also comfortable with that inferred from the above photoionization rate and measurements of the mean free path to ionizing photons (Bolton and Haehnelt 6, although Becker and Bolton 3 recently argued for a slightly larger value), as well as the UV emissivity implied by measurements of the galaxy luminosity function (e.g. Robertson et al. 43).

In this context, it is useful to note that:

$$\Gamma_{\text{HI}} = \varepsilon_{\text{HI}} \sigma_{\text{HI},\text{lim}} \lambda_{\text{mfp}} \frac{\beta}{\beta + 1.5}, \quad (2.1)$$

where ε_{HI} is the average proper ionizing emissivity, $\sigma_{\text{HI},\text{lim}} = 6.3 \times 10^{-18} \text{cm}^2$ is the photoionization cross section at the Lyman limit, λ_{mfp} is the mean free path of ionizing photons at the Lyman limit, and β is the intrinsic, unhardened spectral index of the ionizing radiation.

¹These studies assume that reionization is complete at these redshifts. If the universe is in fact partly neutral, then a higher photoionization rate should be required in the ionized regions. In our rough estimate here, we neglect this given the other significant uncertainties involved.

2.2 Viability of Transmission Through a Partially Neutral IGM

This expression assumes that the mean free path to ionizing photons propagating through a clumpy IGM scales as $\nu^{3/2}$ (54). Inserting typical numbers we find:

$$\Gamma_{\text{HI}} = 5.0 \times 10^{-13} \text{ sec}^{-1} \left[\frac{\varepsilon_{\text{HI}}}{3 \text{ photons/atom/Gyr}} \right] \quad (2.2)$$

$$\times \left[\frac{\beta}{2} \right] \left[\frac{3.5}{1.5 + \beta} \right] \left[\frac{1+z}{6.5} \right]^3 \left[\frac{\lambda_{\text{mfp}}}{9.1 \text{ pMpc}} \right]. \quad (2.3)$$

In other words, to get transmission through the forest for plausible values of the ionizing emissivity, we require the mean separation between neutral islands to be $\lambda_{\min} \gtrsim \lambda_{\text{mfp}} \gtrsim 9.1 \text{ pMpc}$. This is a minimal requirement in that it assumes the neutral islands set the mean free path, when in fact Lyman limit systems and cumulative absorption in the mostly ionized gas may also play a role. On the other hand, the required minimum separation between the neutral islands would go down if a smaller Γ_{HI} suffices to allow transmission through the forest, or if the ionizing emissivity is in fact higher. However, the mean free path to ionizing photons has recently been measured at $z = 5.16$ to be $\lambda_{\text{mfp}} = 10.3 \pm 1.6 \text{ pMpc}$ (52), only somewhat larger than our assumed λ_{mfp} here. While there are still uncertainties, and while the measured mean free path scales steeply with redshift ($\lambda_{\text{mfp}} \propto (1+z)^{5.5}$), viable models are unlikely to have neutral islands spaced much more closely than this.

We can then use this requirement on λ_{mfp} to get some sense of which volume-averaged neutral fractions are plausible at $z \sim 5.5$. In the simulation outputs considered here (see §2.3), the mean separation between neutral islands is $\lambda_{\text{mfp}} = 17.0 \text{ pMpc}$, 5.3 pMpc , and 2.7 pMpc for $\langle x_{\text{HI}} \rangle = 0.05$, 0.22 , and 0.35 , respectively. The first case certainly satisfies the requirement described above, the second case is just a bit on the small side, while the third case is uncomfortably small. Given the uncertainties in this argument, and the possibility that the neutral islands are a bit larger than in our simulation (which would increase their mean separation at fixed filling factor), we consider all three cases, but refrain from considering still more neutral models. We regard the latter case ($\langle x_{\text{HI}} \rangle = 0.35$) as an extreme scenario intended mostly for illustration.

2.3 Simulations and Mock Spectra

Finally, it is worth keeping in mind that any remaining neutral islands will likely be photoionized on a short timescale. For example, using Eq. 1 in Lidz and Malloy (27) with $C = 3$, $M_{\min} = 10^9 M_{\odot}$, and $\zeta = 20$, the redshift interval over which the volume average ionized fraction transitions from $\langle x_i \rangle = 0.8$ to $\langle x_i \rangle = 1$ is only $\Delta z \sim 0.5$. However, it is possible that we are catching this – likely brief – phase in $z \sim 5.5$ Ly- α forest spectra and the possibility of testing this remains tantalizing.

2.3 Simulations and Mock Spectra

With the above discussion to frame the range of possibilities, we move to describe the numerical simulations used in this analysis and our approach to constructing mock Ly- α forest absorption spectra before reionization completes. We use simulated density and ionization fields generated from a dark matter simulation of (31) which tracks 1024^3 dark matter particles in a simulation volume with a co-moving sidelength of $L = 130$ Mpc/ h . We assume that the gas closely follows the dark matter. In this work, we focus on redshift $z = 5.5$, but consider several possible neutral fractions. In practice, we obtain ionization fields with higher (lower) neutral fractions by using simulation outputs at higher (lower) redshifts. This should be an appropriate approximation since the statistical properties of ionized regions at a given neutral fraction are most sensitive to the neutral fraction and are relatively insensitive to the redshift at which the neutral fraction was attained (see McQuinn et al. 32 and Furlanetto et al. 21).

We generate mock quasar spectra according to the usual “fluctuating Gunn-Peterson” approach (e.g., Croft et al. 14), with a few refinements to capture the main effects of incomplete reionization. First and foremost, we do not assume a fully ionized IGM. The transmission in the Ly- α forest is sensitive to the precise ionized fractions in the ionized phase of the IGM. In order to simplify our study, as mentioned in the previous section, we rescale the simulated photoionization rates in the ionized regions to match the observed mean transmitted flux through the Ly- α forest. We do this assuming ionization equilibrium,

2.3 Simulations and Mock Spectra

and a constant value of the UV background (with a photoionization rate per atom of Γ_{HI}) within ionized regions. Specifically, simulated pixels with $x_i > 0.9$ are considered highly ionized while less ionized pixels are considered fully neutral.¹ This simplified approach allows us to consider a range of different possibilities for the ionization state of the IGM quickly. We comment on the shortcomings of this approach when appropriate. The optical depth of a given pixel, i , in the simulation can then be found by summing over contributions from neighboring pixels (Bolton & Haehnelt 2008):

$$\tau_\alpha(i) = \frac{c\sigma_\alpha\delta R}{\pi^{1/2}} \sum_j \frac{n_{\text{HI}}(j)}{b(j)} H(a, x), \quad (2.4)$$

where $b(j) = (2k_B T(j)/m_p)^{1/2}$ is the Doppler parameter, $T(j)$ is the temperature of pixel j , δR is the pixel proper width, $\sigma_\alpha = 4.48 \times 10^{-18} \text{ cm}^2$ is the Ly α scattering cross section, m_p is the proton mass, $H(a, x)$ is the Hjerting function, and $n_{\text{HI}}(i)$ is the number density of hydrogen atoms at pixel i , found using the simulated density field. To calculate the Doppler parameter, we assume that the gas obeys a modified temperature-density relationship

$$T(\delta) = \begin{cases} T_0(1 + \delta)^{\gamma-1} & \text{if ionized} \\ 1,000 \text{K} & \text{if neutral,} \end{cases} \quad (2.5)$$

where δ is the matter overdensity in units of the cosmic mean and we choose $T_0 = 2 \times 10^4 \text{ K}$ and $\gamma = 1.3$ as the temperature at mean density and slope of the temperature-density relation, respectively. For simplicity, we assume the ionized gas lies on the aforementioned temperature-density relation, although there should be significant scatter around this relation close to reionization (e.g. Lidz and Malloy 27). We do not expect this to impact our conclusions significantly. The neutral gas should be colder than the ionized gas, of course, with a temperature set perhaps by low levels of X-ray pre-heating. Here we adopt

¹After effectively thresholding the ionization field in this way, we end up with neutral fractions which are $\approx 20\%$ higher than in the original simulation. Throughout the paper, we refer to increased, thresholded neutral fractions.

2.3 Simulations and Mock Spectra

$T = 1,000$ K for the neutral gas; this choice is likely a bit large (it was chosen partly for ease in computing the Hjerting function below), but we have checked that we get nearly identical results for colder temperature choices.

The Hjerting function is a convolution of a Lorentzian profile, which incorporates the natural line profile of the Lyman-series lines, with a Maxwell-Boltzmann distribution, which accounts for the effects of thermal broadening on the line profile. The Hjerting function is defined by:

$$H(a, x) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2} dy}{a^2 + (x - y)^2},$$

where $a = \Lambda_\alpha \lambda_\alpha / 4\pi b(j)$, $\Lambda_\alpha = 6.265 \times 10^8 \text{ sec}^{-1}$ is the damping constant, $\lambda_\alpha = 1215.67 \text{ r A}$ is the Ly α wavelength, x is the relative velocity of pixel i and pixel j in units of the Doppler parameter, defined as $x = [v_H(i) - u(j)] / b(j)$, where $u(j) = v_H(j) + v_{\text{pec}}(j)$. The peculiar velocity field is generated by applying linear perturbation theory to the underlying density field.¹ In detail, the natural line profile is only approximately described by a Lorentzian (26), with asymmetric corrections becoming important far from line center. In this study, the precise shape of the damping wing far from line center is unimportant: we make use only of the gradual recovery in transmission around saturated neutral regions, rather than the detailed shape of this recovery, which is also strongly influenced by neighboring neutral regions. We hence expect the Lorentzian form to be a good approximation for our present purposes.

In addition to including absorption from the hydrogen damping wing, we also include absorption from primordial deuterium. As a result of big bang nucleosynthesis, primordial hydrogen should be accompanied by traces of deuterium, with a relative abundance by number of $D/H = 2.5 \times 10^{-5}$ (Cooke et al. 12). Due to its slightly increased reduced mass, Lyman series transitions in deuterium will be shifted blueward by 82km/s compared

¹This was done because the full peculiar velocity field was not readily available, but this approximation should not impact our results.

2.3 Simulations and Mock Spectra

to the same transitions in hydrogen. We account for deuterium by scaling the number density of hydrogen in a given pixel by the relative abundance and shifting the resulting optical depths blueward by 82km/s. Additionally, the Doppler parameter is adjusted to $b_D(j) = (2k_B T(j)/(2m_p))^{1/2}$ to account for the increase in mass.

In this work, we focus mostly on $z = 5.5$ and adopt a mean transmitted flux at this redshift of $\langle F \rangle = 0.1$, consistent with determinations from e.g., (5). In some cases, we test the sensitivity of our results to the mean transmitted flux by considering $\langle F \rangle = 0.05$ as well. In general, the lower the mean transmitted flux, the more challenging it is for us to identify any remaining neutral islands. On the other hand, the likelihood that neutral islands remain increases towards high redshift and decreasing mean transmitted flux. As mentioned previously, we rescale the simulated photoionization rates in the ionized regions to a uniform value, normalized so that an ensemble of mock spectra matches the observed mean transmitted flux. It is important to note that the mean transmitted flux is a very steep function of redshift near $z \sim 5.5$, and that the sightline-to-sightline scatter in this quantity is substantial (17), and so one may want to carefully test for sensitivity to the precise redshift binning used.

We use the same approach as described above to generate Ly β mock spectra, with $\lambda_\beta = 1025.72\text{r A}$, $\Lambda_\beta = 1.897 \times 10^8 \text{ sec}^{-1}$, and $\sigma_\beta = 7.18 \times 10^{-19} \text{ cm}^2$. However, in generating Ly β mock spectra, we must also account for foreground Ly α absorption due to gas at lower redshifts, $\lambda_\alpha(1 + z_{\text{Ly}\alpha}) = \lambda_\beta(1 + z_{\text{Ly}\beta})$, where $z_{\text{Ly}\alpha}$ is the redshift of the foreground Ly α absorber and $z_{\text{Ly}\beta}$ is the redshift of the Ly β absorber. We will assume we are investigating quasar spectra at $z_{\text{Ly}\beta} = 5.5$ for this work, such that the corresponding foreground Ly α absorption in the Ly β spectra occurs at redshift $z_{\text{Ly}\alpha} = 4.5$. We adjust Γ_{HI} for the foreground Ly α absorption to match measurements of the mean transmission from (2) at these redshifts ($\langle F \rangle \approx 0.31$ at $z = 4.5$).¹ The optical depth of a pixel in a Ly β spectrum

¹We generate foreground Ly α absorption by considering the absorption from regions in the same simulation box, but demand that they are widely-separated from the high redshift regions of interest ($> 10 \text{ Mpc}/h$). This enforces that the underlying density fields sourcing the Ly β absorption and the foreground Ly α ab-

2.3 Simulations and Mock Spectra

is then the sum of the contribution from the foreground Ly α absorption and the intrinsic Ly β absorption $\tau_{\beta}^{\text{tot}}(z_{\text{Ly}\beta}) = \tau_{\beta}(z_{\text{Ly}\beta}) + \tau_{\alpha}(z_{\text{Ly}\alpha})$.

In Fig. 2.1, we show an example mock Ly α spectrum for a particular line of sight through the simulation. We show only a portion of the line of sight in order to exhibit smaller-scale features. The top figure shows the Ly α transmission when the hydrogen damping wing is neglected (black) and when it is included (dashed red), while the bottom panel shows the underlying thresholded ionization field. We have neglected peculiar velocities in creating this figure in order to facilitate a comparison between the spectrum and the underlying ionization field.

From this figure, we see that the damping wing indeed has a significant effect on the transmission, but that its effect is hard to discern without knowing the damping-wing-less transmission. This is the case for two reasons. First, the forest here is very absorbed and the damping wing absorption becomes mixed with resonant absorption from neighboring ionized regions. Second, the damping wing from a particular neutral region may overlap with the damping wing from another neutral region, altering the shape of the resulting absorption. Specifically, we see that, in the example spectra, the region at $v \approx 4500$ km/s is sandwiched between HI regions to the left and right, both within 1000km/s. Therefore, the optical depths for the corresponding pixels likely have significant contributions from resonant absorption, damping wing absorption from the HI region to the left, and damping wing absorption from the HI region to the right. While detecting individual instances of damping wing absorption in this case seems impossible, we will show that detecting the presence of damping wing absorption *on average* should be feasible through the stacking of high-redshift quasar spectra.

sorption are uncorrelated, as should be the case for actual spectra.

2.4 Dark Gap Statistics

With the mock spectra of the previous section in hand, we now consider the size distribution of regions of saturated absorption – dark gaps – and its dependence on the underlying neutral fraction. Using such dark gap statistics has been widely discussed as a potential probe of the IGM neutral fraction (see, e.g., Fan et al. 17, Gallerani et al. 22, Mesinger 33, McGreer et al. 30). In a fully ionized IGM, the size of dark gaps in quasar spectra should grow with increasing redshift, simply owing to the increasing mean density of the universe and as a result of any decline in the intensity of the UV radiation background. However, once quasar spectra start to probe the tail end of reionization, the increase in dark gap size should accelerate due to the presence of islands of neutral hydrogen.

In Fig. 2.2, we have plotted the size distribution of dark gaps, $dP(L)/d\ln L$, in blue for the $\langle x_{\text{HI}} \rangle = 0.22$ mock spectra, assuming a mean transmission of $\langle F \rangle = 0.1$. Additionally, the dashed curves display the two underlying populations of dark gaps: those sourced by ionized gas (magenta) and those sourced by neutral gas (cyan). For clarity, we only show dark gaps larger than $L_{\text{sat,min}} = 0.7 \text{ Mpc}/h$ ($\sim 90 \text{ km/s}$), since smaller saturated regions will be predominantly ionized. Additionally, we have neglected peculiar velocities when generating spectra here. Two important points become apparent from this figure. First, at $L \sim 8.5 \text{ Mpc}/h$ ($\sim 1100 \text{ km/s}$), dark gaps transition from being primarily sourced by ionized gas to being primarily sourced by neutral gas. This reinforces our intuition that, in a partially neutral IGM, the largest dark gaps should correspond to the remaining neutral islands. Second, the dark gaps being composed of two different populations gives rise to a bimodality in the size distribution. This suggests that the behavior of the large- L tail of the size distribution may offer additional information about the neutral fraction, with a steep decline suggesting a highly ionized IGM and a more gradual decline, or the emergence of a second peak, suggesting a significantly neutral IGM. Such a “knee” in the dark gap size distribution is also mentioned in (33).

2.4 Dark Gap Statistics

Additionally, we can consider the large- L tail of the size distribution and its dependence on neutral fraction at a fixed mean transmission. In Fig. 2.3, we plot an expected histogram of dark gap sizes for 10 spectra for $\langle x_{\text{HI}} \rangle = 0$ (magenta), 0.05 (cyan), 0.22 (blue), and 0.35 (black), again assuming that $\langle F \rangle = 0.1$. Three trends become obvious from this plot. First, as the neutral fraction is increased (at fixed $\langle F \rangle$), the number of large saturated regions increases and, second, as the neutral fraction is increased, the size of the largest dark gaps also increases. For example, in the $\langle x_{\text{HI}} \rangle = 0.22$ model, the largest dark gaps are roughly five times bigger than in the fully ionized model. Additionally, we again see hints of the underlying dark gap size distribution being bimodal as the neutral fraction is increased, supporting the idea that the *shape* of the dark gap size distribution may be a diagnostic for the underlying neutral fraction.

Given these trends, it should be possible to compare dark gap distributions from observed spectra against models at various neutral fractions and use this to constrain the mean neutral fraction of the IGM. This approach is appealing in that it does not require especially high-resolution or high signal-to-noise spectra. However, it does require comparison with simulated models of the dark gap size distribution and so the conclusions reached will be somewhat model dependent. Additionally, the distributions are dependent on the assumed mean transmission, which is itself uncertain. In particular, estimates of the mean transmission at high redshift may be impacted by continuum fitting errors, given the inherent difficulty in estimating the unabsorbed continuum level in highly-absorbed spectra.

In order to investigate the impact of possible continuum fitting errors, we generate mock spectra in the fully ionized model with $\langle F \rangle = 0.03$ but then rescale the flux in each simulated pixel by a multiplicative factor – to mimic the effect of continuum misplacement – such that the measured mean transmitted flux appears to be $\langle F_{\text{meas}} \rangle = 0.1$. This case is shown as the magenta dashed line in Fig. 2.3. Here the dark gap size distribution is shifted towards sizes than one would expect in an ionized model at $\langle F \rangle = \langle F_{\text{meas}} \rangle = 0.1$. However, the shape of the size distribution is still quite different than in the partly neutral models. Importantly,

the dark gap distribution in the ionized model still lacks the distinctive bump at large sizes that is the hallmark of a partly neutral IGM in our models.

2.5 Stacking Toy Spectra

In this section we describe our basic approach of stacking Ly α and Ly β spectra in order to detect the presence of HI damping wings and absorption due to deuterium, respectively. While the forest is too absorbed at these redshifts to easily detect damping wings or deuterium absorption due to individual neutral regions, here we demonstrate that the presence of such features can be revealed *on average* by stacking regions of transmission over many spectra.

This section serves as a proof of principle by applying a simplified stacking approach to mock spectra generated using an idealized IGM model. Specifically, we consider an ensemble of sightlines through our simulation box and assume that the IGM is entirely ionized with the exception of a single HI island with mean density and varying length, L , inserted randomly along each line of sight. We then generate mock spectra assuming these density and ionization fields. The stacking in this section is always done starting at the HI/HII boundaries of a given HI region moving outward.

2.5.1 HI Damping Wing

Our stacking approach can be clearly demonstrated by considering the damping wing from neutral hydrogen. Due to the natural width of the Ly α line, a neutral hydrogen gas parcel should cause Ly α absorption over a range of frequencies. Far from line center, this absorption will have an optical depth roughly following (Miralda-Escude and Rees 37):

$$\tau_{\text{Ly}\alpha}^{\text{DW}}(\Delta v) \approx \frac{\tau_{\text{GP}} R_\alpha c}{\pi} \left[\frac{1}{\Delta v} - \frac{1}{\Delta v + v_{\text{ext}}} \right] \quad (2.6)$$

2.5 Stacking Toy Spectra

where τ_{GP} is the Gunn-Paterson optical depth, $R_\alpha \equiv \Gamma_\alpha \lambda_\alpha / 4\pi c$, $\Gamma_\alpha = 6.265 \times 10^8 \text{ sec}^{-1}$ is the Ly α decay constant, Δv is the separation from the HI/HII boundary in velocity space, v_{ext} is the extent of the hydrogen region in velocity space, and c is the speed of light. For a large neutral region, this equation implies that $\tau_{\text{Ly}\alpha}^{\text{DW}}(|v| < 600 \text{ km/s}) \geq 1$ at $z \sim 5.5$. This excess absorption is referred to as the hydrogen “damping wing”. While both neutral gas and highly ionized gas can cause absorption in quasar spectra, only a significantly neutral hydrogen patch will result in damping wing absorption, owing to the greatly reduced optical depth in the wing compared to line center. As such, detecting damping wing absorption would be a smoking gun for the presence of significantly neutral hydrogen islands. Note that the transmission profile will differ from the simple form of Eq. 2.6, owing mostly to neighboring neutral regions, however the gradual recovery to transmission around saturated neutral regions should be a distinctive indicator that highly neutral regions remain in the IGM.

In Fig. 2.4 we show the results of stacking transmission outside of neutral regions in the toy mock spectra described earlier in this section, neglecting deuterium for the time being. Namely, we show the stacked transmission outside neutral islands of length $L = 0.76 \text{ Mpc}/h$ ($\sim 100 \text{ km/s}$) in black, $L = 1.27 \text{ Mpc}/h$ ($\sim 170 \text{ km/s}$) in blue, $L = 5.34 \text{ Mpc}/h$ ($\sim 700 \text{ km/s}$) in cyan, and stacked transmission neglecting the damping wing in red. Additionally, we have plotted the analytic curves corresponding to Eq. 2.6 for the various L values, shown with dashed curves. We have applied a single multiplicative factor to these curves to account for average resonant absorption from ionized gas. Together, this figure implies that damping wing absorption from isolated neutral regions has a significant impact on quasar spectra, extending $\sim 1000 \text{ km/s}$ past the HI/HII boundaries, which may be observable through stacking as expected from Eq. 2.6.

In providing a toy example of how the hydrogen damping wing can affect spectra, we have neglected many important challenges that such a measurement would face. For example, we assumed perfect knowledge of the underlying ionization state of the IGM in

order to determine where to stack and we assumed that we could discriminate between neutral and highly ionized absorption systems. However, the presence of such a large and potentially-observable feature provides motivation for us to apply the stacking approach in a more realistic manner. In §2.6 and in Appendix B, we describe several such challenges and subtleties along with potential resolutions.

2.5.2 Deuterium

With the stacking approach of the previous section in mind, we now consider absorption due to deuterium. As noted in §2.3, primordial hydrogen should be accompanied by traces of deuterium, with a relative abundance of $\sim 2.5 \times 10^{-5}$ (Cooke et al. 12). Due to its slightly increased reduced mass, atomic transitions in deuterium are shifted blueward by 82km/s compared to the same transitions in hydrogen. This implies that absorption due to neutral hydrogen in the IGM should be accompanied by additional absorption from deuterium, shifted blueward by 82km/s. We can estimate the optical depth for Ly α absorption in deuterium at cosmic mean density by simply scaling the hydrogen Ly α optical depth by the deuterium abundance:

$$\tau_{D,\alpha} = \left[\frac{D}{H} \right] \times \tau_{\text{GP}} \approx 8.25x_{\text{HI}} (1 + \delta) \left[\frac{1+z}{6.5} \right]^{3/2}. \quad (2.7)$$

Thus, we see that while the relative abundance of deuterium is extremely small, the Gunn-Peterson optical depth is so large that the resulting deuterium optical depth is still of order 10 in Ly α .

An appealing aspect of searching for damping wing absorption is that the optical depth in the wing is large enough to cause significant absorption in the presence of neutral islands, but small enough to be negligible for ionized absorption systems. We see this again in the case of deuterium absorption, suggesting that it may be useful as an additional “smoking gun” indicator for underlying neutral hydrogen. However, an obvious problem with detecting deuterium in Ly α spectra is that the feature should be narrow and well within the broad

2.5 Stacking Toy Spectra

range of velocities where the hydrogen damping wing is significant. Specifically, according to Eq. 2.6, at $\Delta v = 82\text{km/s}$, the damping wing optical depth for an extended neutral region should be $\tau_{\text{Ly}\alpha}^{\text{DW}}(\Delta v \approx 82\text{km/s}) \approx 8$. Therefore, the feature should be completely wiped out in Ly α spectra by the hydrogen damping wing.

However, the damping wing optical depth in the Ly β line is much smaller. Specifically, according to Eq. 2.6, the damping wing optical depth scales as

$$\begin{aligned} \frac{\tau_{\text{Ly}\beta}^{\text{DW}}}{\tau_{\text{Ly}\alpha}^{\text{DW}}} &= \frac{\tau_{\text{GP},\beta}}{\tau_{\text{GP},\alpha}} \times \frac{R_\beta}{R_\alpha} = \frac{f_\beta \lambda_\beta}{f_\alpha \lambda_\alpha} \times \frac{(\Gamma_\beta + \Gamma_{\text{H}\alpha}) \lambda_\beta}{\Gamma_\alpha \lambda_\alpha} \\ &= \frac{f_\beta^2}{f_\alpha^2} \left(1 + \frac{f_{\text{H}\alpha}}{f_\beta} \frac{\lambda_\beta^2}{\lambda_{\text{H}\alpha}^2} \right) = .0410, \end{aligned} \quad (2.8)$$

and should therefore be significantly narrower in Ly β than in Ly α . In the above expression, f_α , f_β , and $f_{\text{H}\alpha}$ are the oscillator strengths of the Ly α , Ly β , and Balmer- α transitions, respectively, with λ denoting the corresponding wavelengths and Γ denoting the corresponding decay constants. By modifying Eq. 2.6 for Ly β , we see that $\tau_{\text{Ly}\beta}^{\text{DW}}(|\Delta v| \gtrsim 25\text{km/s}) \leq 1$. Therefore we find that *the hydrogen damping wing should not wipe out deuterium absorption features in Ly β* . Furthermore, while the hydrogen damping wing optical depth is reduced by a factor of roughly f_β^2/f_α^2 when considering Ly β , the total optical depth in the deuterium line is only reduced relative to deuterium Ly α by $f_\beta \lambda_\beta / f_\alpha \lambda_\alpha \approx 1/6$, such that the optical depth should still be of order 1 for deuterium Ly β . Therefore, not only should a deuterium absorption feature survive the hydrogen damping wing, but it should still have a strong enough optical depth to cause significant absorption if neutral islands in fact remain.

Naturally, it should be very difficult to detect individual deuterium absorption features from the diffuse IGM, as the Ly- β spectra will be very absorbed when the universe is neutral enough to produce the features in the first place. However, the feature may nonetheless be observable *on average* through the stacking of high-resolution quasar spectra. In order to demonstrate the strength of the deuterium absorption feature in stacked spectra, we

2.5 Stacking Toy Spectra

incorporate deuterium into the same toy sightlines from §2.5.1 to produce mock Ly β spectra, neglecting foreground Ly α absorption for the time being. We are then able to stack transmission outside of neutral regions in the spectra, starting at the HI/HII boundaries and moving outward. However, since deuterium absorption will only occur on the blue side of neutral regions, we need only stack those regions of transmission. In fact, this offers a clean test for detecting deuterium. Namely, we can separately stack transmission redward and blueward of neutral regions and compare. Excess absorption on the blue side of neutral regions, on average, could signal the presence of deuterium absorption. This is especially appealing since there should be no sources of contamination that would cause a similar, and significant, red/blue asymmetry.¹

In Fig. 2.5, we show the results of stacking transmission in Ly β redward (red) and blueward (black) of the toy neutral regions across the full ensemble of mock quasar spectra. As in Fig. 2.4, all stacking begins at HI/HII boundaries. We can see the blueward transmission clearly exhibits excess absorption due to deuterium extending roughly $\sim 80\text{km/s}$ from the HI/HII boundary. Thus, in this idealized scenario, the presence of deuterium in islands of neutral hydrogen leaves a very clear signature in the stacked Ly β transmission.

Before proceeding further, we should point out one important caveat here. In our simulated models, the transition between fully neutral and highly ionized regions is, by construction, perfectly sharp. If this transition is more gradual in reality, then the narrow deuterium feature could be overwhelmed by absorption from mostly ionized hydrogen in this transition region. A minimal scale for this transition region is set roughly by the mean free path to ionizing photons through the neutral IGM, which is only $\lambda_{\text{HI}} \sim 1/(n_{\text{HI}}\sigma_{\text{HI}}) \approx$

¹One source of asymmetry we do find, which can be seen in Fig. 2.9, results from the fact that, when dealing with realistic spectra, we force there to be transmission in Ly β at the locations where stacking begins. This results in a small selection effect, where selected neutral absorption systems have a reduced probability of having nearby neutral regions and have correspondingly-smaller nearby optical depths. For deuterium, this smaller optical depth is shifted blueward, causing *less* absorption on the blue side of the line for $\Delta v \gtrsim 82\text{km/s}$. However, this asymmetry is minor and *opposes* the asymmetry from deuterium absorption.

2.6 Steps of Approach

$6\text{proper kpc}/h \approx 0.8\text{km/s}$. This minimal scale is two orders of magnitude smaller than the scale of the deuterium feature and hence does not present a worry. However, if the edges of the ionized regions tend to experience a reduced ionizing background, this might obscure the deuterium feature, even in the case of a partly neutral IGM. We believe the possibility of detecting this deuterium feature is enticing enough to warrant further investigation.

As was the case in §2.5.1, we have made several simplifying assumptions and have additionally neglected foreground Ly α absorption from the lower-redshift IGM. However, the clear presence of deuterium absorption revealed through the simplified stacking approach provides motivation to also consider applying it to more realistic spectra, as will be discussed in §2.6.

2.6 Steps of Approach

In §2.5, we demonstrated the utility of stacking idealized quasar spectra in order to reveal the presence of the HI damping wing and deuterium absorption. The success of this approach in the toy case provides motivation for us to apply it to realistic mock spectra. In doing so, we must confront the simplifying assumptions made in §2.5.

The most obviously unrealistic assumption made in §2.5 is that we can precisely identify the HI/HII boundaries underlying our spectra. In practice, we will only have access to the level of transmission at each point along the spectra. However, based on Fig. 2.5, the recovery from saturated absorption to transmission occurs within $\lesssim 15\text{km/s}$ in Ly β from the edge of the neutral zone, and should therefore provide a relatively good indicator of the HI/HII boundary. Therefore, we choose to identify stacking locations based on where transmission recovers *in Ly β* . To be clear, for the case of the hydrogen damping wing, we are stacking transmission in the Ly α forest, but we are choosing where to start the stacking based on features in the Ly β forest. A drawback of this approach, when searching for the hydrogen damping wing, is that we are only able to stack regions of the Ly α forest with

2.6 Steps of Approach

corresponding regions in the Ly β forest that are not contaminated by Ly γ absorption. This effectively reduces the amount of usable spectra, since, for a quasar at $z = 5.5$, the pure Ly α forest will extend $4.5 \leq z \leq 5.5$, but Ly γ absorption will contaminate the Ly β forest at $z \lesssim 5.16$. If presented with a limited number of spectra, it may be worth searching for the damping wing by using only the Ly α regions of the spectra.

By stacking at the precise locations of HI/HII boundaries in §2.5, we were also ensuring that our sample of absorption systems was all neutral. However, when we modify our approach to begin stacking at locations where transmission recovers from saturated absorption, we may start stacking transmission outside of ionized absorption systems together with transmission outside of neutral absorption systems, diluting our signal. Since the signal we are aiming to find is small to begin with, it is important that we minimize this contamination from ionized regions. To do this, we take advantage of the main argument of §2.4, namely that regions of saturated absorption sourced by neutral gas should be significantly larger, on average, than those sourced by ionized gas. Therefore, we choose to stack only transmission outside of *large* regions of saturated absorption. Furthermore, since true neutral regions should cause saturated absorption in Ly β , we choose to stack only outside of large saturated regions which are fully absorbed in Ly β , where we define “large” to be $> 500\text{km/s}$ ($\gtrsim 4 \text{ Mpc}/h$) in Ly β . Note that this choice is tuned for the case of $\langle F \rangle = 0.1$: a different choice may be better for other values of the mean transmitted flux. At any rate, in applying these tests to real data, one would likely vary this size scale across a range of possible values.

Additionally, an appealing feature of the search for deuterium absorption is that it offers a very clean test for its detection, namely a red/blue asymmetry in the transmission outside of plausibly neutral regions. The disparity in the size distribution of saturated regions sourced by neutral and ionized gas suggests a similar test may be possible for the detection of the HI damping wing. Namely, while large regions of saturated absorption are likely to be sourced by neutral gas, small regions of saturated absorption are likely to be

2.6 Steps of Approach

sourced by ionized gas. Therefore, to find evidence of excess absorption outside of neutral regions due to the HI damping wing, we compare the stacked transmission outside of large absorption systems, plausibly sourced by neutral gas, to that outside of small absorption systems, likely sourced by ionized gas. A significant amount of excess absorption outside of the former compared to the latter, extending further than any possible density correlations, would suggest the presence of damping wing absorption.

Furthermore, in §2.5.2, we discussed how the damping wing is greatly weakened in Ly β compared to in Ly α . Therefore, an additional test for the presence of damping wing absorption could be to take the ratio of the stacked Ly β transmission to the stacked Ly α transmission, where stacking occurs in the same physical regions in both cases. In the event that there is significant damping wing absorption, this ratio should also slowly recover to some constant value at large Δv . We further discuss and develop this approach in Appendix B.

When dealing with realistic spectra, we must adjust our approach to accommodate the presence of noise (and finite spectral resolution). While noise should average out in stacked regions, the presence of noise will also obfuscate the precise boundaries between saturated absorption and transmission. We choose to handle this by smoothing our noisy spectra over a scale of 100 km/s (~ 0.75 Mpc/ h) and defining any pixel, i , with transmission $F_i < 3\tilde{\sigma}_N$ to be consistent with saturated absorption, where $\tilde{\sigma}_N$ denotes the standard deviation of the smoothed noise. We then define regions in the smoothed spectra where the flux goes from $F < 3\tilde{\sigma}_N$ to $F > 3\tilde{\sigma}_N$ as the transitions from saturated absorption to transmission, and therefore as potential points to start stacking. When stacking transmission, however, we stack the transmission in the *unsmoothed* spectra.

Another concern is that damping wing absorption sourced by DLAs may erroneously be attributed to a significantly neutral IGM. However, in Appendix A, we estimate the expected rate of DLAs occurring in $z \sim 5.5$ quasar spectra and find it is small enough to

be ignored. Additionally, DLAs may be discriminated from diffuse neutral islands based on the presence of metal lines and the relative sizes of their absorption in Ly α and Ly β .

Finally, as mentioned previously, we approximate the ionizing background in the ionized regions as uniform and ignore scatter in the temperature density relation. Accounting for these fluctuations *might* lead to a more gradual recovery in the transmission around absorbed regions – in the case of a fully ionized universe – than in our models. Further investigation of this issue would certainly be required if a gradual recovery is indeed found in real spectra. In Appendix B, we discuss a possible empirical test that may help in this regard.

2.7 Results

Having considered the subtleties of the previous section, we are now ready to apply the three-pronged approach to more realistic mock spectra. In each section, we first consider the ideal case where no noise has been applied to give an idea of the potential constraining power of the different methods. Subsequently, we add realistic levels of noise and consider realistic spectra resolution to give an idea of the constraining power of the approaches applied to Keck HIRES spectra for the deuterium feature and damping wing, and spectra with slightly higher resolution than SDSS for the dark gap size distribution.

2.7.1 Detecting the Damping Wing

We first consider the ability to uncover the presence of the hydrogen damping wing by strategically stacking regions of transmission in the the Ly α forest of $z \approx 5.5$ noiseless mock quasar spectra. As discussed in §2.6, our aim is to compare the average transmission outside of plausibly neutral absorption systems to the transmission outside of likely ionized systems.

We identify the plausibly neutral absorption systems by requiring the regions be completely absorbed in Ly β , and also that the regions of saturated absorption are at least

2.7 Results

$L_{\text{sat}} > 500 \text{ km/s}$ ($\sim 4 \text{ Mpc}/h$) *in Ly* β . We begin stacking at the point in the $\text{Ly}\alpha$ spectrum which corresponds to the recovery from absorption to transmission *in Ly* β . We identify the likely ionized absorption systems by requiring that they are *below* a maximum length $L_{\text{max}} = 300 \text{ km/s}$ *in Ly* β .

In Fig. 2.6, we show the results of applying this approach to realistic mock spectra generated assuming various ionization states of the IGM. In the top panel, we show the stacked transmission outside of plausibly neutral absorption systems (solid) and likely ionized absorption systems (dashed), using a volume-averaged neutral fraction of $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and $\langle x_{\text{HI}} \rangle = 0$ (magenta). The curves agree with our expectations, namely that transmission outside of neutral regions should recover more slowly and exhibit a rough damping wing shape with a large extent in velocity space. We see that the excess absorption extends *farther* than the $\lesssim 1000 \text{ km/s}$ expected from an isolated damping wing. However, as discussed in Appendix C, we find that the spatial clustering of neutral regions is responsible for this effect.

While, for several reasons discussed earlier, the shape of the absorption is distorted compared to Fig. 2.4, it can be seen for all significantly neutral ionization states. An important check is to apply the stacking procedure to a fully ionized IGM and ensure that we do not make a false detection. The results of this check are shown by the magenta curves in Fig. 2.6. As we can see, the resulting stacked transmission outside of plausibly neutral regions lacks an overall damping wing shape and stays roughly fixed near the mean transmission.

We can also see that the transmission outside of small absorption systems is very sensitive to the underlying neutral fraction. We expect this, however, since this stacked transmission depends strongly on the average transmission in regions which are not in saturated absorption, denoted $\langle F|F > 0 \rangle$. Since the dark pixel covering fraction in our mock spectra increases with $\langle x_{\text{HI}} \rangle$, mock spectra with larger neutral fractions must have larger values for

2.7 Results

$\langle F|F > 0 \rangle$ to maintain $\langle F \rangle = 0.1$. As such, Fig. 2.6 shows that the stacked transmission outside of small absorption systems increases monotonically with $\langle x_{\text{HI}} \rangle$.

We estimated the stacked transmission from a large ensemble of simulated spectra to produce a smooth estimate of the average transmission around saturated regions in each model. The transmission curves outside of individual absorption systems are, however, quite noisy on their own such that, from saturated region to saturated region, there is significant scatter about the mean-value curves shown in the top panel. In order to estimate how confidently we can distinguish the solid and dashed curves with a reasonable number of quasar spectra, we scale the number of identified absorption systems to what we would expect using ~ 20 spectra. Specifically, we take the difference between the dashed and solid curves and divide by the scatter in each bin. The scatter of each bin is simply the scatter in stacked transmission outside of large absorption systems, scaled by $1/\sqrt{N_{\text{sat,large}}}$, added in quadrature with the scatter in the stacked transmission outside of small absorption systems, scaled by $1/\sqrt{N_{\text{sat,small}}}$. Here we scale to estimate the plausible scatter around the mean after estimating the transmission around saturated regions using 20 quasar absorption spectra.

The results of this are shown in the bottom panel of Fig. 2.6 for the same ionization states. The results appear to be very encouraging, indicating that, assuming noiseless spectra, the solid and dashed curves are $\gtrsim 5\sigma$ statistically-significantly different (*even for* $\langle x_{\text{HI}} \rangle = 0.05!$). In addition, we see that the difference roughly follows a damping wing shape and remains significant for $\gtrsim 3000 \text{ km/s}$. We should emphasize that, while the deuterium absorption feature will necessarily be a $\lesssim 80 \text{ km/s}$ feature and require high resolution spectra to be seen, the damping wing feature extends an order of magnitude farther in velocity space and should be accessible to lower-resolution spectra.

In Fig. 2.7, we show the same results as in Fig. 2.6, but assume a lower mean transmission of $\langle F \rangle = 0.05$, consistent with spectra at $5.7 \lesssim z < 6$ (Becker et al. 5). From the figure, we see that these results are very similar to those for $\langle F \rangle = 0.10$, but with the significance

2.7 Results

curves peaking at a $\sim 70\%$ lower value and with the stacked transmission recovering to a lower mean. Overall, this provides encouragement for applying the approach to higher- z spectra, suggesting that a range of physically interesting neutral fractions could be probed.

It is also interesting to consider these results when spectra are generated according to the specifications of existing data. In Fig. 2.8 we show the same results as in the bottom panel of Fig. 2.6 except we have adjusted the spectra to mimic HIRES spectra. Namely, we have assumed a spectral resolution with FWHM = 6.7km/s and bins with size $\Delta v_{\text{bin}} = 2.1\text{km/s}$ (e.g. Viel et al. 50). Additionally, we have assumed a signal to noise of SNR = 10 at the continuum per 2.1km/s pixel and that we have 20 such spectra. While we are currently only aware of 10 such spectra, this case is still interesting since spectra with significantly worse spectral resolution should also be adequate for this test.

From this figure, we can see that, despite the degradation of the spectra, the damping wing is still visible with the significance curve peaking at $\gtrsim 5\sigma$ ($\gtrsim 8\sigma$) significance for the $\langle x_{\text{HI}} \rangle = 0.22$ (0.35) ionization state. However, this figure suggests that, in the $\langle x_{\text{HI}} \rangle = 0.05$ case, it is less-clear whether the damping wing is detectable.

An important effect of adding noise to the mock spectra is that it obscures the precise location where spectra should be stacked and also increases the fraction of selected saturated regions which are, in fact, ionized. We find that for the spectra in this section $\sim 30\%$, 40% , and 75% of identified plausibly neutral regions are in fact ionized for $\langle x_{\text{HI}} \rangle = 0.35$, 0.22 , and 0.05 , respectively. This is compared to $\sim 7\%$, 10% , and 20% contamination when noise is neglected.

Statistical significances in this section are only estimates. In reality, the statistical significance with which the damping wing can be detected will depend on how extended the significance curves are, along with how correlated the errors in neighboring bins are. We discuss this in §2.8.

2.7.2 Deuterium Feature Results

We now turn to consider the prospects for identifying deuterium absorption in realistic Ly β mock spectra. As discussed in §2.6, our aim is to identify plausibly neutral absorption systems in the Ly β spectra and compare the stacked transmission moving blueward and redward away from the absorption. We identify the plausibly neutral regions in the same manner as for the damping wing.

In Fig. 2.9, we show the results of applying the stacking approach to the same mock spectra as in the previous section. The top panel shows the mean transmission blueward (solid) and redward (dashed) moving away from plausibly neutral absorption systems for the same ionization states as in the previous section. We can very clearly see excess absorption in the partially neutral spectra, extending $\sim 80\text{km/s}$, consistent with our expectations from Fig. 2.5. Additionally, we also find that, in the fully ionized case, the blueward and redward stacked transmission match up very well.

As in the previous section, we can construct a rough significance curve for the difference between the blueward and redward transmission. Specifically, in the bottom panel of Fig. 2.9 we show the excess blueward absorption in units of the standard deviation of the stacked blueward transmission assuming 20 quasar spectra. We can see that the significance of the red/blue asymmetry extends $\sim 70\text{km/s}$ ($\sim 0.3 \text{ Mpc}/h$) and is $\gtrsim 3\sigma$ for all of the partially neutral models considered, with increasing significance for models with higher neutral fractions. Additionally, we see that the curve corresponding to the fully ionized model shows no statistically-significant deviation from red/blue symmetry. Thus, this is indeed a very clean test for the presence of deuterium. However, the signal itself is an order of magnitude smaller in velocity-space extent and is found with significantly less statistical significance than the damping wing signal. Therefore, we expect that high-resolution, high-signal-to-noise spectra will be necessary to search for it.

As in §2.7.1, we can reproduce Fig. 2.9 assuming spectra with specifications mimicking Keck HIRES. Unfortunately, we find that, with a signal to noise per pixel in the continuum

2.7 Results

of 10, the deuterium feature is hard to observe. Because of this, we consider using 20 HIRES-style spectra with a signal to noise per pixel of 30 in the continuum. While this signal-to-noise value is higher than those for existing spectra we found in the literature, it is not unreasonable to assume such spectra may become available in the future. Furthermore, this may provide additional motivation to obtain such spectra. Regardless, after applying the stacking approach with twenty $\text{SNR} = 30$ HIRES spectra, we obtain the results shown in Fig. 2.10. This figure shows that the feature should be observable with modest statistical significance. Specifically, for $x_{\text{HI}} = 0.35$ (0.22) the significance curve peaks at $\sim 3.7\sigma$ ($\sim 3\sigma$). Additionally, when these curves are generated assuming MIKE-style spectra, with spectral resolution $\text{FWHM} = 13.6\text{km/s}$ and velocity bin size $\Delta v_{\text{bin}} = 5.0\text{km/s}$, we obtain similar curves as in Fig. 2.10 but with the signal being statistically significant over a smaller range of velocities.

Again, important effects of adding noise to the mock spectra are that it obscures the precise location where spectra should be stacked and increases the fraction of selected plausibly neutral regions which are, in fact, ionized. We find that for the spectra in this section $\sim 15\%$, 20% , and 40% of identified plausibly neutral regions are in fact ionized for $\langle x_{\text{HI}} \rangle = 0.35$, 0.22 , and 0.05 , respectively. This is compared to $\sim 7\%$, 10% , and 20% contamination when noise is neglected. As expected, we find a smaller level of contamination than in the previous section, owing to the increased signal to noise of the spectra used. However, for the case of deuterium, the effect of noise on the stacking location is more apparent. Fig. 2.9 demonstrates that, without noise, deuterium absorption imprints a feature on stacked noiseless spectra extending $\approx 80\text{km/s}$, but only extending $\approx 60\text{km/s}$ in stacked noisy spectra, as shown in Fig. 2.10.

The above results suggest that stacking Ly β transmission in high- z spectra can indeed be used to detect the presence of primordial deuterium, and hence that of hydrogen, but that high-resolution and high signal-to-noise spectra will be required. Nevertheless, it would certainly be interesting to apply this approach to existing HIRES/MIKE spectra as it

provides an additional test, independent of the damping wing search, for the presence of underlying neutral hydrogen in the IGM. As such, a detection with modest levels of statistical significance could lend credence to a claimed detection of the HI damping wing.

2.7.3 Dark Gap Statistics

We now shift our focus away from stacking and toward the distribution in lengths of regions of saturated absorption (dark gaps). As discussed in §2.4, the dark gap size distribution in quasar spectra should contain information about the underlying ionization state of the IGM. Specifically, in a more neutral IGM, the typical sizes of dark gaps should be larger and the shape of the dark gap size distribution should have a more gradual decline, and possibly show a hint of bimodality, toward large L .

We continue this discussion in this section by considering plausible dark gap size distributions that could be observed with moderate-resolution, moderate-signal-to-noise spectra. Specifically, we consider spectra with spectral resolution $\text{FWHM} = 100\text{km/s}$, bin size $v_{\text{bin}} = 50\text{km/s}$, and a signal-to-noise ratio of 10 at the continuum. These spectra are of only slightly better quality than SDSS spectra. Additionally, since we are not concerned with Ly β transmission, we are able to use the entire Ly α forest for each spectra.

In Fig. 2.11, we show the resulting dark gap size histogram expected for 20 such spectra for $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and 0 (magenta). In generating this figure, we use the same ensemble of mock spectra as in §2.7.1 and §2.7.2, except with their spectral resolution and bin size modified as mentioned. We maintain the requirement that $\langle F \rangle = 0.1$.

This figure qualitatively agrees with Fig. 2.3, where noiseless spectra with finer spectral resolution were used, but shows a shift toward larger L due to smoothing the spectra. Additionally, the ionization states are not as easily distinguishable as in Fig. 2.3, with the $\langle x_{\text{HI}} \rangle = 0.05$ distribution looking practically identical to the fully ionized scenario. However, for the other neutral fractions considered, the situation looks very encouraging.

The distributions for $\langle x_{\text{HI}} \rangle = 0.22$ and 0.35 show a more gradual decline toward large L than the fully ionized case and also reveal the clear emergence of a bimodal distribution. Additionally, the largest dark gaps in these ionization states are roughly twice as large as in the fully ionized case.

2.8 Forecasts

Having discussed the results of the proposed stacking approaches applied to realistic mock spectra, we now consider the ability of these methods to constrain the ionization state of the $z \sim 5.5$ IGM. Specifically, in this section we focus on the ability of each method to rule out the hypothesis of a fully ionized IGM.

In both the case of the HI damping wing and deuterium absorption feature, we would like to compare models representing different ionization states and estimate the $\Delta\chi^2$ between $\langle x_{\text{HI}} \rangle \neq 0$ models and the fully ionized model, assuming a reasonable number of spectra. Let $F_{\langle x_{\text{HI}} \rangle}(\Delta v)$ denote the mean behavior for a model with given neutral fraction, $\langle x_{\text{HI}} \rangle$, as a function of stacked velocity separation and let $F_{\text{ion}}(\Delta v)$ denote the mean behavior of the ionized model, also as a function of stacked velocity separation. The precise definitions of what is meant by behavior will be discussed later in this section. In this case, the $\Delta\chi^2$ value between two models can be calculated by

$$\Delta\chi^2_{\langle x_{\text{HI}} \rangle} = \Delta F_{\langle x_{\text{HI}} \rangle} \cdot C^{-1} \cdot \Delta F_{\langle x_{\text{HI}} \rangle}^T \quad (2.9)$$

where C is the covariance matrix of the $\langle x_{\text{HI}} \rangle$ model, representing the correlation between stacked pixels, and $\Delta F_{\langle x_{\text{HI}} \rangle}(\Delta v) \equiv F_{\text{ion}}(\Delta v) - F_{\langle x_{\text{HI}} \rangle}(\Delta v)$, with $\Delta F_{\langle x_{\text{HI}} \rangle}^T$ denoting its transpose. For simplicity, rather than estimating the full covariance matrix and its inverse, we approximate pixels at sufficiently wide separations as independent. We then coarsely sample the pixels – on the scale where they can be well approximated as independent – and assume a diagonal covariance matrix for the coarsely sampled pixels. Specifically, we

estimate $\Delta\chi^2_{\langle x_{\text{HI}} \rangle}$ by simply adding up the squared statistical significance of each coarsely-sampled bin, $[\Delta F_{\langle x_{\text{HI}} \rangle}(\Delta v_i)/\sigma_{\langle x_{\text{HI}} \rangle}(\Delta v_i)]^2$, where $\sigma_{\langle x_{\text{HI}} \rangle}(\Delta v_i)$ is the standard deviation of $F_{\langle x_{\text{HI}} \rangle}(\Delta v_i)$.

2.8.1 Deuterium

Perhaps it is best to consider the case of the deuterium absorption feature first. In the case of a fully ionized IGM, the transmission looking blueward and redward from absorption systems should be symmetric on average, with excess blueward absorption only occurring when the IGM is significantly neutral. Therefore, we may calculate the $\Delta\chi^2_{\langle x_{\text{HI}} \rangle}$ between stacked transmission looking redward ($F_{\langle x_{\text{HI}} \rangle, \text{red}}(v)$) and blueward ($F_{\langle x_{\text{HI}} \rangle, \text{blue}}(v)$) from plausibly neutral absorption systems for each ionization state $\langle x_{\text{HI}} \rangle$ to estimate our ability to rule out the hypothesis of a fully ionized IGM in each case.

Thus, in the context of Eq. 2.9, we have

$$\Delta F_{\langle x_{\text{HI}} \rangle}(\Delta v) \equiv F_{\langle x_{\text{HI}} \rangle, \text{red}}(\Delta v) - F_{\langle x_{\text{HI}} \rangle, \text{blue}}(\Delta v) \quad (2.10)$$

$$C_{ij}^{-1} = \delta_{ij}/\sigma_{\langle x_{\text{HI}} \rangle, \text{blue}}^2(v_i) \quad (2.11)$$

where $\sigma_{\langle x_{\text{HI}} \rangle, \text{blue}}(v_i)$ is the standard deviation of the stacked transmission blueward of plausibly neutral absorption systems, assuming a given number of spectra, and we have assumed that we have already resampled $\Delta F_{\langle x_{\text{HI}} \rangle}(v)$ at sufficient velocity separations such that neighboring bins can be approximated as independent. At this point, the only missing ingredient is the minimum separation between two stacked pixels for them to be considered independent. We calculate the correlation function between stacked pixels in Ly β , and find that the correlation has a width of FWHM $\approx 80\text{km/s}$ and, as such, we do not expect to get more than one independent bin within the scale of the deuterium feature. Therefore, a rough estimate of the $\Delta\chi^2_{\langle x_{\text{HI}} \rangle}$ value obtainable in each ionization state can be estimated simply by the peak value in the “significance curves” in Fig. 2.10. Thus, if the underlying neutral fraction of the IGM is $\langle x_{\text{HI}} \rangle = 0.22$ (0.35), then we expect to be able to rule out a

fully ionized IGM with $\sim 3\sigma$ ($\sim 3.7\sigma$) confidence, assuming 20 HIRES/MIKE spectra with signal to noise of 30 per pixel at the continuum. Unfortunately, we do not expect to be able to rule out the hypothesis of a fully ionized IGM if the underlying neutral fraction is $\langle x_{\text{HI}} \rangle \lesssim 0.05$.

2.8.2 HI Damping Wing

Assessing the statistical significance with which we can observe the HI damping wing is slightly more complicated than the deuterium case since the test for its detection is not as simple. When faced with actual spectra, we would look for the presence of significant and extended absorption outside of large absorption systems compared to that outside of small absorption systems.

Therefore, the behavior we would like to compare in each case is the stacked transmission outside of plausibly neutral absorption systems ($f_{\text{large}}(\Delta v)$) and the transmission outside of likely ionized absorption systems ($f_{\text{small}}(\Delta v)$). Let us denote

$$F(\Delta v) \equiv f_{\text{small}}(\Delta v) - f_{\text{large}}(\Delta v) \quad (2.12)$$

as the difference in these stacked transmissions where $F_{\langle x_{\text{HI}} \rangle}(\Delta v)$ and $F_{\text{ion}}(\Delta v)$ represent this behavior for the ionization state with neutral fraction $\langle x_{\text{HI}} \rangle$ and the fully ionized state, respectively. Thus, in the context of Eq. 2.9, we have

$$\Delta F_{\langle x_{\text{HI}} \rangle}(\Delta v) = F_{\text{ion}}(\Delta v) - F_{\langle x_{\text{HI}} \rangle}(\Delta v) \quad (2.13)$$

$$C_{ij}^{-1} = \delta_{ij} / \sigma_{F_{\langle x_{\text{HI}} \rangle}}^2(v_i) \quad (2.14)$$

where $\sigma_{F_{\langle x_{\text{HI}} \rangle}}(v_i)$ denotes the standard deviation of $F_{\langle x_{\text{HI}} \rangle}(v)$ at v_i . The resulting $\sqrt{\Delta\chi^2}$ value indicates the expected significance with which we could rule out a *fully ionized* IGM if the neutral fraction were, in fact, $\langle x_{\text{HI}} \rangle$. Again, for Eq. 2.14, we have assumed that $\Delta F_{\langle x_{\text{HI}} \rangle}(\Delta v)$ has been resampled at velocity separations such that the pixels can be treated as independent. We find that the correlation function between pixels of stacked transmission

2.9 Conclusion

in the Ly α forest within the scale of the HI damping wing has FWHM $\approx 100\text{km/s}$. While this scale is large, the excess absorption due to the presence of damping wing absorption leaves a feature extending $\gtrsim 3000\text{km/s}$, leaving us with $\gtrsim 30$ independent bins within the scale of the feature.

In this manner, we are able to calculate a rough estimate for the $\Delta\chi^2$ values for the ionization states considered thus far. Assuming the same type of spectra as in Fig. 2.6, namely 20 HIRES spectra with signal to noise in the continuum of 10 per pixel, we find that if the IGM is, in fact, 5%, 22%, or 35% neutral, then we should be able to rule out a fully ionized IGM at 5.3σ , 19.2σ , or 26.3σ , respectively. In the case of $\langle F \rangle = 0.05$, this reduces to 14.8σ , 8.7σ , and 2.2σ , respectively.¹ While we are only aware of ~ 10 such spectra that exist at the moment, we still regard this estimate as somewhat conservative. We found that excess stacked absorption due to the damping wing extends thousands of km/s, and therefore it is not necessary to have the state-of-the-art in spectral resolution to measure it. Especially with such extended correlation among neighboring pixels, it is unclear how much is gained by resolution improvements beyond $\sim 100\text{km/s}$.

2.9 Conclusion

In this work, we developed empirical tests of the possibility that the Epoch of Reionization is not yet complete by $z \sim 5.5$. Specifically, we proposed three measurements that can be made with existing, and future, high-redshift quasar spectra in order to investigate this region of reionization history parameter space.

First, we discussed using the dark gap size distribution in quasar spectra as a means of constraining the $z \sim 5.5$ neutral fraction. We find that not only do the typical sizes of dark gaps increase with neutral fraction but that the *shape* of the size distribution is also sensitive to the neutral fraction. Specifically, the presence of dark gaps sourced by significantly neutral hydrogen islands introduces a bimodality in the dark gap size distribution. We find

2.9 Conclusion

that this bimodality should be observable at $z \sim 5.5$, provided that $\langle x_{\text{HI}} \rangle \gtrsim 0.2$, and should not be affected by continuum fitting errors.

Next, we proposed a method for searching for hydrogen damping wing absorption by strategically stacking regions of transmission in the Ly α forest. Specifically, we searched for excess absorption in stacked transmission outside of plausibly neutral regions compared to that outside of likely ionized regions. We found that the presence of the hydrogen damping wing will result in excess absorption extending $\sim 5000 \text{ km/s}$ past ionization boundaries of neutral regions. Furthermore, this excess absorption should be detectable with $\gtrsim 5.3\sigma$ statistical significance for $\langle x_{\text{HI}} \rangle \gtrsim 0.05$, using 20 HIRES-style spectra with a signal-to-noise value per pixel of 10 at the continuum.

Lastly, we proposed a similar stacking measurement utilizing the Ly β forest in order to search for deuterium absorption associated with significantly neutral hydrogen islands at $z \sim 5.5$. We proposed searching for this feature by looking for excess absorption in stacked Ly β transmission blueward of plausibly neutral regions compared to the corresponding redward transmission. We find that this feature should be observable in principle but will likely require additional high-resolution spectra in order to be detected. Specifically, we find that the feature should be observable at $z \sim 5.5$ with $\sim 3\sigma$ ($\sim 3.7\sigma$) confidence using 20 HIRES-style spectra with a signal-to-noise value per pixel of 30 at the continuum if $\langle x_{\text{HI}} \rangle = 0.22$ (0.35). While we are not aware of this many available spectra with such specifications, this provides motivation for acquiring such spectra in the future, possibly through the follow-up observation of SDSS quasars.

While we have taken many steps to ensure that the analysis of mock spectra presented in this work is realistic, there are still additional complexities that will be faced when one is presented with actual spectra. For example, we treat all portions of our spectra as being at $z = 5.5$ when, in reality, the redshift will evolve along the lines of sight. In addition, we ignored spatial fluctuations in the UV radiation field and in the temperature density relation. Additional work will certainly be required to definitively interpret future

2.9 Conclusion

measurements along the lines we suggest here. However, we believe the signatures explored here are well-worth further investigation and should ultimately improve our understanding of the reionization history of the IGM.

Appendix A: Contamination from DLAs?

A potential concern is that damping wings from super Lyman-limit systems and damped Ly- α absorbers (DLAs) might produce “false positives” and contaminate our search for diffuse neutral islands. Since DLAs and super Lyman-limit systems are mostly associated with galaxies and the circumgalactic medium (see Wolfe et al. 51 for a review), we would like to distinguish these absorbers from the more diffuse and spatially extended islands of neutral hydrogen that are the subject of our search. In addition, note that it is difficult to fully resolve and model high column density absorbers in cosmological simulations (e.g. Rahmati and Schaye 42 and references therein) – especially given our present aim of capturing the large-scale features of reionization – and so the impact of these systems is not captured in our present modelling.

Fortunately, we don’t expect these dense absorbers to be a big contaminant, provided we make use of the Ly- β forest – which helps owing to the lower cross section in the wing of the line (compared to Ly- α) – and confine our search to fairly extended neutral islands. The Ly- β line profile for a high column density absorber can be approximated by a Lorentzian, so that the optical depth at velocity offset Δv is:

$$\tau_{\beta,\text{DLA}}(\Delta v) = N_{\text{HI}} \frac{\sigma_{\beta,0}}{\pi} \frac{R_\beta}{(\Delta v/c)^2 + R_\beta^2}. \quad (2.15)$$

For comparison, a fully neutral and isolated absorber of co-moving length L_{neut} produces saturated Ly- β absorption over a velocity extent slightly larger than $\Delta v_{\text{neut}} = H(z)L_{\text{neut}}/(1+z)$. We can then determine the column density required for a DLA to produce as long a saturated region in the Ly- β forest as produced by a neutral island of length L_{neut} . We

2.9 Conclusion

consider a DLA to produce saturated absorption at velocity separations where $\tau_{\beta,\text{DLA}} \geq 3$. Provided the extent of the absorber is large enough that $\Delta v_{\text{neut}}/c \gg R_\beta$ (which is a good approximation for the extended neutral islands of interest), this critical column density, $N_{\text{HI},\text{crit}}$, is given by:

$$N_{\text{HI},\text{crit}} = 7.2 \times 10^{21} \text{ cm}^2 \left[\frac{\tau_{\beta,\text{DLA}}}{3} \right] \left[\frac{1+z}{6.5} \right] \left[\frac{L_{\text{neut}}}{3.8 \text{ Mpc}/h} \right]^2. \quad (2.16)$$

The fiducial value of L_{neut} in the above equation corresponds to $\Delta v_{\text{neut}} = 500$ km/s – this is the minimum saturated stretch included in our stacks when we search for neutral regions (see §2.6). The column density $N_{\text{HI},\text{crit}}$ required for a DLA to produce this much saturated absorption is quite large, and the abundance of DLAs with column densities larger than $N_{\text{HI},\text{crit}}$ is very small.

Quantitatively, taking the Gamma function fit to the column density distribution of DLAs from Prochaska et al. (41)¹ (which accounts for the sharp cutoff in the observed abundance of DLAs at high column densities), we find that the number of DLAs with $N_{\text{HI}} \geq N_{\text{HI},\text{crit}}$ is only $dN(> N_{\text{HI},\text{crit}})/dz = 1.5 \times 10^{-3}$. For reference, the redshift extent of the forest between the Ly- α and Ly- β emission line at these redshifts is roughly $\Delta z \approx 1$, and so such high column density DLAs should be exceedingly rare. Since $N_{\text{HI},\text{crit}}$ is only a little smaller than the exponential cut-off in the column density distribution function, the results are rather sensitive to the precise choice of $N_{\text{HI},\text{crit}}$. Given that smaller column-density DLAs might still leak into our stack if they happen to be next to saturated ionized regions, it is worth checking this dependence. However, even choosing $N_{\text{HI},\text{crit}} = 2 \times 10^{20}$ cm² yields only $dN/dz = 0.43$, which is still smaller than the abundance of neutral islands

¹This is for the case where we do not attempt to further optimize the analysis for the decrease in transmission. It is possible that further gains could be made, with Fig. 2.7 representing a best-possible-case scenario.

¹Specifically, we use their highest redshift bin fit, which includes DLAs between redshifts $3.5 \leq z \leq 5.5$.

2.9 Conclusion

we seek to detect. From these estimates, we expect very minimal contamination from DLAs leaking into our stacked sample of possible neutral regions. Note also that deuterium Ly- β absorption from these high column density absorbers will be in the saturated part of the HI Ly- β line, and so DLAs should not contaminate our search for the deuterium signature of neutral islands either.

A separate possible worry is that DLAs could instead contaminate our sample of *small* saturated regions that we use for comparison purposes (as described in §2.6). Our small saturated sample is meant to reflect absorption around saturated yet ionized regions, and so should not contain significant damping wing absorption. In principle, wings from any DLAs in *this* sample could influence the transmission around the small saturated regions. It seems unlikely that this is a significant worry, since the saturated yet ionized regions are likely vastly more abundant than even the small column density DLAs and super Lyman-limit systems. In addition, we can simply inspect the profile of the small saturated sample to see whether it shows any hint of damping wing absorption that might arise from either small isolated neutral regions or DLAs.

Although contamination from DLAs does not appear to be a big worry for our tests, a more detailed examination would certainly be warranted if possible neutral islands are discovered in real data. We may also be able to remove DLA-contaminated regions by flagging spectral regions in the Ly- α and Ly- β forest that have the same redshift as strong metal absorbers, which generally accompany DLAs (see e.g. Wolfe et al. 51).

Appendix B: Further Utilizing the Ly β Forest

In order to infer the presence of the HI damping wing, we would like to compare the stacked transmission outside of plausibly neutral absorption systems to what that transmission would have been in the absence of the damping wing. Up to this point, we have been using the stacked transmission outside of *small* absorption systems as a proxy for the latter

2.9 Conclusion

quantity. From there, we argued that any extended excess absorption outside of large, plausibly neutral absorption systems compared to small, likely ionized absorption systems is indicative of the presence of damping wing absorption.

However, we do have another handle on what transmission would be like in the absence of the HI damping wing and that is the transmission in the Ly β forest. As discussed in §2.5.2, the HI damping wing should be significantly reduced in the Ly β forest compared to the Ly α forest. Specifically, we saw that the damping wing optical depth in Ly β falls to less than one at velocity separations $\gtrsim 25\text{km/s}$ from neutral gas. Therefore, at separations greater than this, stacked transmission in the Ly β forest should provide information on what the shape of the Ly α transmission would have been in the absence of the damping wing, with foreground Ly α absorption only altering the stacked Ly β transmission by an overall constant $\langle F_\alpha(z_\alpha) \rangle$. Using information from stacked Ly β transmission has the appeal that it does not require using physically different regions of space in order to estimate the damping-wing-less transmission outside of selected plausibly neutral absorption systems. This provides protection from problems arising from unanticipated differences between the small, likely ionized absorption systems and the large, plausibly neutral absorption systems.

Thus, we would like to find a way to estimate the Ly α transmission in the absence of the damping wing by using only the Ly β transmission. In principle, this can be done by using simulations to model the relationship between the two and generating a (ionization-state-dependent) mapping that takes a measurement of stacked Ly β transmission outside of large absorption systems and maps it to an estimate of the damping-wing-less Ly α transmission in the same regions. From there, the ratio of the stacked Ly α transmission to this estimate of the damping-wing-less Ly α transmission would leave us with an estimate of $e^{-\tau_{\text{DW}}(v)}$. In the left-hand panel of Fig. 2.12, we show the recovered $e^{-\tau_{\text{DW}}(v)}$ curve after applying this approach to each of the ionization states considered thus far, and then normalizing each curve to peak at 1. Specifically, a mapping between stacked Ly β transmission and stacked damping-wing-less Ly α transmission for each ionization state was generated using a large

2.9 Conclusion

ensemble of mock spectra and then applied to groups of 20 spectra. The error bars in the figure show the scatter in the estimated damping wing absorption between realizations of 20 spectra. For ease of viewing, we show only the error bars for $\langle x_{\text{HI}} \rangle = 0$ and 0.35. This figure demonstrates that the approach works well and recovers a damping wing shape for an IGM with $\langle x_{\text{HI}} \rangle \gtrsim 0.05$ (in the absence of noise).

A few things are worth pointing out about this process. First, the recovered damping wing profiles are only useful to the extent that they provide confidence that we are, in fact, observing neutral hydrogen in the IGM. The stacked profile of the HI damping wing is a complicated entity and, as such, we *do not* expect to be able to, for example, fit the recovered curves to Eq. 2.6 and estimate $\langle x_{\text{HI}} \rangle$. Secondly, it is comforting to note that not only is no damping shape recovered in the case of $\langle x_{\text{HI}} \rangle = 0$, but even if a mapping corresponding to a significantly neutral IGM is applied to a measurement of a fully ionized IGM, we do not recover a damping wing shape. Therefore, we do not expect this approach to yield false positives. Lastly, this process relies on simulations in order to map the stacked Ly β transmission to damping-wing-less Ly α transmission and is therefore somewhat model-dependent. However, we do not expect the specifics of reionization to significantly impact this mapping and we are also not interested in the fine details of the results here. We are primarily interested in whether damping wing absorption can be measured *at all* in the case of a somewhat neutral IGM and, as such, we are comfortable with this level of model dependence.

Finally, we find that this mapping is relatively simple. Namely, for velocity separations $\gtrsim 100 - 200 \text{ km/s}$, the stacked Ly β transmission and the stacked damping-wing-less Ly α transmission differ by roughly a constant multiplicative factor. Thus, in aiming to recover the *shape* of the stacked damping wing absorption, it appears to be a very good approximation to simply divide the stacked Ly α transmission by the stacked Ly β transmission. In the right-hand panel of Fig. 2.12 we simply take groups of 20 spectra and divide their stacked Ly α transmission by their stacked Ly β transmission (outside of large absorption

2.9 Conclusion

systems) and give the result unity amplitude. Qualitatively, the results look very similar to those obtained from the mapping procedure (shown on the left-hand side) but are without any model-dependence. Additionally, we again find that, in the case of a fully ionized IGM, we do not recover a damping wing shape. Thus, this provides another check which may be performed with actual spectra in order to bolster confidence that a damping wing is in fact being observed.

A potential concern for our Ly α stacking approach in general could be that, while we make the approximation that ionized regions are exposed to a uniform ionizing background, the ionizing background will in fact be fluctuating spatially. It is then possible that, in scenarios where the ionizing background is weaker closer to the stacking location and stronger farther from the stacking location, an extended recovery could be imprinted on the stacked transmission despite the IGM being fully ionized. If one were not careful, and if these spatial fluctuations occurred on scales comparable to the damping wing feature, a false detection could be possible. One tool we have to protect against this is the fact that the scale of the damping wing is significantly smaller in Ly β than in Ly α . Therefore, any extended recovery in stacked transmission which occurs over similar scales in Ly α and Ly β is unlikely to be caused by the damping wing.

Appendix C: Extended damping wing Absorption from Correlated HI Islands

As mentioned in §2.7.1, and seen in Fig. 2.6, 2.7, & 2.8, stacked Ly α transmission outside of large saturated regions in a significantly neutral IGM displays excess absorption extending significantly past the scale of an isolated damping wing. To explain this, we seek to model the expected transmission outside of a neutral region, incorporating the correlation between the neutral region at the origin and neighboring neutral regions. In order to simplify the calculation, while capturing the main effect, we ignore correlations between the neighbors

2.9 Conclusion

themselves – i.e., we only include the correlation between the neutral region at the origin and the neighbors and ignore inter-neighbor correlations.

Ignoring correlations between the neighboring neutral regions, we can approximate them as following a Poisson distribution and consider the total absorption contributed by these neutral islands or “clouds” following (54). Suppose that on average m clouds contribute to the absorption at a given region of the spectrum. Let $F_k \equiv e^{-\tau_1} e^{-\tau_2} \dots e^{-\tau_k}$ denote the transmission when k clouds reside along the line of sight and impact the given spectral region, with τ_i being the optical depth of the i th cloud. If we allow the clouds to be placed independently and if they have equal optical depths, then:

$$\langle F_k \rangle \equiv \langle e^{-\tau_1} e^{-\tau_2} \dots e^{-\tau_k} \rangle = \langle e^{-\tau_1} \rangle \langle e^{-\tau_2} \rangle \dots \langle e^{-\tau_k} \rangle = (e^{-\tau})^k. \quad (2.17)$$

Using this expression, we can calculate the ensemble-averaged transmission by averaging over all the possible numbers of intervening clouds:

$$\langle F \rangle = \sum_{k=0}^{\infty} \text{Pois}(k; m) \langle F_k \rangle = \sum_{k=0}^{\infty} \frac{e^{-m}}{k!} m^k \langle e^{-\tau} \rangle^k = e^{-m} \sum_{k=0}^{\infty} \frac{(me^{-\tau})^k}{k!} \quad (2.18)$$

$$= e^{-m} e^{me^{-\tau}} = e^{-m(1-e^{-\tau})}. \quad (2.19)$$

Let us define the quantity τ_{eff} as

$$e^{-m(1-e^{-\tau})} \equiv e^{-\tau_{\text{eff}}} \quad (2.20)$$

$$\tau_{\text{eff}} \equiv m(1 - e^{-\tau}). \quad (2.21)$$

We now have the absorption from neighboring clouds, characterized by the parameter τ_{eff} .

We just need to adapt this slightly to the problem at hand. Suppose that – with certainty – there is a neutral region located at $v = 0$ and let us consider the excess absorption, above random, contributed by neighboring neutral regions. First, the optical depth at $v = \Delta v$ that is contributed by a neighboring cloud located at $v = v'$ will depend on $|\Delta v - v'|$: clouds closer to $v = \Delta v$ will have larger optical depths at that corresponding frequency. We can

2.9 Conclusion

account for this by substituting $\tau \rightarrow \tau(|\Delta v - v'|)$, according to Eq. 2.6. Next, the expected number of HI islands in a region with velocity extent dv' nearby our stacking location can be approximated as

$$m \approx dv' \langle n_{\text{HI}} \rangle (1 + \xi_{\text{HI,HI}}(v')) \quad (2.22)$$

where $\langle n_{\text{HI}} \rangle$ is the average number of HI islands per interval dv' and $\xi_{\text{HI,HI}}(v')$ is the correlation function between the centers of neutral regions separated by v' . From here, we can model the effective optical depth at a given velocity separation due to neighboring HI islands, $\tau_{\text{eff}}(\Delta v)$, as

$$\tau_{\text{eff}}(\Delta v) = \int dv' \langle n_{\text{HI}} \rangle (1 + \xi_{\text{HI,HI}}(v')) \left[1 - e^{-\tau(|\Delta v - v'|)} \right]. \quad (2.23)$$

In other words, the excess effective optical depth from neighboring systems involves the convolution of the absorption profile around each region with the correlation function of the neutral regions. This is analogous to the “two-halo” term in the halo model (e.g. Cooray and Sheth 13).

Thus, the model for the overall stacked transmission outside of neutral regions, also incorporating the damping wing from the central neutral region, becomes:

$$\langle F \rangle (\Delta v) = e^{-\tau_{\text{DW}}(\Delta v)} e^{-\tau_{\text{eff}}(\Delta v)}. \quad (2.24)$$

This model requires two inputs. First, it requires the correlation function between the centers of neutral regions, $\xi_{\text{HI,HI}}(v')$, which can be calculated from a model of the underlying ionization field. Second, the optical depth profile for neutral regions, $\tau(|\Delta v - v'|)$, is a function of the size of the neutral regions, per Eq. 2.6. While the neutral regions in the mock spectra take on a range of sizes, we make the simplifying approximation here – but not in the body of this paper – that the neutral regions at a given neutral fraction have one typical size and denote this L_{typical} with a corresponding extent in velocity space v_{ext} . This

2.9 Conclusion

effectively results in the optical depth profile of an individual neutral island in the right hand side of Eq. 2.23 being described by a piecewise function

$$\tau_{\text{DW}}(\Delta v) = \begin{cases} \infty & \text{if } |\Delta v| < v_{\text{ext}}/2 \\ \frac{\tau_{\text{GP}} R_\alpha c}{\pi} \left[\frac{1}{\Delta v - v_{\text{ext}}/2} - \frac{1}{\Delta v + v_{\text{ext}}/2} \right] & \text{otherwise.} \end{cases} \quad (2.25)$$

Next, we would like to compare this model against results using mock spectra. We do this by first generating mock spectra which *only include absorption from neutral islands*, since this is the only type of absorption incorporated in our model, and stack these mock spectra at the HI/HII boundaries. To be clear, while the model curve described above adopted a fixed L_{neut} for the purpose of calculating a τ_{eff} , the mock spectra here are generated using the same simulated ionization fields as throughout the rest of the paper, with a wide range of sizes for the underlying neutral regions.

To obtain a model for the stacked transmission, we first calculate the correlation function between the centers of neutral regions using the mock underlying ionization fields and also choose a value for L_{typical} to be used in the optical depth profile. Additionally, in Eq. 2.23, the $(1 + \xi_{\text{HI,HI}}(v'))$ term effectively breaks our integral into two pieces: the first representing the mean absorption from neutral regions and the second representing the excess or reduced absorption at $v = \Delta v$ due to the clustering of neutral islands. For our case, we are only concerned with the *excess* absorption, so we make the replacement $(1 + \xi_{\text{HI,HI}}(v')) \rightarrow \xi_{\text{HI,HI}}(v')$.

Therefore, by providing a value for L_{typical} and measuring $\xi_{\text{HI,HI}}(v')$, we can get a estimate for the mean transmission outside of neutral regions which incorporates absorption from spatially-correlated neighboring regions. In the left panel of Fig. 2.13, we plot an example of this for $\langle x_{\text{HI}} \rangle = 0.22$. We show the modelled damping wing absorption from the central neutral region in blue, the modelled absorption from neighboring neutral islands and their damping wings in cyan, and the product of these in black. For comparison, we show the stacked transmission in the mock spectra in magenta, shifted by $v_{\text{ext}}/2$ to account

2.9 Conclusion

for stacking occurring at HI/HII boundaries instead of at the center of neutral regions. All curves have been divided by the mean transmission. After taking $L_{\text{typical}} = 3.2 \text{ Mpc}/h$, we find good agreement between the above model and the stacked transmission. The precise agreement should be taken with a grain of salt, since the model makes several simplifying assumptions, especially that the neutral regions have a fixed size. However, the model does demonstrate that clustered neutral islands should lead to extended excess absorption, significantly beyond the scale of an individual damping wing.

In the right panel of Fig. 2.13, we show the comparison between the stacked transmission (solid) and modelled transmission (dashed) for $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), and 0.05 (cyan), where each curve has been multiplied by the mean transmission for clarity. In generating these plots, we have assumed $L_{\text{typical}} = 2.5 \text{ Mpc}/h$, $3.2 \text{ Mpc}/h$, and $0.75 \text{ Mpc}/h$ for $\langle x_{\text{HI}} \rangle = 0.35$, 0.22, and 0.05, respectively. We again find a very nice agreement between the modelled and stacked transmission, further confirming that spatially-correlated regions are indeed responsible for the significantly-extended excess absorption.

2.9 Conclusion

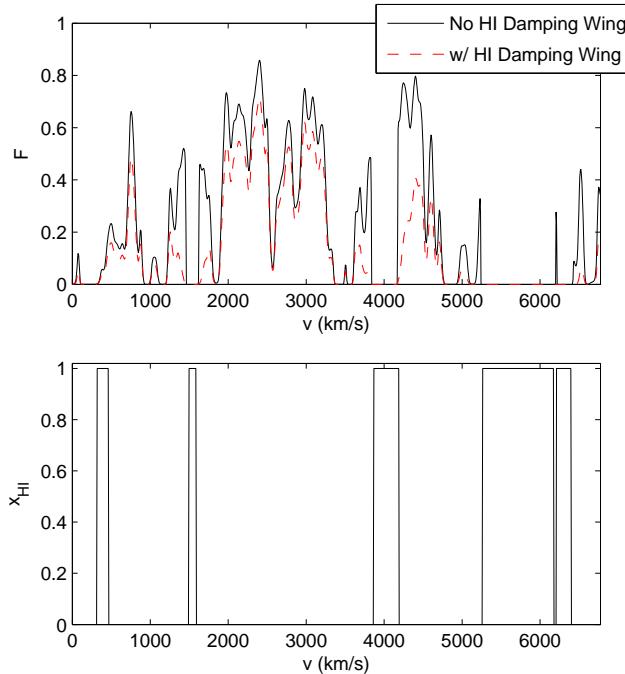


Figure 2.1: Example mock Ly α forest spectrum and corresponding neutral fraction. The top panel shows the Ly α transmission while the bottom panel is the neutral fraction along the line of sight, with ionized regions set to $x_{\text{HI}} \approx 0$ for illustration. The black curve in the top panel shows the transmission through the forest when absorption due to the hydrogen damping wing is neglected, while the red curve includes damping wing absorption. The comparison illustrates that damping wing absorption has a prominent impact, but it is also clear that the presence of the damping wing will be hard to discern by eye. The line of sight is extracted from a model with $\langle x_{\text{HI}} \rangle = 0.22$, but note that we have deliberately chosen a sightline with more neutral regions than typical.

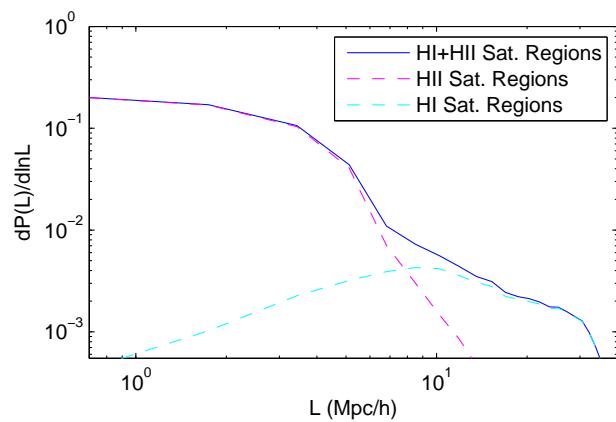


Figure 2.2: Dark gap size distribution for the $\langle x_{\text{HI}} \rangle = 0.22$, $\langle F \rangle = 0.1$ model. The solid blue curve shows the total distribution of dark gaps from an ensemble of mock spectra, where the magenta (cyan) curve shows the same thing but for the dark gaps sourced by ionized (neutral) gas. Here, we have focused on dark gaps with $L > 0.75 \text{ Mpc}/h$. This clearly demonstrates that neutral hydrogen is the dominant source of *large* dark gaps in our mock spectra, provided there is an appreciable neutral fraction.

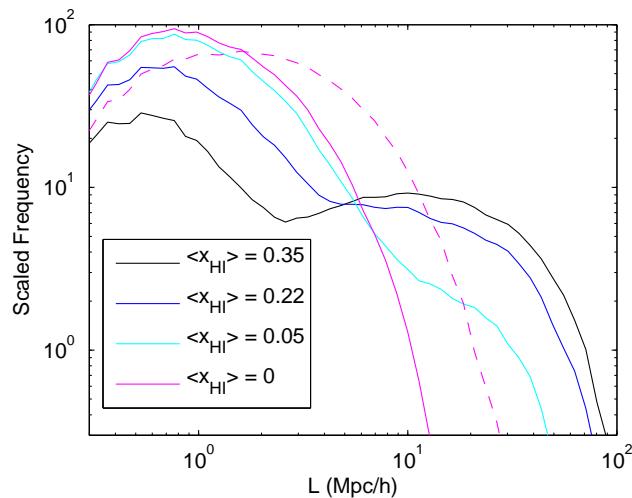


Figure 2.3: Large-length tail of the dark gap size histogram for $\langle x_{\text{HI}} \rangle = 0$ (magenta), 0.05 (cyan), 0.22 (blue), and 0.35 (black) for the case when $\langle F \rangle = 0.1$. The y-axis is scaled to indicate the expected number of dark gaps obtainable from 20 spectra. Bins in this figure are spaced logarithmically. The dashed magenta line indicates the dark-gap size distribution in the fully ionized case when the true transmission is $\langle F \rangle = 0.03$, but continuum fitting errors result in a measured mean transmission of $\langle F_{\text{meas}} \rangle = 0.1$.

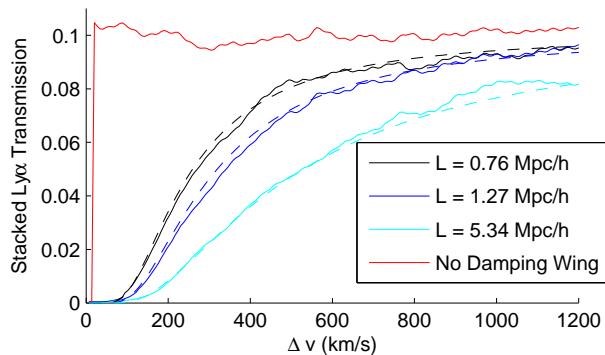


Figure 2.4: Stacking idealized Ly α spectra containing toy HI regions. The above figure shows the stacked transmission outside isolated HI regions with mean density and size $L = 0.76 \text{ Mpc}/h$ ($v_{\text{ext}} \approx 100 \text{ km/s}$), $L = 1.27 \text{ Mpc}/h$ ($v_{\text{ext}} \approx 170 \text{ km/s}$), and $L = 5.34 \text{ Mpc}/h$ ($v_{\text{ext}} \approx 700 \text{ km/s}$) shown in black, blue, and cyan, respectively. The solid red curve shows the stacked transmission outside of the same HI regions *neglecting* the damping wing, which will be the same on average in all cases. In generating these spectra, we assume $\langle F \rangle = 0.1$. In this greatly-idealized case, the presence of the hydrogen damping wing is seen clearly through extended excess absorption compared to the red curve. Furthermore, we can see that the excess absorption closely follows what we would expect analytically based on multiplying Eq. 2.6 by the overall mean transmission. In this figure, all stacking starts at HI/HII boundaries.

2.9 Conclusion

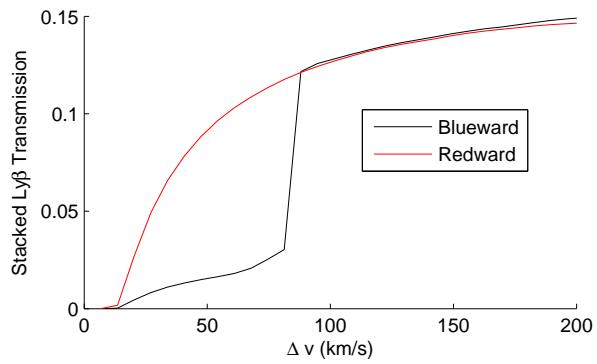


Figure 2.5: Presence of deuterium absorption revealed through stacking idealized Ly β spectra containing toy neutral regions. The red and black curves show the stacked Ly β transmission redward and blueward, respectively, of toy neutral regions of length $L = 5 \text{ Mpc}/h$ ($\approx 700 \text{ km/s}$) randomly inserted into many sightlines, with spectra generated assuming $\langle F_{\text{Ly}\alpha} \rangle = 0.1$. In each case, stacking begins at the underlying HI/HII boundary. We have also mimicked the effect of including foreground Ly α absorption by scaling the feature by the mean transmission in the foreground Ly α forest. This demonstrates that, at least in this idealized case, the presence of deuterium absorption can be easily seen out to $\sim 80 \text{ km/s}$ past the HI/HII boundaries.

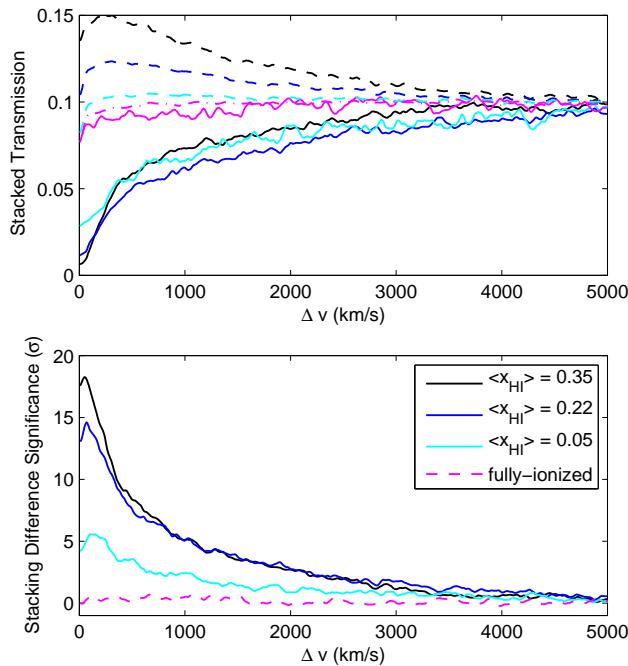


Figure 2.6: Ly α stacking results for various neutral fractions. The top panel shows the mean (noiseless) stacked transmission outside of large absorption systems (solid) and small absorption systems (dashed) in the Ly α forest for neutral fractions $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (red), and 0 (magenta). The transmission here is estimated from a large ensemble of mock spectra to obtain a smooth estimate of the average transmission around saturated regions in each model. The bottom panel shows the statistical significance of the difference between the dashed and solid curves in the top panel assuming a sample of 20 spectra are used in the stacking process.

2.9 Conclusion

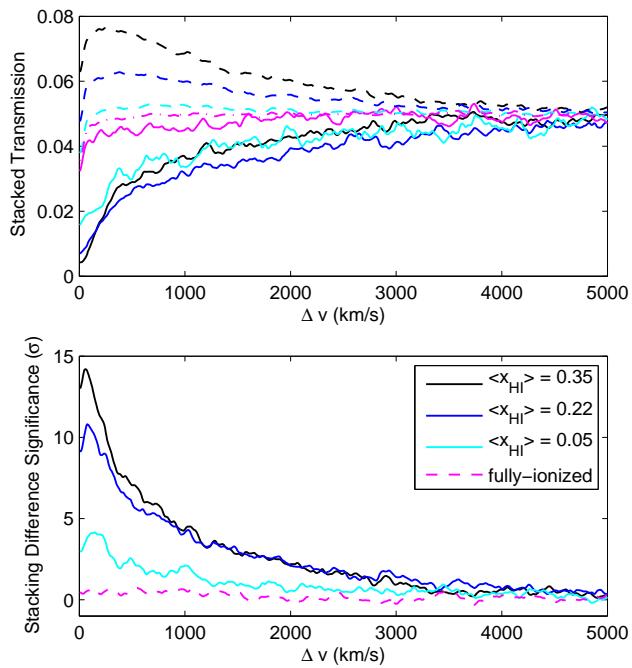


Figure 2.7: Ly α stacking results assuming $\langle F \rangle = 0.05$. The above panels are identical to those in Fig. 2.6 except that mock spectra have been generated assuming $\langle F \rangle = 0.05$.

2.9 Conclusion

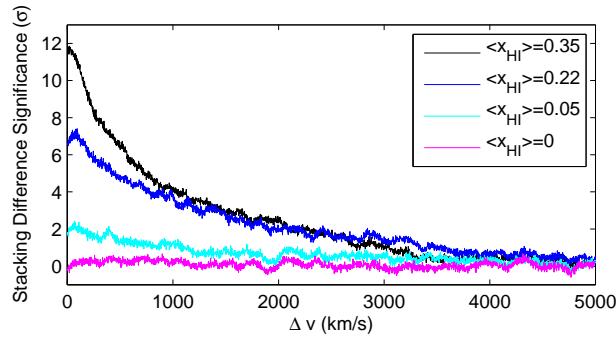


Figure 2.8: Results of Ly α stacking with HIRES-style spectra ($\langle F \rangle = 0.1$). The above panel is identical to the bottom panel in Fig. 2.6 except that the spectra have had the bin size and spectral resolution adjusted to match that of Keck-HIRES spectra. Additionally, we have added noise such that the spectra have a signal-to-noise value of 10 per pixel at the continuum.

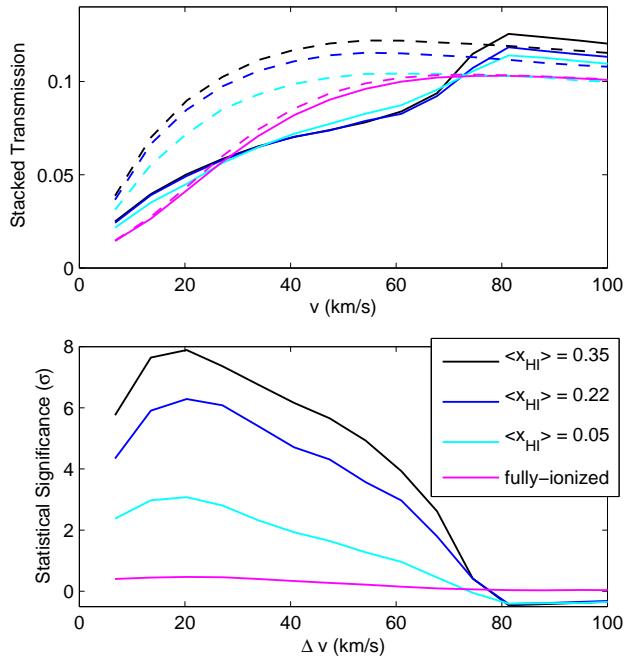


Figure 2.9: Deuterium Ly β stacking results for various neutral fractions. The top panel shows the mean ensemble-averaged noiseless stacked transmission moving blueward (solid) and redward (dashed) away from large absorption systems in the Ly β forest for neutral fractions $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and 0 (magenta). The bottom panel shows the excess blueward absorption in units of the standard deviation of the stacked redward transmission, assuming 20 spectra.

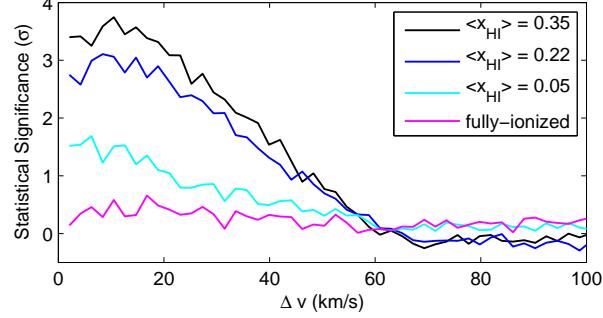


Figure 2.10: Results of Ly β stacking with HIRES-style spectra. The above panel is the same as in the bottom panel of Fig. 2.9, except that it is generated using HIRES-style spectra, with spectral resolution of FWHM = 6.7km/s and additive noise with signal to noise of 30 per 2.1 km/s pixel at the continuum.

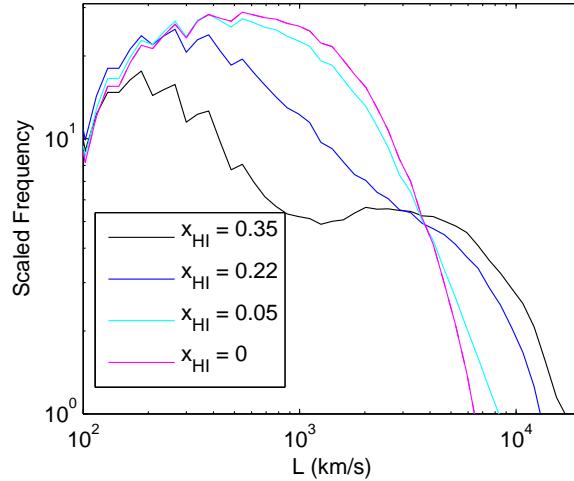


Figure 2.11: Mock dark gap size distribution. This figure is identical to Fig. 2.3 except that it uses spectra with spectral resolution FWHM = 100km/s, bin size $\Delta v_{\text{bin}} = 50\text{km/s}$, and a signal-to-noise ratio of 10 at the continuum. This figure shows the expected histogram of dark gap sizes using 20 spectra with $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), 0.05 (cyan), and 0 (magenta) at fixed $\langle F \rangle = 0.1$.

2.9 Conclusion

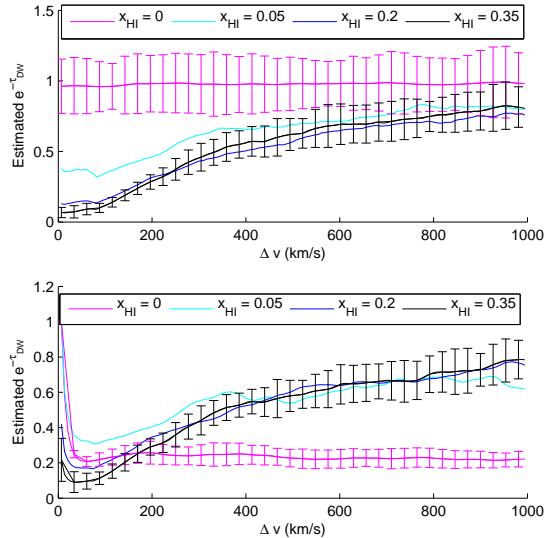


Figure 2.12: Using the Ly β forest to estimate damping-wing-less Ly α transmission. The above figure shows the estimated *shape* of stacked damping wing absorption for $\langle x_{HI} \rangle = 0$ (magenta), 0.05 (cyan), 0.22 (blue), and 0.35 (black). The curves have been normalized to have their mean values peak at 1. Additionally, we show error bars for the fully ionized case and $\langle x_{HI} \rangle = 0.35$ case which indicate the scatter in the curves between groups of 20 spectra. The left-hand plot is obtained by using a large ensemble of mock spectra to model a mapping between stacked Ly β transmission and stacked damping-wing-less Ly α transmission and then applying this to groups of 20 spectra. Meanwhile, the right-hand figure plots the ratio of the stacked Ly α flux to the stacked Ly β flux, providing a simplified estimate of the damping wing contribution to the absorption for each case.

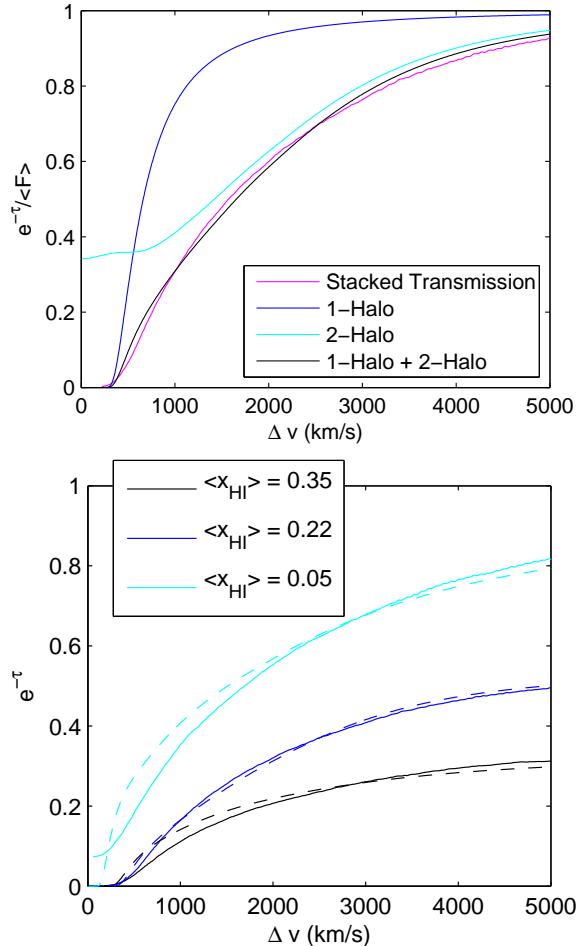


Figure 2.13: Model for the extended damping wing absorption. The left panel shows the components of our model for stacked transmission outside of a neutral region compared to the stacked transmission using mocked spectra (magenta) for $\langle x_{\text{HI}} \rangle = 0.22$. We show the absorption due to the central neutral region (blue), average absorption due to neighboring, clustered neutral regions (cyan), and the product of the two transmissions (black). These are denoted in the legend as “1-Halo”, “2-Halo”, and “1-Halo + 2-Halo” in analogy with the halo model. In the right-hand panel, we show the comparison between the modelled transmission (dashed) and transmission from stacked mocked spectra (solid) for $\langle x_{\text{HI}} \rangle = 0.35$ (black), 0.22 (blue), and 0.05 (cyan). The curves in the right-hand figure have been multiplied by the mean transmission (computed here ignoring resonant absorption for illustration). In this appendix, the stacking is done at the HI/HII boundaries and only damping wing absorption is incorporated to demonstrate the extended excess absorption owing to correlated neighboring systems.

Chapter 3

On Modelling and Measuring the Temperature of the $z \sim 5$ IGM

3.1 Introduction

The temperature of the low density intergalactic medium (IGM) after reionization retains information about when and how the gas was heated during the Epoch of Reionization (EoR) (e.g. ? ? ? ?). The temperature of the IGM in turn impacts the statistical properties of the Ly- α forest towards background quasars and so the absorption in the forest provides “fossil” evidence regarding the timing and nature of reionization. Scrutinized carefully, this fossil may therefore improve our understanding of reionization. For example, the IGM will likely be cooler at $z \sim 5$ if most of the IGM volume reionized at relatively high redshift, near e.g. $z \sim 10$, than if reionization happened later, near say $z \sim 6$. If reionization occurs early, the gas has longer to cool and reaches a lower temperature than if it happens late, at least provided the gas is heated to a fixed temperature at reionization. In addition, the IGM temperature should be inhomogeneous, partly as a result of spatial variations in the timing of reionization across the universe (? ? ?). Careful measurements of the IGM temperature after reionization should hence constrain the average reionization history of

3.1 Introduction

the universe, and may potentially reveal spatial variations around the average history as well.

Two separate phases of reionization are likely relevant for understanding the thermal history of the IGM: an early period of hydrogen reionization during which hydrogen is ionized, and helium is singly ionized by star-forming galaxies, and a later period in which helium is doubly-ionized by quasars, i.e. HeII reionization. Hydrogen reionization completed sometime before $z \sim 6$ or so (e.g. Fan et al. 17, although it might conceivably end as late as $z \sim 5$ – see Lidz et al. 28, McGreer et al. 30, Mesinger 34) , while mounting evidence suggests HeII reionization finished by $z \gtrsim 2.5 - 3$ (see e.g. ? ? and references therein). Many of the existing IGM temperature measurements focus on redshifts of $z \sim 2 - 4$ (? ? ? ? ? ?); in this case the temperature is likely strongly influenced by HeII reionization (e.g. ?) and so these measurements mostly constrain helium reionization rather than hydrogen reionization.

In order to best constrain hydrogen reionization using the thermal history of the IGM, temperature measurements at higher redshift ($z \gtrsim 5$) are required. Indeed, recent work has started to probe the temperature at these early times. In particular, the recent study by ?) includes a measurement at $z = 4.8$; ?) and ?) determine the $z \sim 6$ temperature in the special “proximity zone” region of the Ly- α forest close to the quasar itself; and the analysis in Viel et al. (50) starts to bound the $z \gtrsim 5$ IGM temperature, although these authors focus on placing limits on warm dark matter models.

The temperature at these higher redshifts is unlikely to be significantly impacted by HeII reionization. In addition, the “memory” of intergalactic gas to heating during the EoR gradually fades and so measurements as close as possible to the EoR should, in principle, be most constraining. It is not, however, obvious that the IGM temperature can be measured accurately enough from the $z \gtrsim 5$ Ly- α forest to exploit the sensitivity of the high redshift temperature to the properties of reionization. In particular, the forest is highly absorbed by $z \sim 5$ with $z \gtrsim 6$ spectra showing essentially complete Gunn-Peterson (23) absorption

3.1 Introduction

troughs (5, 17). An interesting question is then: what is the highest redshift at which it is feasible to measure the IGM temperature from the Ly- α forest?

Towards this end, the goal of this paper is to both model the thermal state of the $z \sim 5$ IGM, incorporating inhomogeneities in the hydrogen reionization process, and to quantify the prospects for actually measuring the IGM temperature using $z \gtrsim 5$ Ly- α forest absorption spectra. The outline of this paper is as follows. In §4.2.2, we describe the numerical simulations used in our analysis. In §3.3, we present plausible example models for the reionization history of the universe and describe our approach for modeling inhomogeneous reionization. We adopt a semi-analytic approach for modeling the resulting thermal history of the IGM, as described in §3.4. In this section, we also quantify the statistical properties of the temperature field in several simulated reionization models. Finally, in §3.5 we discuss how to measure the temperature from the $z \sim 5$ Ly- α forest, and forecast how well it may be measured with existing data. Our main conclusions are described in §3.6.

This work partly overlaps with previous work which also recognized the importance of, and modeled, temperature inhomogeneities in the $z \sim 5$ IGM and considered some of the observable implications (? ? ?).¹ One key difference with this earlier work is that we consider a more direct approach for measuring the temperature of the $z \sim 5$ IGM from the Ly- α forest. Our modeling of the thermal state of the IGM is closely related to that in ?, except that we implement a similar general approach using numerical simulations, which allow us to construct mock Ly- α forest spectra and to measure the detailed statistical properties of these spectra. The works of (?) use radiative transfer simulations to model hydrogen reionization and the thermal history of the IGM and so these authors track some of the underlying physics in more detail than we do here. However, our approach here is faster, simpler, and more flexible, while we believe that it nevertheless captures many of the important processes involved.

¹?) also considered temperature fluctuations from hydrogen reionization, but these authors focused on $z \sim 3$ where – as they discussed – these fluctuations should be small and swamped by effects from HeII reionization.

3.2 Simulations

Our analysis makes use of two different types of numerical simulations. First, we use the “semi-numeric” scheme of ?) to model reionization; this algorithm is performed on top of the dark matter simulation of McQuinn et al. (31). The McQuinn et al. (31) simulation tracks 1024^3 dark matter particles in a simulation volume with a co-moving sidelength of 130 Mpc/ h . Using the semi-numeric technique allows us to capture the impact of inhomogeneities in the reionization process, while providing the flexibility to explore a range of possible reionization models. In these models, we assume that the gas distribution closely traces the simulated dark matter distribution. We discuss the impact of this approximation when relevant. As we will describe, we map between the redshift of reionization of each gas element and its temperature at high redshift using the technique of ?); this mapping depends on the density of each gas element. We then produce mock Ly- α forest spectra, according to the usual “fluctuating Gunn-Peterson approximation” (e.g. Croft et al. 14, Miralda-Escude et al. 38) although here we additionally account for the temperature variations from inhomogeneous reionization.

We also make use of one of the smoothed particle hydrodynamic (SPH) simulations from ?). These simulations were run using the code Gadget-2 (?). This simulation tracks 2×1024^3 particles (with equal numbers of dark matter and baryonic particles) in a 50 Mpc/ h simulation box. In these calculations, we ignore the inhomogeneity of the reionization process. We use these simulations to more faithfully capture the gas distribution (for gas elements that reionize at a given time). In constructing mock Ly- α forest spectra from these simulations, we first modify the simulated gas temperatures, according to various prescriptions, in order to test how sensitive the statistical properties of the absorption are to the gas temperature.

3.3 Reionization Histories

In an effort to explore how the thermal state of the post-reionization IGM depends on the reionization history of the universe, we consider several example reionization histories. Our aim is to consider models that result in a wide range of possible thermal histories, while broadly maintaining consistency with current observational constraints on reionization.

For simplicity, we assume (as is common) that early galaxy populations produce ionizing photons at a rate that is directly proportional to the rate at which matter collapses into halos above some minimum mass. The minimum mass describes the host halo mass above which galaxies form readily; here we adopt $M_{\min} = 10^9 M_{\odot}$. We compute the collapse fraction from the Sheth-Tormen halo mass function (?). With these assumptions, the volume averaged ionization fraction ($\langle x_i \rangle$) obeys the following differential equation (? ?):

$$\frac{d\langle x_i \rangle}{dt} = \zeta \frac{df_{\text{coll}}}{dt} - \frac{\langle x_i \rangle}{\bar{t}_{\text{rec}}}. \quad (3.1)$$

The first term on the right hand side of the equation describes the rate at which neutral atoms are ionized, while the second term on the right hand side accounts for ionized atoms that recombine. The recombination time (\bar{t}_{rec}) depends on the clumpiness of the IGM, parametrized by a “clumping factor”, $C = \langle n_e^2 \rangle / \langle n_e \rangle^2$, and the temperature of the IGM. In solving this equation – and for this purpose only – we assume an isothermal IGM. We approximate the clumping factor and the temperature as independent of redshift. Adopting the case-B recombination rate in solving this equation, a temperature of $T_0 = 2 \times 10^4$ K, and $C = 3$ (see e.g. ? ? for a discussion regarding plausible values of the clumping factor) gives

$$\bar{t}_{\text{rec}} = 0.93 \text{Gyr} \left[\frac{3}{C} \right] \left[\frac{1+z}{7} \right]^{-3} \left[\frac{T_0}{2 \times 10^4 K} \right]^{0.7}. \quad (3.2)$$

Solving the differential equation, Eq. 3.1, suffices to compute the average ionization fraction as a function of redshift, given the minimum mass and efficiency, ζ , of the ionizing sources.

3.3 Reionization Histories

In order to model reionization inhomogeneities, we use the “semi-numeric” scheme of ?), which is based on the excursion set model of reionization developed in Furlanetto et al. (21). This scheme captures the tendency for halos – and hence galaxies – to form first in regions that are overdense on large scales, and to reionize before more typical locations in the universe. In the simplest version of the semi-numeric scheme, recombinations are considered only in an average sense and are treated as simply reducing the overall efficiency at which atoms are ionized. Let us denote the resulting efficiency factor as $\tilde{\zeta}(z)$ to distinguish it from the above ionizing efficiency factor ζ . As we explain subsequently, we allow this efficiency factor to be redshift dependent. We can then consider the condition that a region of co-moving size R is ionized. In the initial conditions, the mass enclosed within this co-moving region is $M = 4\pi R^3 \langle \rho_M \rangle / 3$, with $\langle \rho_M \rangle$ denoting the average matter density per co-moving volume. The condition for this region to be ionized is then:

$$\tilde{\zeta}(z) f_{\text{coll}}(> M_{\min} | \delta_M, M) \geq 1. \quad (3.3)$$

In this equation $f_{\text{coll}}(> M_{\min} | \delta_M, M)$ is the conditional collapse fraction, i.e., the fraction of mass in halos above the minimum mass (M_{\min}) in a region of linear overdensity δ_M . Here δ_M denotes the overdensity when the linear density field is smoothed on mass scale M .

In order to tabulate a reionization redshift for many grid cells across the volume of our simulation, we smooth the density field – linearly evolved from the initial conditions – on a range of mass scales, starting from large scales and stepping downward until we reach the scale of each simulation cell. For each simulation cell, and across all smoothing scales considered, we record the highest redshift at which the ionization barrier (Eq. 3.3) is crossed. This highest crossing redshift is considered to be the reionization redshift, z_r , for the cell in question. We tabulate reionization redshifts for each of 512^3 grid cells. This provides us with a reasonable model for the expected spatial variations in the redshift of reionization – and the coherence scale of these inhomogeneities – across the simulation volume.

3.3 Reionization Histories

Note that here we approximate the excursion set model as determining the redshift at which each *volume element* is reionized, although in reality mass elements move from their initial positions, and overdense regions expand less rapidly than typical locations. This approximation is commonly made, and is reasonable given the large size of the ionized regions (21) and the correspondingly large coherence scale of the spatial variations in the reionization redshift.

Another ingredient we use from the McQuinn et al. (31) simulation is the evolved non-linear dark matter density field, interpolated onto the same grid (using CIC interpolation) as the reionization redshifts. For our calculations with this simulation, we generally assume that the gas distribution follows the simulated, gridded dark matter distribution. Note that the smoothing introduced by gridding the dark matter particles is comparable to the Jeans smoothing scale: the co-moving Jeans wavenumber for isothermal gas at 10^4 K, is $k_J = 13h \text{ Mpc}^{-1}$ at $z = 5$ which is comparable to the Nyquist wavenumber of the grid, $k_{\text{Nyq}} = 12h \text{ Mpc}^{-1}$. More relevant, however, is the “filtering scale” – essentially a time-averaged Jeans scale – and this should be smaller than the Jeans scale by around a factor of a few (?). In any case a single global smoothing only roughly approximates the full effect of Jeans smoothing. We will return to discuss this further in §3.5 and §3.5.3. In particular, in order to approximately capture the impact of small scale structure and thermal broadening in our mock quasar spectra, we will add small-scale structure using a lognormal model. Although using the gridded dark matter density field to represent the gas distribution is inadequate for making detailed mock spectra, it suffices for our model of the temperature distribution of the low density gas.

Returning to further consider the semi-numeric modeling, an important caveat is that this algorithm does not return precisely the expected volume-averaged ionization fraction (see the Appendix of ? for a discussion). Here we simply tune $\tilde{\zeta}(z)$ to produce the desired redshift evolution of the ionization fraction. Although this procedure is not ideal, small adjustments to the ionizing efficiency factor have little impact on the size of the

3.3 Reionization Histories

ionized regions at a given volume-averaged ionization fraction, $\langle x_i \rangle$, and so this approach is adequate for our present purposes.

The redshift evolution of the volume-averaged ionization fractions are shown in Fig. 3.1 for three example models. The symbols show the average ionized fraction from the simulation cube at different redshifts. We call the three examples in the figure the “Low-z” model, the “Mid-z” model, and the “High-z” model. The Mid-z model adopts a redshift dependent efficiency factor of the form $\tilde{\zeta}(z) = 35(1 + z/13)^{1.75}$ for $z \leq 12$ and $\tilde{\zeta} = 35$ for $z > 12$. For comparison, the black dashed line shows the solution to Eq. 3.1 for a model with $\zeta = 46$, $C = 3$, and $M_{\min} = 10^9 M_\odot$. Hence the semi-numeric scheme in the Mid-z model has been tuned to return the ionized fraction expected from Eq. 3.1 for a plausible model. The Low-z and High-z models are similar to the Mid-z model, except that the efficiency factor in the semi-numeric models has been adjusted to $\tilde{\zeta}(z) = 12(1 + z/11)^{0.60}$ at $z \leq 10$ and to $\tilde{\zeta}(z) = 12$ at $z > 10$ for the Low-z model and to a constant efficiency factor $\tilde{\zeta}(z) = 70$ for the High-z model. (Although these alternative models were not themselves explicitly tuned to match particular solutions to Eq. 3.1, the general behavior is similar to in the Mid-z model except that reionization happens a little later/earlier in the Low-z/High-z model and so these models also appear reasonable).

It is also useful to quantify the timing and duration of reionization, as well as the optical depth to Thomson scattering (τ_e), in each model. Defining the “completion” of reionization as the redshift where the volume averaged ionization fraction first reaches $\langle x_i \rangle = 1$, the High-z model completes at $z = 9.6$, the Mid-z model at $z = 6.7$, and the Low-z model at $z = 5.8$. As one measure of the “duration” of reionization, we consider the redshift spread over which $\langle x_i \rangle$ evolves from $\langle x_i \rangle = 0.1$ to $\langle x_i \rangle = 1$. This duration is $\Delta z = 2.7, 4.3, 3.5$ for the High-z, Mid-z, and Low-z models. Note that the duration is the longest in the Mid-z model because the ionizing efficiency factor $\tilde{\zeta}(z)$ has the strongest redshift dependence in this case. The electron scattering optical depths are $\tau_e = 0.088, 0.066, 0.052$ for the High-z, Mid-z, and Low-z reionization models respectively. These values assume that the fraction

3.3 Reionization Histories

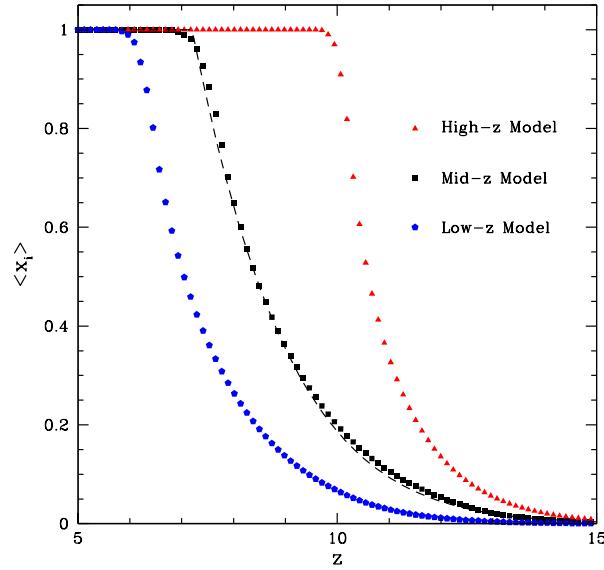


Figure 3.1: Example reionization histories. The red triangles show the simulated volume-average ionization fraction in our semi-numeric High- z reionization model, the black squares are for the Mid- z reionization scenario, and the blue pentagons are for a low redshift (Low- z) reionization model. The black dashed line shows the reionization history computed by solving Eq. 3.1 with $\zeta = 46$, $M_{\min} = 10^9 M_\odot$ and $C = 3$. The semi-numeric efficiency parameters $\tilde{\zeta}(z)$ in the Mid- z case have been tuned to match this model.

3.4 The Thermal State of the IGM

of helium that is singly ionized is identical to the fraction of hydrogen that is ionized, and ignore the slight boost expected from the free electrons produced after HeII reionization, which we do not track in this work.

The present constraint on τ_e from Planck CMB temperature anisotropy data (1), combined with the E-mode polarization power spectrum at large angular scales from WMAP nine year data (?), is $\tau_e = 0.089^{+0.012}_{-0.014}$. Most of the constraining power here comes from the WMAP polarization data. Hence our High-z model produces a τ_e close to the presently preferred value, the Mid-z model is low by $1.6 - \sigma$, while the Low-z model is too low by $2.6 - \sigma$. Hence our lower redshift reionization models are already marginally disfavored, but they are still certainly worthy of further investigation. The Planck collaboration should soon announce new large-scale CMB polarization measurements; the improved frequency coverage of the Planck satellite should help guard against foreground contamination, and further test these models for τ_e . Although the current τ_e constraints allow higher reionization redshift models than the three examples considered here, the $z \sim 5$ IGM temperature is insensitive to the reionization redshift if reionization happens above $z \gtrsim 10$ (§3.4, ?). While viable, we need not consider such models explicitly here since in these cases the $z \sim 5$ temperature will be similar to that in our High-z model.

3.4 The Thermal State of the IGM

We now explore how the thermal state of the $z \sim 4 - 6$ IGM depends on the reionization history of the Universe, using the example histories of the previous section. In this section, we focus mostly on the Low-z and High-z models since they span a fairly wide range of possibilities for the IGM temperature at the redshifts of interest.

The key equation describing the thermal evolution of a gas element in the IGM is (e.g.

3.4 The Thermal State of the IGM

?):

$$\begin{aligned} \frac{dT}{dt} = & -2HT + \frac{2T}{3(1+\delta)} \frac{d\delta}{dt} + \frac{T}{\mu} \frac{d\mu}{dt} \\ & + \frac{2\mu m_p}{3\rho k_B} (\mathcal{H} - \Lambda). \end{aligned} \quad (3.4)$$

The first term on the right hand side accounts for adiabatic cooling owing to the overall expansion of the universe. The second term describes adiabatic heating/cooling from structure formation, i.e. from gas elements contracting/expanding. In the third and fourth terms, μ is the mean mass per free particle in the gas, in units of the proton mass. The third term accounts for the temperature change that occurs because the mean mass per particle may change with time. \mathcal{H} describes the heating term, while Λ is the cooling function of the gas. These terms describe the heat gain and loss per unit volume, per unit time, in the gas.

Let us first summarize the qualitative behavior of the solutions to Eq. 3.4, focusing on the low density intergalactic gas that fills most of the volume of the universe (see also ? ? ? ?). During reionization, most gas elements are rapidly ionized and change their ionization fraction by order unity.¹ The excess energy of the ionizing photons (above the ionization threshold) goes into the kinetic energy of the outgoing electrons, which quickly share their energy with the surrounding gas, and raise its temperature. The first thing to consider is hence the initial temperature reached at reionization. Provided the gas becomes highly ionized, its temperature boost during reionization depends only on the *shape* of the spectrum – and not the amplitude – of the radiation that ionized it.

In detail, we expect gas elements to be ionized by radiation with a range of spectral shapes. This should be the case both because the intrinsic ionizing spectrum will vary from galaxy to galaxy, and because the spectral shape may be hardened by intervening absorption, which will itself vary spatially depending on the column density of neutral

¹Sufficiently overdense regions/clumps may be only gradually ionized as reionization proceeds and the ionizing radiation field incident upon them grows in intensity, but we will neglect these, assuming that partly neutral clumps fill only a small fraction of the volume within mostly ionized regions.

3.4 The Thermal State of the IGM

gas between an ionizing source and an absorber. On the other hand, the ionized regions during hydrogen reionization likely grow under the collective influence of numerous (yet individually faint) dwarf galaxies (e.g. Robertson et al. 43), and so some of these variations may average down, provided gas elements are ionized by a combination of several sources and the ionizing radiation arrives along various different pathways. In any case, modeling the precise temperature input during reionization and its spatial variations requires full radiative transfer simulations and is well beyond the scope of our approach here. We adopt this uniform temperature boost approximation throughout, and discuss plausible values for the input temperature subsequently. Note also that in this case the temperature boost during reionization is independent of density: extra heat is put into the overdense regions since more atoms are ionized in these regions, but the heat is shared across the larger number of particles in the overdense parcel.

After a gas element is reionized, it settles into ionization equilibrium and the UV radiation from the ionizing sources keeps the gas highly ionized (at least for the low density gas parcels that fill most of the volume of the IGM). In ionization equilibrium, each recombination is balanced by a photoionization and the ionizations in turn heat the gas; the average time between recombinations in the low density IGM is long, and so the heat input from photoionization is significantly reduced after a parcel becomes highly ionized during reionization. In addition, the spectral shape of the ionizing radiation incident on a typical gas element should soften – i.e., the average heat input per photoionization should drop – after reionization as the optical depth to ionizing photons decreases (?).¹

The dominant cooling processes are adiabatic cooling from the expansion of low density gas parcels and Compton cooling off of the CMB. As a result of cooling, although gas elements that reionize at the same time start off with identical temperatures, irrespective of their density, parcels with differing densities will not stay at the same temperature. In

¹In reality, the spectral softening depends on how progressed reionization is *globally* since the hardening from absorption depends on the density and ionization state of all of the gas between a source and an absorber. Here we neglect this by fixing the spectral shape incident on each gas element after it is ionized.

3.4 The Thermal State of the IGM

particular, the low density elements expand and cool faster than typical regions, while overdense regions recombine faster and thus – in ionization equilibrium – gain more heat from photoionizations after reionization. In addition, sufficiently overdense regions will be heated by adiabatic contraction. ?) showed that this competition between adiabatic cooling/heating, Compton cooling, and photoionization heating, drives the intergalactic gas to generally land on a tight temperature-density relation of the form $T = T_0(1+\delta)^{\gamma-1}$. Both the temperature at mean density, T_0 , and the slope of the temperature density relation, γ , depend on the reionization redshift; T_0 falls off and γ becomes steeper as the gas cools after reionization, until the gas gradually loses memory of the heating during reionization. In the previous work of ?), however, all of the gas was assumed to reionize at the same redshift. Here we would like to generalize this to incorporate spatial variations in the redshift of reionization (see also ? and ?).

3.4.1 Modeling the Thermal State

In general, to follow the thermal evolution in Eq. 3.4 we should combine this equation with equations specifying the evolution of the ionized fraction of each different particle species. However, for our present application a simpler approach should suffice. In particular, we start by assuming that each parcel is heated to a common temperature, T_r , at its reionization redshift, z_r . We then follow the subsequent thermal evolution after a gas element is ionized by assuming ionization equilibrium and that each element is highly ionized (as in ?). More specifically, we assume that both HI and HeI are highly ionized, but that HeII is not yet ionized, i.e., that HeII reionization starts later than the high redshifts of interest for our study. We further assume the gas is composed of only hydrogen and helium, neglecting metal line cooling, and also molecular hydrogen cooling, which should be very good approximations for the low density IGM. In addition to adiabatic heating/cooling and Compton cooling, we track HI photoheating, HeI photoheating, and recombination cooling of HII/HeII using the rate expressions in the Appendix of ?). We ignore collisional

3.4 The Thermal State of the IGM

ionizations, and HI/HeI/HeII line excitation cooling: neglecting these processes should be a good approximation for the low density and highly ionized gas that fills most of the IGM volume after reionization. Furthermore, we ignore other potential heating sources such as shock heating, galactic winds, blazar heating, etc..(see e.g. ? ? and references therein for a discussion). In the Appendix we also derive approximate solutions using linear perturbation theory (incorporating only HI photoheating, Compton cooling, and adiabatic cooling/heating), that are useful for fast and fairly accurate estimates (see also ?).

In principle we could calculate the adiabatic expansion/contraction term (second term in Eq. 3.4) directly from the McQuinn et al. (31) simulation, at least under the approximation that the gas distribution traces the simulated dark matter density field. Here we instead follow the approach of ?) and compute this term for tracer elements assuming their density evolution obeys the Zel'dovich approximation (?). As mentioned earlier, we do, however, extend this calculation to consider gas elements with a range of different reionization redshifts.

The basic premise here is that gas elements with identical reionization redshifts should land on a well-defined temperature-density relation (as supported by the tests in ? and subsequent work); we can determine this relation by solving Eq. 3.4 for many sample gas parcels. Incorporating, however, the spread in reionization redshifts, and that the reionization redshift of each parcel may correlate with its density, a perfect temperature-density relation will *not* generally be a good description. In other words, we follow sample gas parcels to determine the mapping between the temperature-density relation at a given redshift and the reionization redshift and temperature, i.e., this is used to determine $T_0(z|z_r, T_r)$ and $\gamma(z|z_r, T_r)$. These mappings can then be applied to our reionization simulation to determine the temperature of any gas element, given its reionization redshift and overdensity. To determine $T_0(z|z_r, T_r)$ and $\gamma(z|z_r, T_r)$, we follow the thermal evolution for 20,000 tracer elements for many different reionization redshifts, assuming their density evolves according to the Zel'dovich approximation, and fit separate power-laws for each z , z_r , and T_r .

3.4 The Thermal State of the IGM

In the Zel'dovich approximation, the density field evolves according to the equation:

$$1 + \delta = \frac{1}{\det[\delta_{ij} + D(t)\psi_{ij}]}, \quad (3.5)$$

with ψ_{ij} denoting the initial deformation tensor, and $D(t)$ denoting the linear growth factor (normalized to unity today). The density evolution of a tracer element can then be specified by the eigenvalues of the local initial deformation tensor. As in [?] and [?], we can construct realizations of the density evolution in the Zel'dovich approximation by randomly drawing eigenvalues of ψ_{ij} from the expected probability distribution [?]. We do this following [?].

In our fiducial model, we take the temperature at reionization to be $T_r = 2 \times 10^4$ K (see [? ?] for a discussion of this choice). In calculating the photoionization heating term after a gas parcel reionizes, we assume that the specific intensity of the ionizing radiation is a power-law in frequency close to the hydrogen photoionization threshold, $J(\nu) \propto \nu^{-\alpha}$ with $\alpha = 1.5$. As mentioned previously, the heat input is insensitive to the amplitude of the ionizing radiation, provided the gas is highly ionized. This spectral shape is intended to be somewhat harder than expected for the *intrinsic spectrum* of the ionizing sources, since intervening absorption will harden this spectrum [? ? ?].

3.4.2 Simulated Temperature Field

We now examine the properties of the simulated temperature field, modeled as described in the previous section. First, we consider the mapping between the temperature at mean density, T_0 , and the slope of the temperature-density relation, γ , for gas at various redshifts, given the reionization redshift, z_r , of each gas element. This is shown, for our baseline set of assumptions, in Fig. 3.2 for each of $z = 4.5, 5.0$, and $z = 5.5$. The values of T_0 are close to the temperature at reionization ($T_r = 2 \times 10^4$ K) for gas elements that ionized at redshifts just above $z = 5.5$, since these elements have had very little time to cool. On the other hand, gas parcels with higher z_r have had longer to cool and are hence at lower

3.4 The Thermal State of the IGM

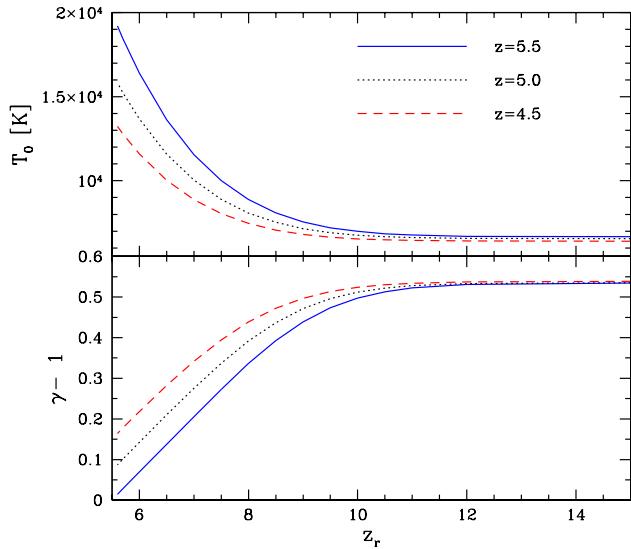


Figure 3.2: Thermal state of gas elements with a given reionization redshift, as a function of that redshift. In each case, the gas elements are heated to a temperature of $T_r = 2 \times 10^4$ K during reionization, and the residual photo-heating after reionization is computed assuming that the (hardened) spectral index of the ionizing sources is $\alpha = 1.5$ near the HI photoionization edge. *Top panel:* The temperature at mean density (T_0) for gas elements at each of $z = 4.5, 5.0$ and 5.5 as a function of their reionization redshift. *Bottom panel:* This is similar to the top panel, except it shows the slope of the temperature-density relation ($\gamma - 1$) rather than T_0 . Note that although we assume that gas elements with a given reionization redshift all land on a well defined temperature-density relation, this will not generally be a good description once we account for the spread in reionization redshift across the universe.

3.4 The Thermal State of the IGM

temperatures. For instance, gas elements that reionized at $z_r = 8$ have cooled down to $T_0 = 8,800$ K by $z = 5.5$, more than a factor of two below the temperature at reionization. The temperatures of gas elements that reionize at sufficiently high redshift, however, become insensitive to the precise redshift of reionization. In particular, gas elements that reionize above $z_r \gtrsim 10$ are all at $T_0 = 6,700$ K at $z = 5.5$, irrespective of z_r . This results mainly because Compton cooling is very efficient at high redshift ($z \gtrsim 10$), and effectively erases the memory of the photoheating during reionization (?). Indeed, this is the main reason that we don't consider still higher redshift reionization models, although they would be allowed by the present τ_e constraints as discussed in §3.3 (but perhaps disfavored by other data sets, see e.g. Robertson et al. 43? for recent summaries.): the thermal state of the IGM is insensitive to higher redshift reionization models.

The bottom panel is similar to the top panel except here we plot $\gamma - 1$ versus z_r . Gas elements that reionize just above $z = 5.5$ are close to isothermal, while elements that ionize at $z_r \gtrsim 10$ have a steeper slope, $\gamma = 1.53$. The T_0 and $\gamma - 1$ curves at $z = 4.5$ and $z = 5.0$ illustrate less sensitivity to z_r , since gas elements at these redshifts have had longer to cool down from their initial temperatures at reionization. Nevertheless, the models at these lower redshifts still certainly do show some dependence on z_r .

We then use the curves plotted in Fig. 3.2 as a mapping to predict the temperature of various grid cells in our simulation given their overdensities, δ , and reionization redshifts, z_r . This procedure allows us to model the temperature field across the entire simulation volume at various redshifts.

Figs. 3.3 and 3.4 show the result of applying the mapping (at $z = 5.5$) to the simulated density field.¹ Specifically, these figures show thin slices (0.25 Mpc/ h thick) through the

¹In practice, we apply the mapping to the simulated density field at slightly higher redshift ($z = 6.9$) since we don't currently have outputs from this simulation at the lower redshifts of interest. Using the higher redshift output artificially reduces the variance in the density field, and the resulting structure in the temperature field somewhat. For our present purposes, this is not important. The main effect of boosting the density variance should be to increase the minimum and maximum density contrasts shown in scatter

3.4 The Thermal State of the IGM

simulation volume, with the top panel showing the reionization redshift and the bottom panel the corresponding temperature of cells in the simulation volume. In the Low-z model (Fig. 3.3), the temperature field has sizable spatial variations on large scales. As anticipated earlier, these result because of the spread in the timing of reionization across the universe. As one can infer from the slice, the regions that are at low-density (when the density field is smoothed on large-scales) – i.e., the “voids” in the density distribution – are the last to reionize. These regions are at the highest temperature shortly after reionization because they have had the least amount of time to cool (see also ??).

In contrast, the temperature field in the High-z model (Fig. 3.4) has mostly lost memory of the heating during reionization and so the temperature variations are more subtle here. This is expected from Figs. 3.1 and 3.2: much of the gas in this model is reionized at $z_r \gtrsim 10$, and efficient Compton cooling mostly removes the memory of reionization in this case. The temperature variations that are apparent in the High-z model instead reflect the usual temperature-density relation, as the competition between cooling and heating after reionization drives overdense regions to larger temperatures. These temperature variations are primarily coherent on the Jeans/filtering scale and so, as evident from the simulation slices, these fluctuations are concentrated mostly on smaller scales than the ones induced by the spread in the timing of reionization.

A further, more quantitative description is provided by constructing scatter plots of the temperatures of many simulated gas elements as a function of their densities. This is shown in Figs. 3.5 and 3.6 for the Low-z and High-z reionization models, respectively. Broadly similar results may be found in earlier work by ?? and ??). The red short-dashed line in each figure shows the median gas temperature at $z = 5.5$, while the green long-dashed line is the median temperature at $z = 4.5$. Fig. 3.5 shows that the temperature of the $z = 5.5$ IGM is generally rather high – and has a large amount of scatter at low densities

plots such as Fig. 3.5. Importantly, this has little impact on the median temperature-density relation and the scatter around this relation for the range of density contrasts in our scatter plots. We have tested this explicitly using a lognormal approximation to the density field at $z = 4.5$ and $z = 5.5$.

3.4 The Thermal State of the IGM

– in the Low- z reionization model. By contrast, the temperature in the High- z reionization model (Fig. 3.6) is smaller – e.g., by 60% for the median temperature near the cosmic mean density – as is the scatter. In the Low- z model the median temperature is a fairly flat function of density for gas less dense than the cosmic mean.

Note that although the regions that have low density – when the density field is averaged on large scales – ionize last and are mostly hotter than denser regions (see Fig. 3.3), this does not fully “invert” the temperature-density relation. This is because the density field on the scale of the simulation grid (and at the Jeans scale) is only somewhat correlated with the larger scale density variations that determine the spread in the timing of reionization in our model. In any case, in agreement with previous work (? ?), the usual temperature-density relation is a poor description of the thermal state of the IGM in the Low- z model at $z = 5.5$. At slightly lower redshifts, $z = 4.5$, the temperature has dropped somewhat and the scatter in the Low- z model has partially subsided, although it is still substantial. The median temperatures in the two reionization models are closer to each other by $z = 4.5$, but they still differ by 30%.

It is also useful to calculate the power spectrum of temperature fluctuations in each model. Since we are assigning a value of T_0 and γ to each grid cell in the simulation volume, we can easily consider the power spectrum of T_0 rather than the power spectrum of the full temperature field. The advantage of considering the power spectrum of T_0 is that this power spectrum vanishes in the case of homogeneous reionization. In the case of homogeneous reionization, the temperature is a power-law in the gas density, and so the full temperature field still has (mostly small scale) fluctuations sourced directly by density inhomogeneities. Hence we consider here the power spectrum of $T_0(\mathbf{x})$, or more precisely the power spectrum of $\delta_{T_0}(\mathbf{x}) = (T_0(\mathbf{x}) - \langle T_0 \rangle)/\langle T_0 \rangle$. Although the power spectrum of this field is not directly observable, it nevertheless helps to characterize the temperature fluctuations from reionization.

3.4 The Thermal State of the IGM

The power spectra in some of our models are shown in Fig. 3.7. Specifically, the curves show $\Delta_{T_0}^2 = k^3 P_{T_0}(k)/(2\pi^2)$, the contribution to the variance of δ_{T_0} per natural logarithmic interval in k , i.e., per $d\ln(k)$. In the Low-z model, the temperature fluctuations peak at a level of around $\sqrt{\Delta_{T_0}^2(k)} \approx 15\%$. We should keep in mind, however, that the scatter in the temperature at lower density is larger than at mean density (see Fig. 3.5). As a result, the power spectra of T_0 shown here hence do not fully capture the impact of inhomogeneous reionization, but they do nevertheless illustrate the spatial scale of the reionization induced inhomogeneities as well as their redshift and model dependence. The power spectra ($\Delta_{T_0}^2$) are evidently fairly flat functions of k . This is not surprising, since the power spectra of the fluctuations in the ionization field are also rather flat functions of k during most of the EoR (e.g. McQuinn et al. 32). At the same redshift, the δ_{T_0} power spectrum in the Mid-z Model is ≈ 3 times smaller in amplitude than in the Low-z model, while the amplitude of variations ($\Delta_{T_0}^2(k)$) in the High-z model are ≈ 300 times smaller than in the Low-z model. As discussed previously, the small fluctuations in the High-z model result because Compton cooling is efficient at high redshift and this rapid cooling effectively erases the memory of heating at higher redshifts. Comparing the black solid and dashed lines illustrate how the fluctuations fade from $z = 5.5$ to $z = 4.5$ in the Mid-z model.

These models illustrate the dependence of the thermal history of the IGM on the timing of reionization; let us briefly summarize our main findings here. The IGM temperature for models in which a significant fraction of the IGM volume is reionized at relatively low redshift, near $z \sim 6$, is correspondingly larger than if most of the gas is reionized at higher redshift. In addition, the late reionization models produce sizable temperature inhomogeneities with fluctuations on scales as large as \sim tens of co-moving Mpc.

3.4.3 Variations around Fiducial Parameters

Before we proceed to discuss the observable signatures of the IGM temperature models, it is interesting to consider how variations around our fiducial assumptions regarding the

3.4 The Thermal State of the IGM

reionization temperature, T_r , and the shape of the ionizing spectrum after reionization might impact the resulting thermal state of the IGM. To investigate this, we consider models where the reionization temperature is $T_r = 3 \times 10^4$ K and $T_r = 1 \times 10^4$ K to contrast with our fiducial model in which $T_r = 2 \times 10^4$ K. The low T_r model requires sources with extremely soft ionizing spectra, and is meant to represent a lower limit to the plausible reionization temperature, while the higher temperature $T_r = 3 \times 10^4$ K case is more reasonable (e.g. ?). In addition, for our fiducial reionization temperature we produce models with (post reionization) spectral shapes of $\alpha = 0.5$ and $\alpha = 2.5$ to compare with our baseline assumption of $\alpha = 1.5$.

First, we consider how our models for $T_0(z = 5.5|z_r)$ and $\gamma(z = 5.5|z_r)$ depend on the reionization temperature and spectral shape, i.e., we regenerate the models of Fig. 3.2 for different values of T_r and α . The results of these calculations are shown in Fig. 3.8. The first feature to note is that T_0 and γ are independent of z_r and T_r in the limit of large reionization redshift: efficient cooling wipes out the memory of the early heating history. On the other hand, the $z = 5.5$ temperature is naturally quite sensitive to the reionization temperature if reionization occurred relatively recently. One consequence of this is that the *scatter* in the $z \sim 5$ temperature will be larger in low redshift reionization models for cases with larger reionization temperatures: a high reionization temperature increases the temperature contrast between recently reionized gas parcels and those that reionized early. The increased scatter in these models may potentially boost the observability of the temperature inhomogeneities induced by spatial variations in the timing of reionization, as we explore subsequently.

Another important point is that increasing T_r in a high reionization redshift model will not help to mimic the $z \sim 5$ temperature in a lower redshift reionization model, since the gas that reionized at high redshift reaches an asymptotic temperature that is insensitive to T_r . On the other hand, decreasing T_r (as in the $T_r = 10^4$ K curves) in a low reionization redshift model will certainly diminish the distinction between this model and higher reionization

3.4 The Thermal State of the IGM

redshift models. As we will see, however, the larger scatter and flatter trend of temperature with density in the low z_r , low T_r model offer potential handles for distinguishing between these models and higher reionization redshift scenarios.

Next we consider how the results vary with changes in the spectral shape, α , as shown in the figure. These variations have a relatively minor effect. Adopting a harder ionizing spectrum (smaller α) after reionization increases the amount of residual late-time photoheating. This thereby raises the asymptotic temperature and the asymptotic value is reached earlier. Quantitatively, the asymptotic temperature is 20% higher in the $\alpha = 0.5$ case than in our fiducial $\alpha = 1.5$ model, and 18% smaller for $\alpha = 2.5$. The dependence on α is relatively mild compared to other uncertainties in our modeling and so we don't consider it further here.

Another perspective is to construct scatter plots in the temperature-density plane and plot median temperature-density relations for various models, as in Fig. 3.5 and Fig. 3.6. This is shown for gas at $z = 5.5$ in Fig. 3.9. Consider first the two High-z models (the bottom two sets of points and dashed lines in the figure), which show the results of assuming $T_r = 2 \times 10^4$ K (bottom-most case with red points and a black dashed line), and $T_r = 3 \times 10^4$ K (shown as blue points and a cyan line, just above the bottom-most model). This shows that the results in this model are insensitive to T_r : this is as anticipated from Fig. 3.8. The next model is a Low-z case and has $T_r = 1 \times 10^4$ K (green points and blue line). This model is clearly closer to the High-z reionization model than our fiducial Low-z case, which is shown in the figure as the upper most black points with red line fit. However, the median temperature at low density and the scatter in the temperature are both larger in the Low-z, low reionization temperature model than in the High-z models. If the scatter in the temperature, and the trend of temperature with density can be measured observationally, this may help break the partial degeneracies between reionization redshift and temperature.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

We now turn to consider the impact of the thermal state of the IGM on the properties of the $z \sim 5$ Ly- α forest, and on the possibility of extracting these signatures to learn about reionization. The effects of temperature on the statistics of the Ly- α forest are discussed, for example, in [?]). The three main effects are: higher temperatures produce more Doppler broadening; the recombination rate of the absorbing gas is temperature dependent with hotter gas recombining more slowly, leading to less neutral gas and less absorption; hotter gas leads to more Jeans smoothing, with the precise impact of this smoothing depending on the entire prior thermal history of the absorbing gas (?)).

The enhanced Doppler broadening and Jeans smoothing in models with high temperatures each act to reduce the amount of small-scale structure in the Ly- α forest. These two effects are not, however, entirely degenerate: Jeans smoothing filters the gas distribution in three dimensions, while Doppler broadening smooths the optical depth field along the line of sight (e.g. [?]). Previous studies suggest that Doppler broadening impacts the small scale structure in the forest more than Jeans smoothing, at least near $z \sim 3$ ([? ? ?]). At $z \sim 5$, we expect Jeans smoothing to have more impact, however: the high opacity in the Ly- α line at these redshifts implies that even slight density enhancements can give rise to noticeable absorption lines, and these slight density variations may be erased by Jeans smoothing. Unfortunately, it is challenging to model the impact of Jeans smoothing while incorporating a realistic model for inhomogeneous reionization and photoheating. This requires hydrodynamic models that resolve the filtering scale, while capturing a large enough volume to model patchy reionization. Furthermore, the filtering scale depends on the entire prior thermal history. In this paper, we defer this challenge to future work and assume that the effect of Jeans smoothing is sub-dominant to that of Doppler broadening. We caution that Jeans smoothing might, however, enhance the impact of patchy reionization and modeling it may be necessary to robustly interpret future measurements.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

In any case, the small-scale structure in the Ly- α forest should be sensitive to the thermal state and the thermal history of the IGM, and so we now consider an approach for estimating the amplitude of small-scale structure in the forest. Here we will use the basic technique described in [?], except applied here to simulated data at higher redshift where there is more absorption in the forest. The first issue we aim to explore here is to what extent the temperature of the IGM is measurable at higher redshift, where the forest is significantly more absorbed. A second goal is to explore the impact of the temperature inhomogeneities modeled in the previous section.

We briefly outline the approach of [?] for measuring the small-scale structure – and thereby extracting constraints on the IGM temperature – here for completeness. In this approach, each spectrum is convolved with a Morlet wavelet filter and the smoothing scale of this filter is tuned to extract the amplitude of the small scale power spectrum in the forest as a function of position across each spectrum. The Morlet filter is a plane wave, multiplied by a Gaussian and in configuration space may be written as:

$$\Psi_n(x) = K \exp(i k_0 x) \exp\left[-\frac{x^2}{2s_n^2}\right]. \quad (3.6)$$

The Fourier space counterpart, when the normalization constant K is fixed so the filter has unit power (see [?]) is:

$$\Psi_n(k) = \pi^{-1/4} \sqrt{\frac{2\pi s_n}{\Delta u}} \exp\left[-\frac{(k - k_0)^2 s_n^2}{2}\right]. \quad (3.7)$$

In the above equation, Δu is the size of each spectral pixel in velocity units and s_n is a suitable smoothing scale (also in velocity units) chosen to extract the small scale power, and we set $k_0 s_n = 6$ (see [?]). Each mock spectrum is convolved with the above filter. We work with the transmission fluctuation field, $\delta_F(x) = (F(x) - \langle F \rangle)/\langle F \rangle$ where $F = e^{-\tau}$ is the transmission and $\langle F \rangle$ is the ensemble-averaged mean transmitted flux.

The transmission fluctuation, convolved with the wavelet filter, is:

$$a_n(x) = \int dx' \Psi_n(x - x') \delta_F(x'), \quad (3.8)$$

3.5 Measuring the Temperature of the $z \sim 5$ IGM

The amplitude of this filtered field, at position “ x ” is given by

$$A(x) = |a_n(x)|^2, \quad (3.9)$$

and characterizes the amount of small-scale structure in the transmission field. We generally smooth this field with a top-hat of length L ,

$$A_L(x) = \frac{1}{L} \int_{-\infty}^{\infty} dx' \Theta(|x - x'|; L/2) A(x'), \quad (3.10)$$

where Θ is a top-hat function. The quantity $A_L(x)$ is a measure of the average small scale power across different portions of a quasar spectrum, and should broadly trace the temperature of corresponding regions in the IGM, with cold regions giving a larger A_L than hot regions.

3.5.1 Hydrodynamic Simulations: Perfect Temperate-Density Relation Models

As a first test, we take high redshift outputs from the hydrodynamic simulation (see §4.2.2) and impose temperature-density relations before producing mock quasar spectra. This test ignores the impact of inhomogeneous reionization, but it nonetheless provides some intuition for how well our approach can constrain the IGM temperature.

Fig. 3.10 shows an example sightline, 50 Mpc/ h in length, extracted from the hydrodynamic simulation at $z = 5$ for each of two different temperature-density relation models. The top panel shows the transmission fluctuation, δ_F , for models with $\gamma = 1.3$ and each of $T_0 = 7.5 \times 10^3$ K and $T_0 = 2.5 \times 10^4$ K while the bottom panel shows the smoothed wavelet amplitudes, A_L , in each model. In this case, the smoothing scale s_n is set to $s_n = 51$ km/s, the pixel size to $\Delta u = 3.2$ km/s, and $L = 1,000$ km/s. In each case the intensity of the ionizing background has been renormalized so that the global mean transmitted flux is

3.5 Measuring the Temperature of the $z \sim 5$ IGM

$\langle F \rangle = 0.20$. This is the mean transmitted flux implied by extrapolating the recent best-fit measurement of ?) to $z = 5$.¹

Although the differences between δ_F along the two example sightlines are generally small, there are some noticeable differences. In particular, it appears that the “spikes” of transmission in the colder model are more prominent. This is mostly a result of the larger Doppler widths in the hot model. At this redshift, the heights of the transmission spikes are often influenced by nearby gas elements that are centered on saturated or highly absorbed parts of the spectrum; the broad Doppler wings from this gas extend into adjoining unsaturated regions and thereby reduce the height of neighboring transmission spikes. The spikes are less impacted by the narrower Doppler wings in the colder model and remain more prominent. Essentially, the forest has become “inverted” at these redshifts in comparison to at lower redshift. At sufficiently low redshift, the forest is mostly transmitted with some prominent absorption lines interspersed. In the low redshift case most of the information about the IGM temperature comes from narrow absorption lines. At high redshift, the forest is mostly absorbed and most of the information about the IGM temperature is instead in the transmission spikes.

In either case, the amount of small scale power in the Ly- α forest is indicative of the temperature of the gas in the IGM. This is illustrated by the bottom panel of Fig. 3.10 for the high redshift case considered here. This panel shows the smoothed wavelet amplitudes along each line of sight. The smoothed wavelet amplitudes are larger in the cold IGM model, with the largest differences occurring near $\Delta v \sim 4500$ km/s, close to several prominent transmission spikes in the models.

One possible complication is that long completely saturated regions will – regardless of temperature – have low wavelet amplitudes, A_L , since there is no small scale structure in such regions. These saturated zones may in fact be *more prominent* in models with

¹Specifically, we use these authors’ smooth functional fit to their measured effective optical depth. This is an (approximate) fit to measurements in bins centered on redshifts from $z = 2.15$ to $z = 4.85$, and so our extrapolation of this fit out to $z = 5$ is only very slight.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

low temperature since cold regions recombine more quickly, and hence have larger neutral fractions and suffer more absorption than hot regions.¹ Fig. 3.10 suggests, however, that this is not a big effect at $z = 5$, $\langle F \rangle = 0.2$, although the saturated regions will be more prominent at higher redshift (see §3.5.3). We can guard against “contamination” from saturated regions by masking them before measuring the probability distribution of the wavelet amplitudes, and by varying the smoothing scale L .

To characterize the variations of the wavelet amplitudes with temperature more quantitatively, we calculate the probability distribution function (PDF) of wavelet amplitudes for ensembles of mock spectra generated from different temperature-density relation models. The results of these calculations are shown in Fig. 3.11. The PDF is quite sensitive to the temperature at mean density in these models. For example, the location of the peak in the wavelet amplitude PDF is at an A_L that is roughly three times larger in the coldest model shown (with $T_0 = 7.5 \times 10^3$ K), compared to the hottest model considered here ($T_0 = 2.5 \times 10^4$ K). For the mean transmitted flux ($\langle F \rangle = 0.20$) and $z = 5$, the wavelet PDF for $s_n = 51$ km/s is sensitive mostly to densities near the cosmic mean. As a result, we find that the wavelet PDFs here depend strongly on T_0 , but are insensitive to γ , the slope of the temperature-density relation. It is hence important to keep in mind that our approach for measuring the IGM temperature is only sensitive to the temperature of the IGM close to the mean density, and it is therefore not possible to extract the full trend of temperature with density shown in our models (e.g., Fig. 3.5).

Although the PDFs depend sensitively on T_0 , it is also clear that the wavelet amplitudes are not *perfect* indicators of the temperature. In the limit that the temperature at mean density were the only quantity that determined A_L , these PDFs should approach delta functions in A_L . That the wavelet PDFs have some breadth is not, however, surprising:

¹Although the precise impact of the temperature on the wavelet amplitude PDFs shown here is not this transparent since we are comparing models at fixed mean transmitted flux.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

the temperature is clearly not the only quantity that determines the small-scale structure in the forest. That said, Fig. 3.11 looks promising and helps to motivate further study.

3.5.2 Degeneracy with the Mean Transmitted Flux

One other potential issue, however, is that the wavelet PDF is also sensitive to the somewhat uncertain value of the mean transmitted flux, $\langle F \rangle$. Although the present statistical uncertainties on this quantity are $\sigma_{\langle F \rangle}/\langle F \rangle \leq 10\%$ near $z \sim 5$ (?), the systematic uncertainties are significantly larger. In particular, it is difficult to estimate the unabsorbed quasar continuum level, especially at the redshifts of interest for this study, where the absorption in the forest is very large (e.g. ?). The *measurement* of the wavelet PDF itself should, however, be fairly robust to uncertainties in the level of the unabsorbed quasar continuum. This is the case because we consider the statistics of the transmission fluctuations, $\delta_F = (F - \langle F \rangle)/\langle F \rangle$, for which a (multiplicative) error in the continuum normalization divides out (see ? for a discussion and tests with lower redshift data).

Nevertheless, we still need to know the mean transmitted flux very accurately: we use this measurement to in turn fix the intensity of the ionizing background at the redshifts of interest, which is itself quite uncertain. The amount of small scale structure in the forest, and hence the model wavelet amplitudes, do depend on the overall mean transmitted flux. As a result, while we should be able to measure the wavelet PDF without knowing the precise continuum normalization, our interpretation of this measurement still requires knowing the mean transmitted flux. To illustrate this, we plot (in the top panel of Fig. 3.12) the $z = 5$ wavelet amplitude PDF in a model with $T_0 = 1.5 \times 10^4$ K and the preferred mean transmitted flux at this redshift ($\langle F \rangle = 0.20$, red short-dashed line). As in Fig. 3.11, the wavelet amplitudes are smaller in this model than in, for example, the cooler model with $T_0 = 7.5 \times 10^3$ K at the *same value of the* mean transmitted flux (blue long-dashed line in the top panel of Fig. 3.12). However, if we allow the mean transmitted flux to increase in the colder model, the resulting wavelet PDF becomes similar to that in the hotter model.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

In particular, the black solid line shows a colder model with the mean transmitted flux increased to $\langle F \rangle = 0.30$; this closely matches the wavelet PDF in the hotter model at the smaller mean transmitted flux ($\langle F \rangle = 0.20$). This particular value of the mean transmitted flux, $\langle F \rangle = 0.30$, is well outside the presently allowed range, given the statistical errors on current measurements. Nevertheless, the wavelet PDF clearly shows some degeneracy between variations in T_0 and in $\langle F \rangle$. This invites further attention, especially given the systematic concerns associated with estimating the unabsorbed continuum level.

One way to help break this degeneracy is to combine the measured wavelet PDF with a measurement of the flux power spectrum on *larger scales*. This quantity is especially sensitive to the mean transmitted flux, and the power spectrum of δ_F has the virtue – like the wavelet PDF – that it is insensitive to the overall normalization of the quasar continuum. On the other hand, on sufficiently large scales, the gentle fluctuations in the underlying quasar continuum still likely contaminate this measurement. However, there should still be a useful range of scales where the structure in the forest dominates over that in the continuum (see e.g. ?) and it is these scales that we will consider to help break the $T_0 - \langle F \rangle$ degeneracy.

To illustrate how the flux power measurement may help break this degeneracy, we plot the amplitude of transmission fluctuations, $\Delta_F^2 = kP_F(k)/\pi$, for a single example wavenumber ($k = 0.003$ s/km in velocity units, or $k = 0.39h$ Mpc $^{-1}$ in co-moving units at $z = 5$) as a function of mean transmitted flux. The 1-D flux power spectrum ($P_F(k)$) is fairly flat on large scales and so the precise k considered here is not especially important. The bottom panel of Fig. 3.12 shows the flux power spectrum at $k = 0.003$ s/km as a function of $\langle F \rangle$ for the two values of T_0 . In each case, Δ_F^2 is a strong function of $\langle F \rangle$. The red triangle and black pentagon show the power spectra in the $T_0 = 1.5 \times 10^4$ K and the $T_0 = 7.5 \times 10^3$ K models respectively, for the values of the mean transmitted flux ($\langle F \rangle = 0.20$ and $\langle F \rangle = 0.30$) at which their wavelet PDFs are degenerate. The large scale flux power spectra in these two models differ by the sizable factor of 1.7. It should be straightforward

3.5 Measuring the Temperature of the $z \sim 5$ IGM

to measure the flux power spectrum on these scales to this level of accuracy, and so this measurement can help pin down $\langle F \rangle$ and break the degeneracy.

One possible concern with this approach is that the precise relationship between Δ_F^2 and $\langle F \rangle$ may be somewhat model dependent, and our inability to perfectly model the forest – especially at high redshifts, potentially close to the EoR – might lead us to draw spurious conclusions. At present, the only way to guard against this possibility is to test the goodness-of-fit of our models for as wide a range of empirical tests as possible. Ideally, one would compare models with measurements of the flux power across a wide range of scales (although significantly larger scales will be subject to contamination from power in the quasar continuum), the wavelet PDF, the mean transmitted flux, and perhaps the statistics of the Ly- β forest as well (e.g. ? ?).

3.5.3 Wavelet Amplitude PDFs in Inhomogeneous Reionization Models

With the results of the previous section as a guide, we now turn to consider the wavelet amplitude PDFs in the more realistic inhomogeneous temperature models developed in §3.4. In this case, we are using the dark matter simulations of McQuinn et al. (31) along with our model temperature distributions. Although the large volume of these simulations allows us to capture the reionization-induced inhomogeneities, they are not – taken as is – adequate for capturing the small-scale structure in the Ly- α forest, which is the basic observable we aim to explore here.

In order to make headway, we add small-scale structure to sightlines extracted from the simulation cube using the log-normal model, as in ?). Briefly, we generate one-dimensional Gaussian random fields δ_G using the one-dimensional linear density power spectrum (scaled to the redshift of interest, and calculated after smoothing the three-dimensional linear power spectrum with a filter of the form e^{-2k^2/k_f^2} and $k_f = 30h \text{ Mpc}^{-1}$ to loosely mimic the effect of Jeans smoothing, ?). From the Gaussian random realizations, we produce lognormal fields at high resolution using the transformation $1 + \delta_{\text{LN}} = e^{\delta_G - \sigma_G^2/2}$, where σ_G^2 is the

3.5 Measuring the Temperature of the $z \sim 5$ IGM

variance of the Gaussian random field. As in [?] the lognormal field is modulated by the larger scale modes captured in the simulation (δ_{sim}) according to $1 + \delta = (1 + \delta_{\text{sim}})(1 + \delta_{\text{LN}})$ with the (subscript-free) symbol δ denoting the density contrast with added small-scale structure. Similarly, using the simulated temperature in a coarse pixel (described by T_0 and γ), the temperature in a fine pixel becomes $T = T_0(1 + \delta)^{\gamma-1}$.

The main disadvantage here is that the resulting sightlines have too much large scale structure: the lognormal field adds both large and small scale modes to the simulation, and the simulation was not deficient in large scale power to begin with. This is partly mitigated by our using a slightly higher redshift simulation output ($z = 6.9$) than the redshift of interest. This simple approach is hence imperfect, but the added small scale structure does nevertheless allow us to reliably model the impact of thermal broadening on the resulting mock Ly- α forest spectra. As a test, we measure the flux power spectrum from the mock “lognormal-enhanced” sightlines and compare them with the flux power spectrum from mock spectra generated from the hydrodynamic simulation. The flux power spectrum from the lognormal spectra is roughly 50% larger than from the hydrodynamic simulations. However, the overall shape of the flux power is fairly well captured in the lognormal case, and importantly, the shape of the flux power spectrum varies in a similar way in both calculations as the temperature and thermal broadening are varied. Hence we believe that this approach suffices to capture the main impact of patchy reionization on the small scale structure in the Ly- α forest. We caution, however, that a detailed comparison with upcoming measurements will certainly require improvements here.

With this cautionary remark, we turn to consider the properties of mock spectra drawn from our inhomogeneous temperature models. Fig. 3.13 shows typical example sightlines from the High-z and Low-z reionization models at $z = 5.5$ and $\langle F \rangle = 0.1$. The trends are broadly similar to those in Fig. 3.10: the colder models have more small scale structure than the hotter models. As a result, the transmission field has more prominent spikes in the colder High-z model, and the smoothed wavelet amplitudes in this model (bottom panel)

3.5 Measuring the Temperature of the $z \sim 5$ IGM

are larger than in the Low-z model. The inhomogeneous models incorporate, however, the reionization-induced temperature variations that are not included in the previous model, although the impact of these variations are generally hard to discern by eye. One can however identify that the prominent cold region in the middle of this sightline, for example, corresponds to a pronounced peak in the smoothed wavelet amplitude field in each model. Note that the wavelet amplitudes and δ_F fluctuations are larger here than in Fig. 3.10 because here we consider $z = 5.5$ and $\langle F \rangle = 0.1$, while in the previous figure we considered $z = 5.0$ and $\langle F \rangle = 0.2$. In addition, the lower transmitted flux considered here leads to more completely absorbed regions in the mock Ly- α forest, and these regions have correspondingly low wavelet amplitudes. As discussed earlier, these regions do not contain information about the IGM temperature.

More quantitatively, the resulting wavelet amplitude PDFs for some example inhomogeneous temperature models are shown in Fig. 3.14. Since the mean transmitted flux at this redshift ($z = 5.5$) is somewhat uncertain, we compare models normalized to each of $\langle F \rangle = 0.2$ (top panel) and $\langle F \rangle = 0.1$ (bottom panel) in order to illustrate the impact of varying $\langle F \rangle$. Extrapolating the best fit measurement from (?), we find $\langle F \rangle = 0.13$ at $z = 5.5$ so this range should approximately bracket the expected value, although the lower end of this range is preferred. Note that there is some evidence that the mean transmitted flux decreases more rapidly above $z \sim 5.5$ or so compared to the evolution expected from lower redshifts (e.g. Fan et al. 17), and so this would further favor the lower value, and perhaps even slightly smaller numbers than considered here. Nevertheless, it is worth considering the mean transmitted flux dependence explicitly: even if $\langle F \rangle = 0.2$ is reached at $z = 5$ rather than $z = 5.5$, the temperature fluctuations could be as large at $z = 5$ as in our $z = 5.5$ models if reionization is more extended than considered here.

In each case, the wavelet filter's smoothing scales are set to $s_n = 34$ km/s and $L = 1,000$ km/s, while the pixel size is $\Delta u = 2.1$ km/s.¹ In each panel, the PDF in the High-z model

¹The pixel size here matches that of typical Keck HIRES spectral pixels. Note that the smoothing scale,

3.5 Measuring the Temperature of the $z \sim 5$ IGM

is compared to Low-z models for two different reionization temperatures, $T_r = 2 \times 10^4$ K and $T_r = 3 \times 10^4$ K. In addition, we plot the wavelet amplitude PDF for a model with a completely homogeneous temperature field. In this isothermal case the temperature is set to $T_0 = 1.44 \times 10^4$ K, matching the median temperature near the cosmic mean density in the Low-z, $T_r = 3 \times 10^4$ K model. Homogeneous reionization models with the same T_0 but differing γ would give very similar results, since the wavelet amplitudes are sensitive to absorbing gas near the cosmic mean density for these values of the mean transmitted flux. Comparing the $T_0 = 1.44 \times 10^4$ K isothermal model with the Low-z, $T_r = 3 \times 10^4$ K case helps to illustrate the impact of temperature inhomogeneities.

First, we consider how the location of the peak in the PDF varies with the reionization model. The qualitative behavior is as expected: the High-z model peaks at the largest wavelet amplitude, while the Low-z models peak at smaller wavelet amplitude. This results because the High-z model is colder than the Low-z models, and so it has more small-scale structure and hence higher wavelet amplitudes. The Low-z model with the higher reionization temperature, $T_r = 3 \times 10^4$ K, has hotter gas at the redshift of interest ($z = 5.5$) than in the Low-z scenario with $T_r = 2 \times 10^4$ K and so the PDF in the former model is peaked at smaller wavelet amplitudes.

These trends occur in both the $\langle F \rangle = 0.1$ and $\langle F \rangle = 0.2$ models. In the smaller $\langle F \rangle$ model the PDFs peak at larger A_L , reflecting the enhanced power in the $\delta_F = (F - \langle F \rangle) / \langle F \rangle$ field for decreasing mean flux. To provide a quantitative comparison, the High-z model has an average wavelet amplitude that is 73% larger than in the Low-z, $T_r = 2 \times 10^4$ K model, while the average amplitude in the $T_r = 3 \times 10^4$ K, Low-z model is 27% smaller than the $T_r = 2 \times 10^4$ K case. These numbers are for $\langle F \rangle = 0.2$, but the fractional differences are similar for $\langle F \rangle = 0.1$.

s_n , and pixel size, Δu , here are similar, but slightly different than used in the previous section. The large scale smoothing, L , is identical.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

Next we consider the *width* of the wavelet amplitude PDFs. The width arises in part because A_L is not a perfect tracer of temperature, and also because the temperature field is inhomogeneous. The latter contribution to the width can in principle be used to constrain the spread in the timing of reionization, and so this quantity is highly interesting. Comparing first the inhomogeneous models in Fig. 3.14, it is clear that the Low-z, $T_r = 3 \times 10^4$ K model has the widest distribution of wavelet amplitudes, followed by the Low-z, $T_r = 2 \times 10^4$ K model, while the High-z model has the narrowest A_L distribution of these models. This is expected, since the High-z model has the smallest temperature fluctuations, while the Low-z, $T_r = 3 \times 10^4$ K model has the largest temperature fluctuations. It is also instructive to compare the isothermal models (magenta dot-dashed lines in each panel) with the Low-z, $T_r = 3 \times 10^4$ K models. The isothermal model has the same median temperature near the cosmic mean density as the Low-z, $T_r = 3 \times 10^4$ K model. This PDF is similar, but narrower, than in the inhomogeneous temperature case. Quantitatively, the fractional width of the distribution, $\sigma_{A,L}/\langle A_L \rangle$, is 18% larger in the Low-z, $T_r = 3 \times 10^4$ K model than in the homogeneous case at $\langle F \rangle = 0.2$ and 11% larger at $\langle F \rangle = 0.1$. These relatively small differences seem challenging to extract, but are sufficiently interesting to merit further investigation.

3.5.4 Forecasts

Finally, we briefly forecast the significance at which various models may be distinguished using *existing* data samples. Here we will be content with rough estimates. For simplicity, we predict the expected error bar on only the first two moments of the wavelet amplitude PDF, and compare this to the difference between some of our models. We consider a sample of N_{los} independent spectra, and assume that each spectrum has sufficient S/N so that we can estimate error bars in the sample variance limit – i.e., we work in the limit that photon noise from the night sky and the quasar itself, as well as instrumental noise, are negligible compared to sample variance (also termed “cosmic variance”). Our error budget hence

3.5 Measuring the Temperature of the $z \sim 5$ IGM

reflects the scatter expected – given the large scale structure of the universe and the limited volume probed by our hypothetical survey – around the true value that would be obtained if we could average over an infinite volume.

It is instructive to first consider the type of quasar spectra that are required to measure the wavelet amplitudes in the sample variance limit. The first obvious requirement is that the spectral resolution needs to be high enough to resolve the thermal broadening scale, which is on the order of ~ 10 km/s. This can be achieved with, for example, Keck HIRES spectra which have a spectral resolution of $\text{FWHM} = 6.7$ km/s. Spectra from the MIKE spectrograph on Magellan would partly resolve the thermal broadening scale: the resolution of these spectra is a factor of ~ 2 worse than HIRES (e.g. ?). Next, we consider the impact of photon and instrumental noise. In particular, we estimate the S/N (at the continuum) per HIRES pixel at which the expected shot-noise is a small fraction of the average wavelet amplitude in plausible models. The mean wavelet amplitude from the noise should be roughly $\langle A_{\text{noise}} \rangle \approx (N/S)^2 / \langle F \rangle$ (? ?). In order for the noise to be sub-dominant, we impose that the noise should be less than 10% of the mean wavelet amplitude in our Low- z , $T_r = 2 \times 10^4$ K model at $\langle F \rangle = 0.2$ (which has $\langle A \rangle = 0.35$). This requires a $S/N \gtrsim 12$ at the continuum, per 2.1 km/s HIRES pixel. This is a fairly stringent requirement for quasars at the high redshifts of interest for our proposed measurements, but this sensitivity has been reached already in previous work. We could likely make a less stringent requirement on the S/N of the data sample: this would just necessitate careful shot-noise subtraction, and boost our error budget somewhat. Currently, we are aware of roughly ~ 10 HIRES spectra in the published literature at $z \sim 5 - 5.7$ that meet our S/N and resolution criteria (e.g. ?). There are substantially larger numbers of lower resolution and S/N spectra from the SDSS that could potentially be followed-up at higher resolution and sensitivity to improve the statistics here. For example, there are 36 SDSS-DR7 quasars in the $z = [5.0, 5.2]$ redshift bin of ?).

We now proceed to estimate the sample variance errors. The first quantity of interest

3.5 Measuring the Temperature of the $z \sim 5$ IGM

is the sightline-to-sightline scatter in the wavelet amplitudes, averaged over the entire Ly- α forest region of each quasar. We define $P_{A,L}(k)$ to be the power spectrum of the fluctuations in the wavelet amplitude after smoothing the amplitudes on scale L , i.e., the power spectrum of $\delta_{A,L}(x) = (A_L(x) - \langle A_L \rangle) / \langle A_L \rangle$. We relate the expected error bars to this power spectrum, and estimate them by measuring $P_{A,L}(k)$ from our simulated models. This approach has the advantage that we can approximately extrapolate $P_{A,L}(k)$ to scales beyond that of our simulation box and roughly account for missing large scale Fourier modes. This power spectrum of $\delta_{A,L}(x)$ is a four point function of the flux and is related to the power spectrum of $A(x)$ (the *unsmoothed* wavelet amplitude power spectrum) from Eqs. 3.9 and 3.10 by

$$P_{A,L}(k) = \left[\frac{\sin(kL/2)}{kL/2} \right]^2 P_A(k). \quad (3.11)$$

This is just a filtered version of the wavelet amplitude power spectrum. From each independent sightline, we estimate the moments of the wavelet amplitude PDF by further averaging over a length scale L_{spec} , comparable to the separation (in velocity units) between the Ly- α and Ly- β emission lines from the quasar. Note the distinction between the two smoothing scales here: L is the smoothing scale over which we are studying the wavelet amplitude variations, while L_{spec} is the scale over which we estimate the moments from each spectrum.

The formula for the sample variance for a single sightline is then (?):

$$\frac{\sigma_{A,L}^2(L_{\text{spec}})}{\langle A \rangle^2} = \int_{-\infty}^{\infty} \frac{dk'}{2\pi} \left[\frac{\sin(k'L_{\text{spec}}/2)}{k'L_{\text{spec}}/2} \right]^2 P_{A,L}(k'). \quad (3.12)$$

In a sample of N_{los} independent Ly- α forest sightlines, the expected ($1-\sigma$) fractional error on $\langle A_L \rangle$ (in the sample variance limit) is:

$$\frac{\delta \langle A_L \rangle}{\langle A_L \rangle} = \frac{1}{\sqrt{N_{\text{los}}}} \frac{\sigma_{A,L}(L_{\text{spec}})}{\langle A_L \rangle}. \quad (3.13)$$

We can compare this estimate of the fractional error in the average wavelet amplitude with the difference between the average amplitudes in some of the models of the previous section.

3.5 Measuring the Temperature of the $z \sim 5$ IGM

Next, we want to consider the expected error bar on the second moment of the wavelet PDF. This second moment provides one diagnostic for the impact of temperature inhomogeneities from patchy reionization. We would like to check whether the variance of the A_L distribution is broad enough to imply patchy reionization. This requires computing the “variance of the variance”; in particular, we want the variance of an estimate of $\sigma_{A,L}^2$ when this quantity is estimated from a sightline of length L_{spec} . We calculate this quantity assuming that A_L approximately obeys Gaussian statistics. In this case, one can show that the desired variance is:

$$\begin{aligned} \text{Var}[\sigma_{A,L}^2(L_{\text{spec}})] &= 2\langle A_L \rangle^4 \int_{-\infty}^{\infty} \frac{dk'}{2\pi} \left[\frac{\sin(k'L_{\text{spec}}/2)}{k'L_{\text{spec}}/2} \right]^2 \\ &\quad \times \int_{-\infty}^{\infty} \frac{dk''}{2\pi} P_{A,L}(k' - k'') P_{A,L}(k'') \\ &\quad + 4\langle A_L \rangle^4 \sigma_{A,L}^2(L_{\text{spec}}). \end{aligned} \quad (3.14)$$

The expected error on an estimate of $\sigma_{A,L}^2$ from a sample of N_{los} independent sightlines is then:

$$\frac{\delta\sigma_{A,L}^2}{\sigma_{A,L}^2} = \frac{1}{\sqrt{N_{\text{los}}}} \frac{\sqrt{\text{Var}[\sigma_{A,L}^2(L_{\text{spec}})]}}{\sigma_{A,L}^2}. \quad (3.15)$$

We can now plug numbers into Eqs. 3.13 and 3.15 to estimate the ability of current samples to constrain some of our models. We assume that $N_{\text{los}} = 10$ sightlines are available for our study, and take $z = 5.5$, $\langle F \rangle = 0.2$ here. We assume the true model is the $T_r = 3 \times 10^4$ K, Low-z case and examine at what significance other models may be distinguished from this case. Conservatively, we assume that $L_{\text{spec}} = 2.5 \times 10^4$ km/s; the velocity separation between the Ly- α and Ly- β emission lines is 5.1×10^4 km/s, and so our assumed value effectively masks-out half of the forest. While one will want to mask-out proximity regions, damped Ly- α systems, prominent metal lines, etc., our choice is certainly conservative. In this case, evaluating Eq. 3.13 in our assumed true model, we find that a measurement of $\langle A_L \rangle$ from $N_{\text{los}} = 10$ sightlines should rule out the High-z ($T_r = 2 \times 10^4$ K) model at $32 - \sigma$,

3.6 Conclusions

and a Low-z model with the lower reionization temperature ($T_r = 2 \times 10^4$ K) at $9 - \sigma$! These forecasts are optimistic, since we have assumed – for example – perfect knowledge of the mean transmitted flux; in practice, the mean transmitted flux may have to be constrained separately from the large-scale flux power spectrum (§3.5.2). Nonetheless, we believe the overall point is robust: existing samples should provide interesting constraints on $\langle A_L \rangle$.

Significantly more challenging is to measure the variance of the A_L distribution well enough to identify signatures of patchy reionization. With the optimistic “true” model considered here (the Low-z, $T_r = 3 \times 10^4$ K case), however, we forecast that $N_{\text{los}} = 10$ spectra are sufficient to rule out the homogeneous $T = 1.44 \times 10^4$ K model¹ at $2.4 - \sigma$, based on the variance of the A_L distribution alone. Note that both of these estimates assumed $\langle F \rangle = 0.2$ and our $z = 5.5$ temperature models but the expected constraints are similar for $\langle F \rangle = 0.1$.

3.6 Conclusions

In this work, we modeled the temperature of the IGM at $z \gtrsim 5$, incorporating the impact of spatial variations in the timing of reionization across the universe. We contrasted the $z \sim 5$ temperature in models where reionization completes at high redshift – near $z = 10$ – with scenarios where reionization completes later, near $z = 6$. In agreement with previous work (? ?)², we found that the properties of the $z = 5$ temperature differ markedly between these two models. The IGM is cooler in the early reionization model, and the usual temperature-density relation is a good description of the temperature state in this case, while the temperature state is more complex and inhomogeneous in the late reionization scenario.

¹Recall that the temperature in this model matches the median temperature for gas near the cosmic mean density in the Low-z, $T_r = 3 \times 10^4$ K model.

²This is also in general agreement with still earlier work by ? and ?, although these two studies did not incorporate inhomogeneities in the timing of reionization.

3.6 Conclusions

We then produced mock $z \gtrsim 5$ Ly- α forest spectra from our numerical models, in effort to explore the observable implications of the IGM temperature as close as possible to hydrogen reionization. In particular, we used the Morlet wavelet filter approach of ?) to extract the small-scale structure across each Ly- α forest spectrum. The small-scale structure in the forest is sensitive to the temperature of the IGM, and the filter we use is localized in configuration space, which makes it well-suited for application in cases where the temperature field is inhomogeneous.

Interestingly, we found that the small-scale structure in the forest is sensitive to the IGM temperature even when the forest is highly absorbed. In particular, the transmission field in between absorbed regions is more spiky if the IGM is cold, compared to hotter models. Using existing high resolution Ly- α forest samples, one should be able to use this difference to distinguish between high redshift and lower redshift reionization models at high significance. It may, however, be necessary to combine measurements of the small-scale structure in the forest with measurements of the larger scale flux power spectrum to help break degeneracies with the mean transmitted flux, which is hard to estimate directly at the high redshifts of interest for these studies.

In addition, we considered the impact of spatial variations in the timing of reionization on the width of the wavelet amplitude distribution. We found that these variations broaden the width of this distribution, but that the broadening is fairly subtle. This likely results in part because the temperature variations we are interested in are coherent on rather large scales, and aliasing – from fluctuations in the transmission field transverse to the line of sight – obscures our ability to measure large scale fluctuations along the line of sight (e.g. ? ?). Nonetheless, we forecast that our Low-z, $T_r = 3 \times 10^4$ K model can be distinguished from a homogeneous temperature model at $2 - 3\sigma$ with existing samples of ten high resolution sightlines. Larger samples could improve on this, and an analysis of the small-scale structure in the Ly- β forest might help as well. In this paper, we focused on the small-scale structure since it is a direct indicator of the temperature, but another approach

3.6 Conclusions

would be to consider instead transmission fluctuations on large scales, especially as probed in “3D” measurements of the Ly- α forest (e.g. ?). This may be possible at $z \gtrsim 4$ with DESI (? ?).

To robustly interpret future measurements, our modeling should be improved in various ways. In particular, we should incorporate inhomogeneous Jeans smoothing effects into our modeling. This might be accomplished by, for example, incorporating our semi-numeric modeling on top of a large dynamic range HPM (?) simulation. These calculations will need to face the competing requirements of capturing the large-scale variations in the timing of reionization, while simultaneously resolving the filtering scale. Nevertheless, we believe that measurements of the $z \gtrsim 5$ IGM temperature should provide a valuable handle on the reionization history of the universe.

Appendix: Approximate Thermal History Calculations

Here we derive an approximate analytic formula for the thermal history of an IGM gas element using linear perturbation theory. Here our derivation is quite similar to the analytic calculation in ?) (their §3.1), except here we include Compton cooling off of the CMB, which is important for our application in which we consider high redshift reionization and the temperature at redshifts close to reionization.

It is instructive to first briefly examine which heating/cooling processes are important, in addition to the usual adiabatic heating/cooling from the contraction/expansion of gas parcels. In particular, Fig. 3.15 compares the relative importance of HI photoheating, HeI photoheating, Compton cooling, HII recombination cooling, HeII recombination cooling, and free-free emission cooling for intergalactic gas at $z \sim 7$. The figure assumes that the gas is in ionization equilibrium and that hydrogen is highly ionized and helium mostly singly ionized. As in the body of the text, we are assuming that HeII reionization has not yet commenced at the redshifts of interest. For the highly-ionized and low density

3.6 Conclusions

intergalactic gas considered here, line excitation cooling and collisional ionizations should be unimportant. The left hand panel considers the various heating and cooling processes for gas of fixed temperature, $T = 10^4$ K, as a function of density while the right hand panel shows the same for gas at the cosmic mean density as a function of temperature. In each panel, the curves in the figure indicate the absolute values of the various rates, so that cooling processes are shown as positive numbers on the plot. The photoheating curves assume that the hardened ionizing spectrum follows a $J(\nu) \propto \nu^{-1.5}$ power-law near the photoionization edges. The dominant processes are clearly HI photoheating and Compton cooling. After these processes in importance are HeI photoheating and HII recombination cooling: these have rates that are roughly 20% smaller than HI photoheating near the cosmic mean density and $T \sim 10^4$ K at $z \sim 7$. It is also interesting to note that at the densities considered here and for the adopted ionizing spectrum, HeI photoheating and HII recombination cooling have nearly equal magnitudes near $T \sim 10^4$ K; since these two processes enter the thermal evolution equation with opposite signs, this leads to a partial cancellation.

As a result, a good approximation to the IGM thermal evolution (at high redshifts before HeII reionization) results from including only adiabatic heating/cooling, Compton cooling, and HI photoheating. Note, however, that in the body of the work we include all of the additional processes considered in Fig. 3.15. The approximate results here are nonetheless useful and fairly accurate, and can in turn help to build intuition. The approximate equation for the thermal evolution is then:

$$\frac{dT}{dt} = -2HT + \frac{2T}{3(1+\delta)} \frac{d\delta}{dt} + \frac{\alpha_0 \bar{n}_e E_J}{3(1+\chi_{He}) k_B} \left(\frac{T}{10^4 K} \right)^{-0.7} (1+\delta) + \frac{4 \sigma_T a_{\text{rad}} T_\gamma^4}{3 m_e c} (T_\gamma - T). \quad (3.16)$$

Here α_0 is the (case-A) recombination coefficient for hydrogen at $T = 10^4$ K¹, E_J is the

¹We assume the case-A recombination coefficient and use the approximation $\alpha_A = 4.2 \times 10^{-13} (T/10^4 K)^{-0.7} \text{ cm}^3 \text{s}^{-1}$ (?) in this Appendix.

3.6 Conclusions

average energy injected into the gas per photoionization¹, T_γ is the CMB temperature (at the scale factor of interest), σ_T is the Thomson scattering cross section, and $a_{\text{rad}}T_\gamma^4$ is the energy density in the CMB. The first two terms describe adiabatic cooling/heating, the third term is from photoionization heating, and the last term accounts for Compton cooling. This equation assumes that the gas is in photoionization equilibrium, adopts an electron number density of $n_e = n_H + n_{He}$, and assumes the number density of free particles in the gas is $n_{\text{tot}} = n_e + n_H + n_{He} = 2(n_H + n_{He})$.

As in [?], we assume a solution of the form $T = T_0(1 + \delta)^{\gamma-1}$ and linearize ($T \approx T_0[1 + (\gamma - 1)\delta]$) to find equations for T_0 and $\gamma - 1$, as functions of scale factor. Let us first introduce two constants to make the notation more compact:

$$\mathcal{A} = (10^4 K)^{0.7} \frac{\alpha_0 \bar{n}_e(0) E_J}{3(1 + \chi_{\text{He}}) k_B H_0 \sqrt{\Omega_m}}. \quad (3.17)$$

Here $\bar{n}_e(0)$ denotes the present day ($z = 0$), spatially averaged, electron number density. Note that the constant \mathcal{A} has dimensions of $[\mathcal{A}] = [K]^{1.7}$. In our fiducial model with $\alpha = 1.5$, the numerical value of \mathcal{A} is $\mathcal{A} = 5.77 \times 10^5 K^{1.7}$. Second, we introduce

$$\mathcal{B} = \frac{1}{H_0 \sqrt{\Omega_m} t_{\text{Comp}}(0)}; \quad t_{\text{Comp}} = \frac{3m_e c}{4\sigma_T a_{\text{rad}} T_\gamma^4}. \quad (3.18)$$

Here $t_{\text{comp}}(0)$ a characteristic timescale for Compton cooling today ($z = 0$); this timescale falls off towards high redshift as $t_{\text{comp}} \propto a^4$. In our assumed cosmology, the numerical value of this constant is $\mathcal{B} = 1.15 \times 10^{-2}$.

Using the high redshift approximation for the Hubble parameter, $H \approx H_0 \sqrt{\Omega_m} a^{-3/2}$, and setting $\delta = 0$ to find an equation for $T_0(a)$ valid in linear theory, Eq. 3.16 gives the following equation for T_0 :

$$\frac{d(a^2 T_0)}{da} = \mathcal{A} a^{0.9} (a^2 T_0)^{-0.7} - \mathcal{B} a^{-7/2} (a^2 T_0) + \mathcal{B} T_\gamma(0) a^{-5/2}. \quad (3.19)$$

¹ Assuming a power-law spectral index, $J_\nu \propto \nu^{-\alpha}$ (with the power law accounting for hardening from absorption), $E_J = h\nu_{HI}/(\alpha + 2)$.

3.6 Conclusions

A similar equation follows for $\gamma - 1$ (again valid to linear order in δ , and with the approximations to the thermal evolution equation in Eq. 3.16):

$$\frac{d(\gamma - 1)}{da} = \left[\frac{2}{3} - (\gamma - 1) \right] \frac{1}{a} + \mathcal{A}a^{0.9}(a^2 T_0)^{-1.7} [1 - 1.7(\gamma - 1)] - \frac{\mathcal{B}T_\gamma(0)}{a^2 T_0} a^{-5/2}(\gamma - 1) \quad (3.20)$$

We can safely neglect the third term in Eq. 3.19, and we find a solution for $T_0(a)$ of the form (with the initial condition that the gas element is ionized at scale factor a_r to a temperature T_r):

$$u = u_r \exp \left[0.68\mathcal{B}a^{-5/2} - 0.68\mathcal{B}a_r^{-5/2} \right] + 0.68\mathcal{A}(0.68\mathcal{B})^{0.76} e^{0.68\mathcal{B}a^{-5/2}} \int_t^{t_r} dt' t'^{-1.76} e^{-t'}.$$

$$u = (a^2 T_0)^{1.7}; \quad u_r = (a_r^2 T_r)^{1.7}; \quad t = 0.68\mathcal{B}a^{-5/2}; \quad t_r = 0.68\mathcal{B}a_r^{-5/2} \quad (3.21)$$

The corresponding solution to Eq. 3.20 for the evolution of γ does not have a simple closed analytic form, but the equations can be solved numerically. Comparing the solutions for $T_0(z)$ from Eq. 3.21 and $\gamma(z)$ from Eq. 3.20 with results given in the body of the text, we find that the approximate solutions are good to better than 10% accuracy. This accuracy is, in fact, somewhat better than might be expected given the approximations made, and may reflect cancelations between some of the neglected terms (such as the compensating omissions of HeI photoheating and HII recombination cooling, as highlighted in Fig. 3.15). Nevertheless, the approximate solutions seem quite useful and so we include them here.

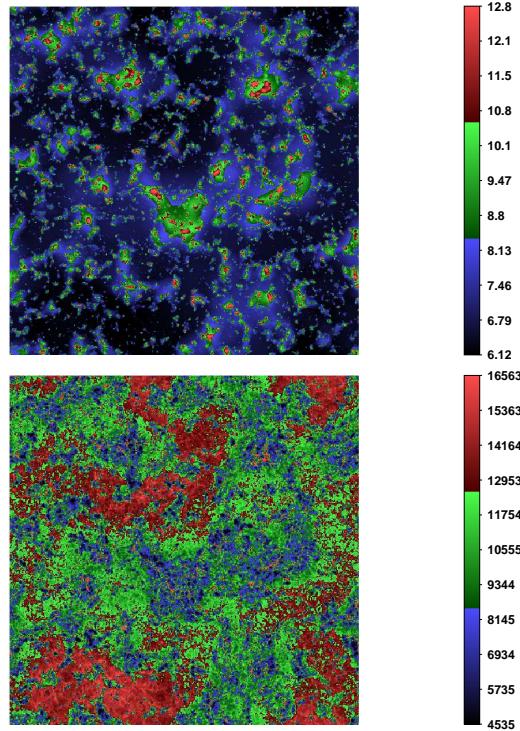


Figure 3.3: Reionization redshifts and temperatures at $z = 5.5$ in the low- z reionization model. *Top panel:* The reionization redshifts for a narrow slice ($0.25 \text{ Mpc}/h$ thick) through the simulation. Each slice is $130 \text{ Mpc}/h$ on a side. The red regions indicate locations with the highest reionization redshifts across the simulation slice, while the dark regions are the last to be reionized. *Bottom panel:* The temperature of the same slice as in the top panel. The red areas in this panel show the hottest locations in the slice, and correspond to the dark regions in the top panel that are reionized late. The dark blue regions in the temperature slice, on the other hand, are the coolest regions that reionized first. The color scales are chosen so that 99% of simulation cells in the slice shown here have redshifts and temperatures falling between the minimum and maximum values on the color bar.

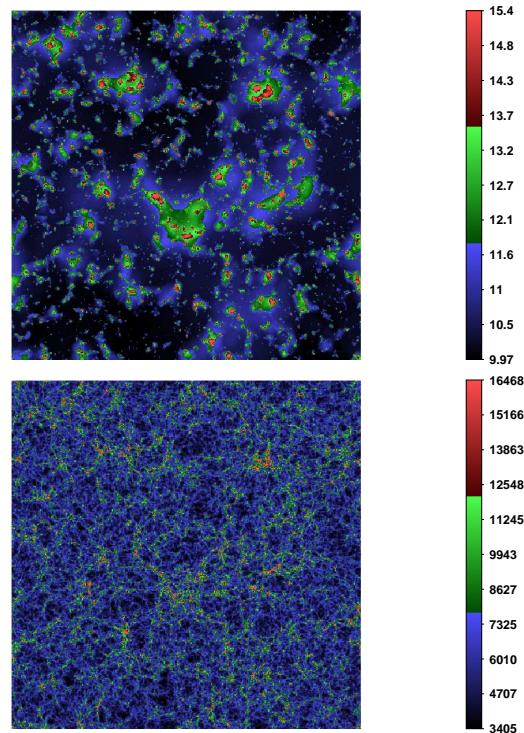


Figure 3.4: Reionization redshifts and temperatures at $z = 5.5$ in the high-z reionization model. Identical to Fig. 3.3, except this figure shows the contrasting High-z model. Note that the color scale in this case also encompasses 99% of the reionization redshifts and temperatures in the simulation slice, but that these ranges are different than in the previous figure.

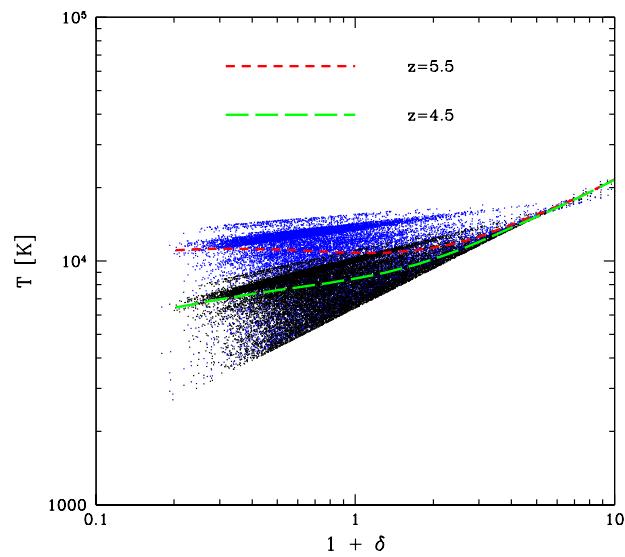


Figure 3.5: Temperature density relations at $z = 4.5$ and $z = 5.5$ in the Low-z reionization model. The blue points show the temperature and density of gas elements from the simulation at $z = 5.5$, while the black points are the same at $z = 4.5$. The red short dashed line shows the median simulated temperature as a function of density at $z = 5.5$. The green long dashed line is the same at $z = 4.5$.

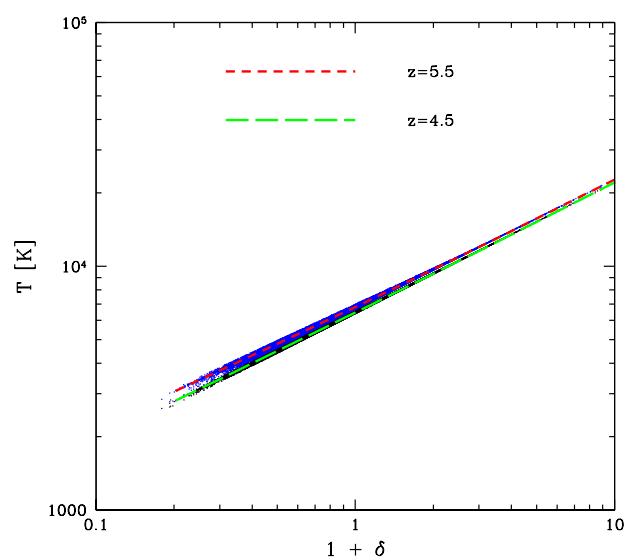


Figure 3.6: Temperature density relations at $z = 4.5$ and $z = 5.5$ in the High-z reionization model. Identical to Fig. 3.5, except the results here are for the High-z reionization model.

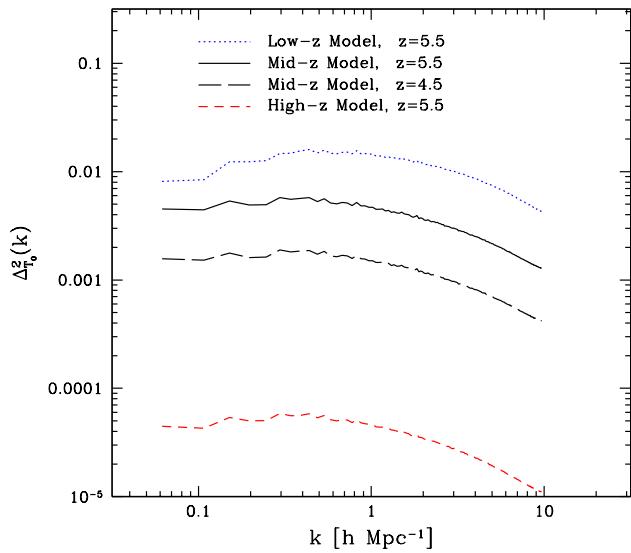


Figure 3.7: Power spectrum of temperature fluctuations in various models. The curves show the power spectrum of $\delta_{T_0}(\mathbf{x}) = (T_0(\mathbf{x}) - \langle T_0 \rangle)/\langle T_0 \rangle$ from the simulated models. The blue dotted line, the black solid line, and the red short-dashed line are the power spectra at $z = 5.5$ in the Low-z, Mid-z, and High-z models respectively. The black long-dashed line shows the δ_{T_0} power spectrum at $z = 4.5$ in the Mid-z model to illustrate how the temperature fluctuations fade with time.

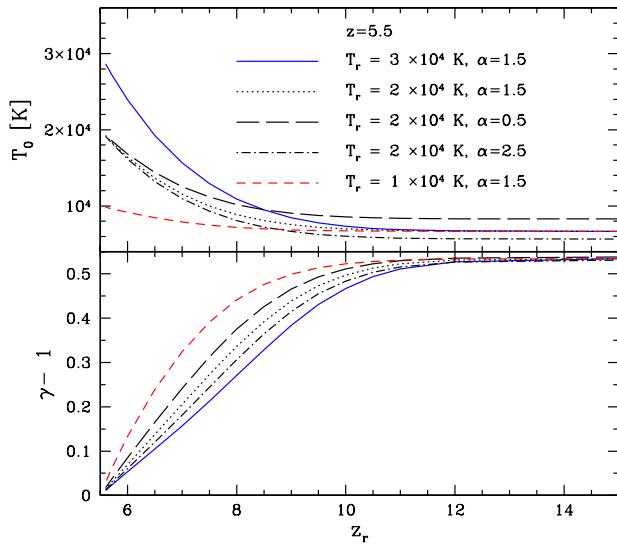


Figure 3.8: Thermal state at $z = 5.5$ for various reionization temperature and spectral shape models. This is similar to the $z = 5.5$ curves in Fig. 3.2, except here we vary the reionization temperature, T_r , and the spectral shape, α . Increasing T_r leads to a higher T_0 and a flatter γ for recently reionized gas parcels, while parcels that reionize at sufficiently high redshifts are insensitive to T_r . A harder ionizing spectrum after reionization (smaller α) leads mostly to a slightly larger value of the asymptotic temperature achieved at high z_r . The harder spectrum also slightly hastens the transition of γ to its asymptotic value.

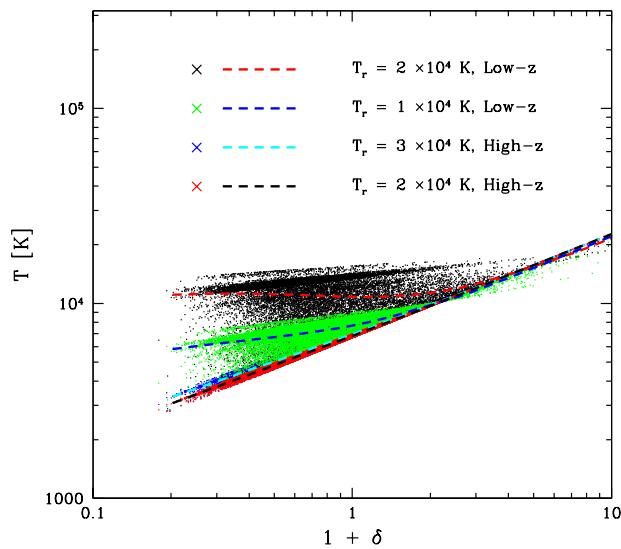


Figure 3.9: Temperature density relation at $z = 5.5$ for various reionization temperatures in the High-z and Low-z models. The “X”’s in the legend indicate the color of the points in the corresponding models, while the dashed lines in the same models have different colors to promote visibility. The models in the legend are listed from top to bottom: the highest points and line (indicating the median temperature at various densities) show the $T_r = 2 \times 10^4$ K, Low-z model; next is the $T_r = 1 \times 10^4$ K, Low-z model; then the $T_r = 3 \times 10^4$ K, High-z model; and finally the $T_r = 3 \times 10^4$ K, High-z model.

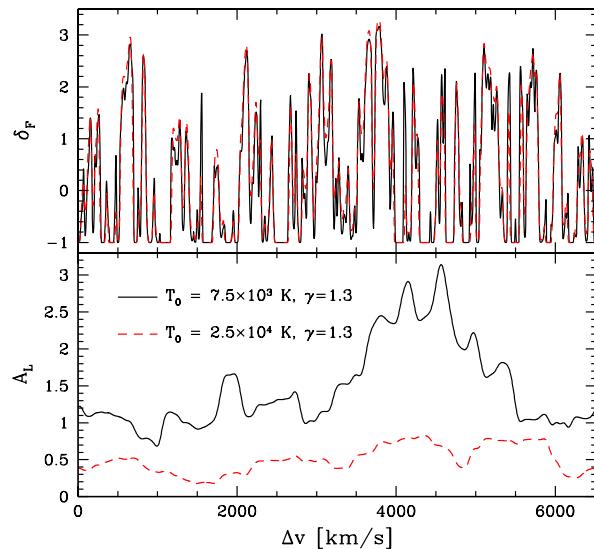


Figure 3.10: Example sightlines and wavelet amplitudes for two different models of the IGM temperature at $z \sim 5$. The top panel shows $\delta_F(x)$ for an example sightline with $T_0 = 2.5 \times 10^4$ K, $\gamma = 1.3$ (red dashed) and the same sightline except with $T_0 = 7.5 \times 10^3$ K, $\gamma = 1.3$ (black solid). The bottom panel shows the smoothed wavelet amplitudes, A_L , along each spectrum. The lower temperature model has more small scale structure and larger wavelet amplitudes. The smoothing scale $s_n = 51$ km/s here, while $\Delta u = 3.2$ km/s and $L = 1,000$ km/s.

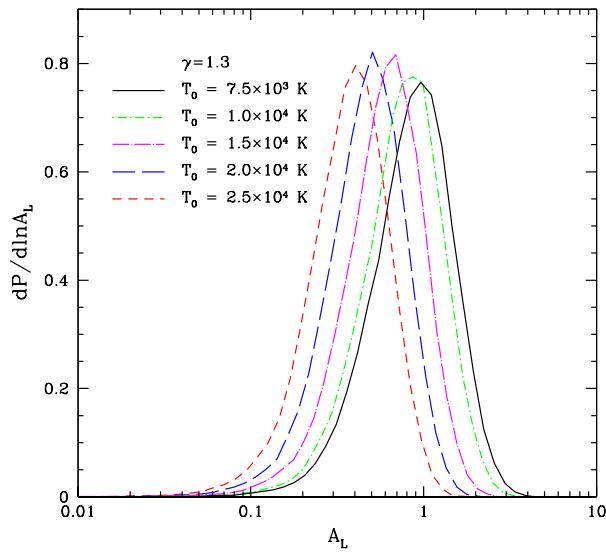


Figure 3.11: Probability distribution of A_L for various T_0 models at $z \sim 5$. Each model here assumes a perfect temperature density relation with $\gamma = 1.3$, and in each case the mean transmitted flux has been fixed – by adjusting the intensity of the ionizing background – to $\langle F_\alpha \rangle = 0.20$. As in Fig. 3.10, the smoothing scale has been set to $s_n = 51 \text{ km/s}$, while $\Delta u = 3.2 \text{ km/s}$ and $L = 1,000 \text{ km/s}$.

3.6 Conclusions

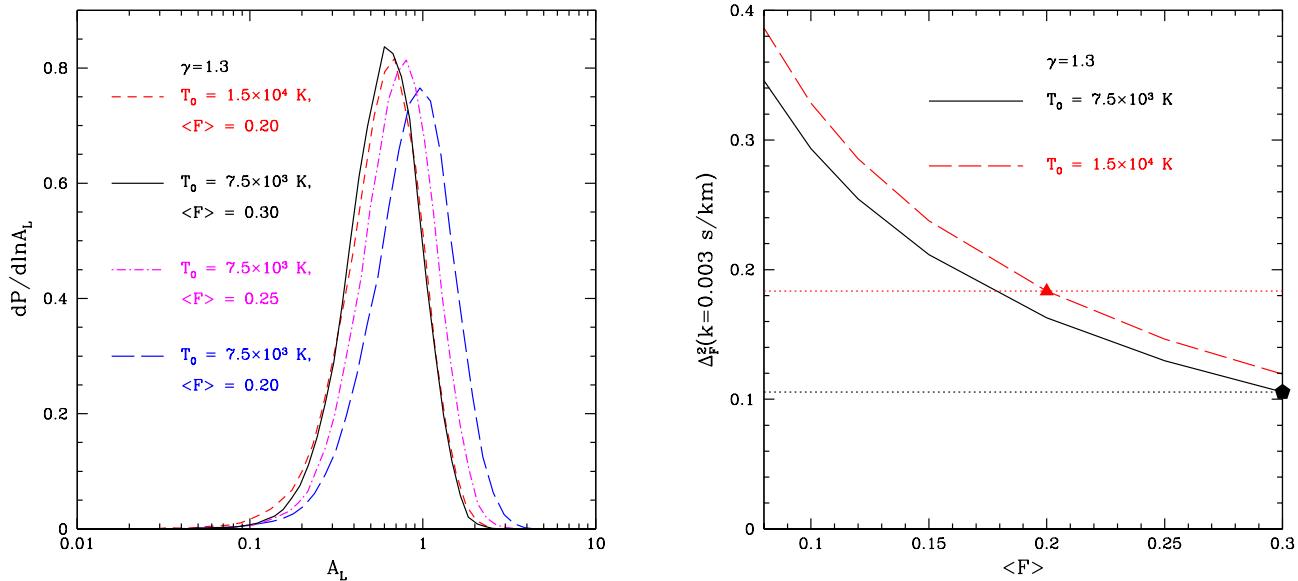


Figure 3.12: Degeneracy with $\langle F \rangle$. *Top panel:* Although the PDF of A_L is sensitive to T_0 , this effect is degenerate with the impact of varying $\langle F \rangle$. For instance, the model with $T_0 = 1.5 \times 10^4$ K and $\langle F \rangle = 0.20$ is closely mimicked by a colder model with $T_0 = 7.5 \times 10^3$ K, yet a larger mean transmission of $\langle F \rangle = 0.30$. *Bottom panel:* This illustrates that the degeneracy can be broken by measuring the (relatively) large scale flux power spectrum. The curves here show the flux power spectrum, evaluated at a single convenient (larger-scale) wavenumber of $k = 0.003$ s/km, in each T_0 model as a function of $\langle F \rangle$. The triangle and pentagon show the flux power for each model at the $\langle F \rangle$ for which the wavelet amplitude PDFs are degenerate in the two models. The large scale flux power in these two models differs appreciably and can be used to break the degeneracy. The red dotted and black dotted horizontal lines are intended only to guide the eye.

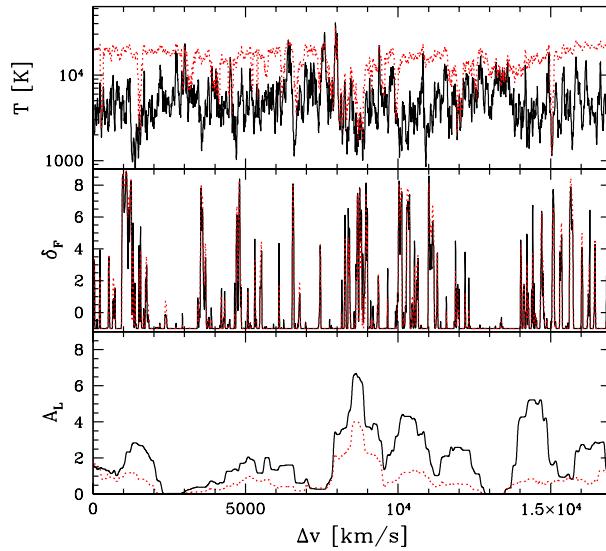


Figure 3.13: Example sightlines and wavelet amplitudes from the Low-z and High-z reionization models. In the models here, the global mean flux is $\langle F \rangle = 0.1$ and $z = 5.5$. In each panel the red dotted line shows a sightline through the $T_r = 3 \times 10^4$ K, Low-z reionization model while the black solid line is the same sightline, except in this case the temperature field is drawn from the High-z reionization model (with $T_r = 2 \times 10^4$ K). The simulated density and temperature fields have small scale structure added according to the lognormal model, as described in the text. *Top panel:* The simulated temperature field. *Middle panel:* The transmission field, δ_F . *Bottom panel:* The smoothed wavelet amplitude with $L = 1,000$ km/s, $s_n = 34$ km/s, and $\Delta u = 2.1$ km/s. The transmission fluctuations and wavelet amplitudes are larger than in Fig. 3.10, mostly because of the lower mean transmitted flux adopted here.

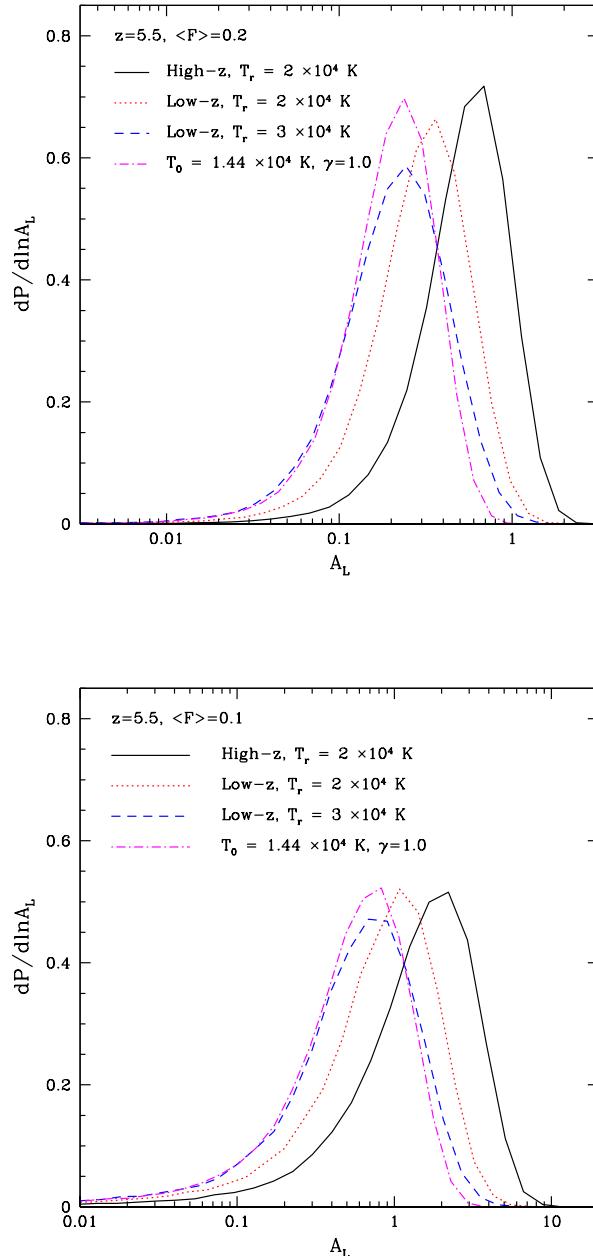


Figure 3.14: Probability distribution of A_L for various reionization and temperature models at $z = 5.5$. *Top panel:* In this panel all models are normalized to $\langle F \rangle = 0.2$. The solid black curve shows the wavelet amplitudes for the High-z reionization model (with $T_r = 2 \times 10^4$ K), while the red dotted and blue dashed curves show Low-z reionization models with reionization temperatures of $T_r = 2 \times 10^4$ K and $T_r = 3 \times 10^4$ K respectively. The magenta dot-dashed line shows a *homogeneous* temperature model for comparison. In this case, the temperature was set to match the median temperature in the Low-z, $T_r = 3 \times 10^4$ K model for gas at the cosmic mean density; the broader distribution in the Low-z model reflects the impact of inhomogeneous

3.6 Conclusions

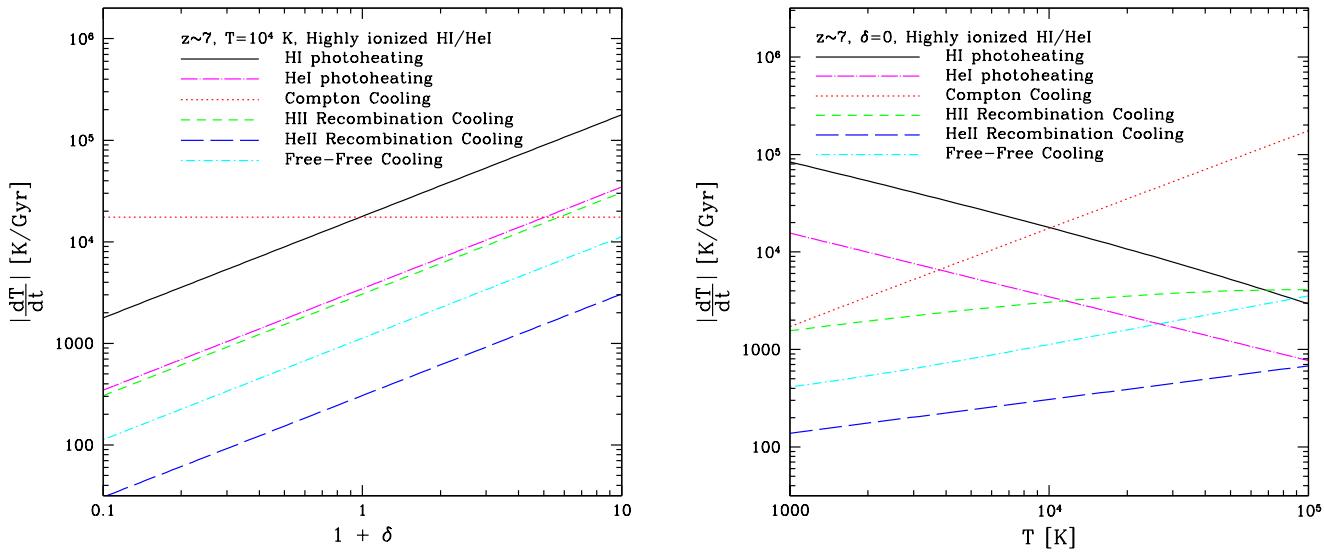


Figure 3.15: Heating/cooling rates at $z \sim 7$. *Left panel*: The (absolute value of) the rates for relevant processes in the IGM at $T = 10^4$ K as a function of density, assuming that hydrogen is highly ionized and that helium is mostly singly-ionized. *Right panel*: Similar to the left panel except the rates are shown as a function of temperature for gas at the comic mean density.

Chapter 4

Identifying Ionized Regions in Noisy Redshifted 21-cm Observations

4.1 Introduction

The Epoch of Reionization (EoR) is the time period when early generations of galaxies first turn on and gradually photoionize neutral hydrogen gas in the surrounding intergalactic medium (IGM). The IGM is expected to resemble a two phase medium during reionization. One phase consists of highly ionized regions, termed ‘ionized bubbles’, that form around clustered groups of ionizing sources, while the other phase is made up of intervening mostly neutral regions that shrink and eventually vanish as reionization progresses. A primary goal of reionization studies is to determine the size distribution and volume-filling factor of the ionized bubbles. This should, in turn, significantly improve our understanding of high redshift galaxy and structure formation. A wide variety of current observations have started to provide tantalizing hints regarding the timing and nature of the EoR (e.g., Fan et al. 17, Totani et al. 49, ? , ? , ? , ? , ? , ?), but we still await a more detailed

4.1 Introduction

understanding.

A highly anticipated way of improving our knowledge of the EoR is to directly detect intergalactic neutral hydrogen from the EoR using the redshifted 21 cm transition (e.g., ?, ?, ?, ?). Indeed, several radio telescopes have been constructed, or are presently under construction, in effort to detect this signal, including the Giant Metrewave Radio Telescope (GMRT) (?), the Low Frequency Array (LOFAR) (?), the Murchison Widefield Array (MWA) (?), and the Precision Array for Probing the Epoch of Reionization (PAPER) (?). This method provides the most direct, and potentially most powerful, way of studying reionization, but several challenges need first to be overcome. In particular, upcoming surveys will need to extract the faint cosmological signal in the presence of strong foreground emission from our own galaxy and extragalactic point sources, and to control systematic effects from the instrumental response, polarization leakage, calibration errors, and other sources of contamination (e.g., ?, ?, ?, ?, ?, ?, ?, ?). In addition, thermal noise will prevent early generations of 21 cm experiments from making detailed maps of the reionization process. Instead, detections will mostly be of a statistical nature (?). For example, a primary goal of these experiments is to measure the power spectrum of 21 cm brightness temperature fluctuations by binning together many individually noisy Fourier modes (? , ?, ?, ?, ?).

It is unclear, however, how best to analyze the upcoming redshifted 21 cm data. Most previous work has focused only on the power spectrum of 21 cm brightness temperature fluctuations (e.g., ?, ?, ?, ?). This statistic does not provide a complete description of the 21 cm signal from the EoR, which will be highly non-Gaussian, with large ionized regions of essentially zero signal intermixed with surrounding neutral regions. The power spectrum, and especially its redshift evolution, do encode interesting information about the volume-averaged ionized fraction and the bubble size distribution (e.g., ?). However, these inferences are somewhat indirect and likely model dependent, and so it is natural to ask if there are more direct ways of determining the properties of the ionized regions.

The approach we explore here is to check whether it may be possible to directly identify ionized regions in noisy redshifted 21 cm observations by applying suitable filters to the noisy data. Our aim here is to blindly identify ionized bubbles across an entire survey volume, rather than to consider targeted searches around special regions, such as those containing known quasars (e.g., ? , ?). Since the 21 cm signal from reionization is expected to have structure on rather large scales – $\gtrsim 30 h^{-1}\text{Mpc}$ co-moving (Furlanetto et al. 21, ? , ? , McQuinn et al. 32) – it may be possible to make crude images of the large scale features even in the regime where the signal to noise per resolution element is much less than unity. Even if it is only possible to identify a few unusually large ionized regions in upcoming data sets, this would still be quite valuable. Any such detection would be straightforward to interpret, and would open-up several interesting possibilities for follow-up investigations. Towards this end, we extend previous work by (?) and (?), who considered the prospects for detecting ionized regions using an optimal matched filter. A matched filter is constructed by correlating a known ‘template’ signal with a noisy data set in order to determine whether the template signal is present in the noisy data. Matched filters are used widely in astrophysics: to name just a few examples, matched filters are used to detect clusters in cosmic microwave background (CMB) data (?), to identify galaxy clusters from weak lensing shear fields (e.g., ? , ?), and are central to data analysis efforts aimed at detecting gravitational waves (e.g., ?).

The outline of this paper is as follows. In §4.2 we describe the mock 21 cm data sets used in our investigations. We use the mock data to first consider the ability of future surveys to make maps of the redshifted 21 cm signal (§4.3). In §4.4, we then quantify the prospects for identifying individual ionized regions using a matched filter technique. In §4.5 and §4.6 we consider variations around our fiducial choice of reionization history and redshifted 21 cm survey parameters. We compare with previous related work in §4.7, and conclude in §4.8. Throughout we consider a ΛCDM cosmology parametrized by $n_s =$

$1, \sigma_8 = 0.8, \Omega_m = 0.27, \Omega_\Lambda = 0.73, \Omega_b = 0.046$, and $h = 0.7$, (all symbols have their usual meanings), consistent with the latest WMAP constraints from (?).

4.2 Method

Briefly, our approach is to construct mock redshifted 21 cm data sets and check whether we can successfully identify known ‘input’ ionized regions in the presence of realistic levels of instrumental noise and the degrading impact of foreground cleaning. Here we describe the ingredients of our mock data sets: our simulations of reionization and the 21 cm signal, our model for thermal noise, and our approach for incorporating the impact of foreground cleaning.

4.2.1 The 21 cm Signal

First, let us describe the underlying 21 cm signal and our reionization simulations. The 21 cm signal will be measured through its contrast with the cosmic microwave background (CMB). The brightness temperature contrast between the CMB and the 21 cm line from a neutral hydrogen cloud with neutral fraction x_{HI} and fractional baryon overdensity δ_ρ is (?):

$$\delta T_b = 28x_{\text{HI}}(1 + \delta_\rho) \left(\frac{T_S - T_\gamma}{T_S} \right) \left(\frac{1+z}{10} \right)^{1/2} \text{ mK}. \quad (4.1)$$

Here T_γ denotes the CMB temperature and T_S is the spin temperature of the 21 cm line. Here and throughout we neglect effects from peculiar velocities, which should be a good approximation at the redshifts and neutral fractions of interest (e.g., ?, ?). Furthermore, throughout we assume that the spin temperature is globally much larger than the CMB temperature, i.e., we assume that $T_S \gg T_\gamma$. In this case the 21 cm signal appears in emission and the brightness temperature contrast is independent of T_S . This is expected to be a good approximation for the volume-averaged ionized fractions of interest for our present study,

although it will break down at earlier times (e.g., ?). With these approximations,

$$\delta T_b = T_0 x_{\text{HI}}(1 + \delta_\rho), \quad (4.2)$$

where $T_0 = 28 [(1+z)/10]^{1/2} \text{mK}$. Throughout this paper, we refer to the brightness temperature contrast in units of T_0 .

4.2.2 Semi-Numeric Simulations

In order to simulate reionization we use the ‘semi-numeric’ scheme described in (?) (see also e.g., ?, for related work and extensions to this technique). This scheme is essentially a Monte Carlo implementation of the analytic model of (21), which is in turn based on the excursion set formalism. The (?) algorithm allows us to rapidly generate realizations of the ionization field over large simulation volumes at various stages of the reionization process. The results of these calculations agree well with more detailed simulations of reionization on large scales (??).

We start by generating a realization of the linear density field in a simulation box with a co-moving side length of $1 h^{-1} \text{Gpc}$ and 512^3 grid cells. The ionization field, x_i , is generated following the algorithm of (?), assuming a minimum host halo mass of $M_{\min} = 10^8 M_\odot$, comparable to the atomic cooling mass at these redshifts (?). Each halo above M_{\min} is assumed to host an ionizing source, and the ionizing efficiency of each galaxy is taken to be independent of halo mass. In our fiducial model, we adjust the ionizing efficiency so that the volume-averaged ionization fraction is $\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$. We focus most of our analysis on this redshift and on this particular model for the volume-averaged ionized fraction. However, we consider additional redshifts in §4.5, as well as variations around our fiducial ionization history in effort to bracket current uncertainties in the ionization history (see e.g., ?, ?).

From the linear density field and the ionization field we generate the 21 cm brightness temperature contrast following Equation 4.2. Using the linear density field here – rather

than the evolved non-linear density field – should be a good approximation for the large scales of interest for our study; we focus on length scales of $R \gtrsim 20 h^{-1}\text{Mpc}$ and high redshift ($z \gtrsim 6$) in subsequent sections.

4.2.3 Redshifted 21 cm Surveys and Thermal Noise

We mostly consider two concrete examples of upcoming redshifted 21 cm surveys. The first is based on the current, 128-tile version of the MWA (?) and the second is based on an expanded, 500-tile version of the MWA (as described in ?, and considered in previous work such as ? ?). These two examples are intended to indicate the general prospects for imaging and bubble identification with first and second generation 21 cm surveys, respectively. Similar considerations would apply for other experiments, but we choose these as a concrete set of examples. We mainly focus on the 500-tile configuration in this paper because of its greater sensitivity. In §4.5.4, we shift to considering 128-tile configurations and in §4.6, we consider a LOFAR-style interferometer for comparison. Hereafter, we refer to the 500-tile configuration as the MWA-500 and the 128-tile version as the MWA-128.

Throughout this paper, we work in co-moving coordinates described by Cartesian labels (x - y - z), with Fourier counterparts (k_x - k_y - k_z). The Fourier modes can be connected directly with the u - v - ν coordinate system generally used to describe interferometric measurements. Here u and v describe the physical separation of a pair of antennae in units of the observed wavelength, while ν describes the corresponding observed frequency. The instrument makes measurements for every frequency, ν , in its bandwidth, and for every antenna tile separation, (u, v) , sampled by the array. In order to shift to a Fourier space description, the interferometric measurements must first be Fourier-transformed along the frequency direction. With our Fourier convention, the relation between the two sets of coordinates is given by:

$$k_x = \frac{2\pi u}{D} \quad k_y = \frac{2\pi v}{D} \quad k_z = \frac{2\pi}{\Delta\chi}, \quad (4.3)$$

4.2 Method

where D is the co-moving distance to the survey center and $\Delta\chi$ is the co-moving distance corresponding to a small difference in observed frequency of $\Delta\nu$ (e.g., ?). For small $\Delta\nu/\nu$, we can express $\Delta\chi$ as

$$\Delta\chi \approx \frac{c(1+z_{\text{fid}})}{H(z_{\text{fid}})} \frac{|\Delta\nu|}{\nu}, \quad (4.4)$$

where $H(z_{\text{fid}})$ is the Hubble parameter at the fiducial redshift, and $|\Delta\nu|/\nu$ is the absolute value of the fractional difference between two nearby observed frequencies.

In order to test the prospects for imaging and bubble identification with the MWA, we must corrupt the underlying 21 cm signal described in §4.2.2 with thermal noise. We do this by generating a Gaussian random noise field in the \mathbf{k} -space coordinate system described above, using an appropriate power spectrum. We assume that the co-variance matrix of the thermal noise power is diagonal in \mathbf{k} -space. We add the resulting noise field to the underlying 21 cm signal (Equation 4.2). The power spectrum of the thermal noise is given by (? , ?):

$$P_N(k, \mu) = \frac{T_{\text{sys}}^2}{B t_{\text{int}}} \frac{D^2 \Delta D}{n(k_\perp)} \left(\frac{\lambda^2}{A_e} \right)^2. \quad (4.5)$$

Here μ is the cosine of the angle between wavevector $k = |\mathbf{k}|$ and the line of sight, so that $k_\perp = \sqrt{1 - \mu^2}k$ is the transverse component of the wavevector. We assume a system temperature of $T_{\text{sky}} = 280[(1+z)/7.5]^{2.3}$ K (?) and a total observing time of $t_{\text{int}} = 1000$ hours, which is an optimistic estimate for the observing time in one year. At our fiducial redshift of $z_{\text{fid}} = 6.9$, the co-moving distance to the center of the survey is $D = 6.42 \times 10^3 h^{-1}\text{Mpc}$. In this equation, λ denotes the observed wavelength of the redshifted 21 cm line, $\lambda = 0.211(1+z)\text{m}$, and A_e is the effective area of each antenna tile. We determine A_e by linearly extrapolating or interpolating from the values given in Table 2 of (?); the effective area at $z_{\text{fid}} = 6.9$ is $A_e = 11.25\text{m}^2$. We assume that the full survey bandwidth of 32 MHz is broken into individual blocks of bandwidth $B = 6$ MHz to protect against redshift evolution across the analysis bandwidth (?). The co-moving survey depth depth corresponding to a $B = 6$ MHz chunk is $\Delta D = 69 h^{-1}\text{Mpc}$. The $n(k_\perp)$ term describes the configuration of the

antenna tiles. More specifically, it is the number density of baselines observing modes with transverse wavenumber k_{\perp} (?). Following (?) and (?), we assume the antenna tiles are initially packed as closely as possible in a dense compact core, and that the number density of antenna tiles subsequently falls off as r^{-2} out to a maximum baseline of 1.5 km. The radius of the dense antenna core is set by the requirement that the antenna density falls off as r^{-2} outside of the core, and that it integrates to the total number of antennae. For the MWA-500, this gives $r_c = 20$ m, while for the MWA-128, the core radius is $r_c \approx 8$ m. Equation 4.5 gives the noise power spectrum in units of mK², and so we divide by T_0^2 to combine with the simulated 21 cm signal expressed in units of T_0 .

Note that the volume of the MWA survey differs somewhat from that of our reionization simulation. In particular, the transverse dimension of the simulation is smaller than that of the MWA by a factor of ~ 3 , while the simulation is deeper in the line-of-sight direction by about the same factor, as compared with the full MWA bandwidth. However, we remove the long wavelength modes along the line-of-sight direction to mimic foreground cleaning (§4.2.4), and so we do not, in practice, use the longer line-of-sight scales in our simulation box. As we will see, the ionized regions in the simulation are substantially smaller than the transverse length of the box. Transverse slices should therefore be representative of what the actual MWA will observe from a fraction of its larger field of view. We have checked that the coarser transverse k -space sampling in the simulation compared to in the actual MWA survey does not impact our results.

4.2.4 Foregrounds

Next, we need to consider contamination from foreground emission at the frequencies of interest. The relevant foregrounds include diffuse Galactic synchrotron radiation, extragalactic point sources, and Galactic Bremsstrahlung radiation. While these foregrounds are many orders of magnitude brighter than the cosmological 21 cm signal, they are expected to individually follow smooth power laws in frequency. Over a sufficiently small frequency

range, the summed contributions can also be approximated as following a smooth power law, while the 21 cm signal will vary rapidly. This allows the foregrounds to be removed from the data by, for example, fitting a low-order function along each line of sight and subtracting it. While this procedure is effective at removing foreground contamination, it also removes long wavelength modes along the line of sight from the signal itself, and hence prevents measuring these modes. Several related methods for foreground removal have been discussed in the literature (e.g., ? , ? , ? , ?). In this work, we approximately mimic the degrading effects from foreground removal by subtracting the running mean from the noisy signal along each line of sight, rather than including realizations of the foregrounds in our simulation and excising them with one of the above algorithms. We generally remove the running mean over a bandwidth of 16 MHz, which corresponds to a co-moving distance of $L_{\text{fg}} = 185 h^{-1} \text{ Mpc}$ at redshift $z_{\text{fid}} = 6.9$; we consider the impact of other choices of L_{fg} in §4.5.3. We defer more detailed models of foreground contamination, and foreground removal algorithms, to future work.

4.3 Prospects for Imaging

Having described our mock 21 cm data sets, we now turn to consider the prospects for constructing direct ‘images’ of the redshifted 21 cm signal. Previous studies already suggest that the prospects for imaging with the MWA-500 are limited (?). Here we emphasize that even a crude, low-resolution image of the redshifted 21 cm signal may be quite interesting, especially given that the ionized regions during reionization may be rather large scale features. We hence seek to quantify the imaging capabilities further using our corrupted reionization simulations. Here our work complements recent work in a similar vein by (?), who considered the prospects for imaging with LOFAR. While the central idea in this section is similar to this previous work, we focus on the MWA while (?) considered LOFAR. In order to construct the best possible images from the noisy mock 21 cm data, we apply a Wiener filter. We assess the ability of the MWA to image the redshifted 21 cm sky by

4.3 Prospects for Imaging

comparing the filtered (recovered) noisy signal with the underlying noise-free 21 cm input signal.

4.3.1 The Wiener Filter

The Wiener filter is the optimal filter for extracting an input signal of known power spectrum when it is corrupted by additive noise, also with known power spectrum. As described in (?), this filter is optimal in that it minimizes the expectation value of the integrated squared error between the estimated signal field and the true signal field. The estimate of the true signal is a convolution of the Wiener filter and the corrupted signal in real space, and so is a product of the two quantities in Fourier space,

$$\tilde{S}(\mathbf{k}) = C(\mathbf{k})W(\mathbf{k}), \quad (4.6)$$

where $C(\mathbf{k})$, $W(\mathbf{k})$, and $\tilde{S}(\mathbf{k})$ are the Fourier transforms of the corrupted signal, Wiener filter, and estimated signal, respectively. Requiring that the filter be optimal in the least-square sense results in $W(\mathbf{k})$ taking the form

$$W(k, \mu) = \frac{P_S(k)}{P_S(k) + P_N(k, \mu)}, \quad (4.7)$$

where $P_S(k)$ and $P_N(k, \mu)$ are the power spectra of the signal and noise, respectively. We note that, while the signal power spectrum is roughly isotropic¹, the noise power spectrum depends on μ and consequently so does the filter. The filter keeps a unity weighting for k -modes where $P_S(k) \gg P_N(k)$ and significantly downweights k -modes where $P_S(k) \ll P_N(k)$. This can allow for partial recovery of the original signal, provided that the signal power dominates for some k -modes.

The Wiener filter requires an estimate of the signal power spectrum, $P_S(k)$, and of the total (signal plus noise) power spectrum, $P_S(k) + P_N(k, \mu)$, as inputs. These may not be

¹Redshift-space distortions and redshift evolution across the observed bandwidth break isotropy (e.g. ?). However, for the bandwidth considered here ($B = 6$ MHz) and the neutral fractions of interest, the signal should be approximately isotropic.

4.3 Prospects for Imaging

precisely known. However, since the filter is the outcome of a minimization problem – i.e., it minimizes the expected difference between the estimated and true fields – the accuracy of the filter should be insensitive to small changes about its optimal value. In other words, the accuracy of the filter is not expected to change greatly by using estimates of the signal and noise power spectra rather than the true spectra.

Furthermore, we do expect to have an estimate of the underlying signal power spectrum; measuring this statistic is a major goal of redshifted 21 cm surveys. Specifically, the underlying signal power can be estimated by cross-correlating redshifted 21 cm measurements made over two different time intervals (after foregrounds have been removed). The statistical properties of the signal should be identical across the two different time periods, but the thermal noise contributions will be independent. The cross-correlation between two time chunks then provides an unbiased estimate of the signal power (e.g., ?). Estimates of the noise power spectrum can then be made by subtracting the estimated signal power from the power measured over the entire integration time, which contains both the signal and noise contributions. The Wiener filter does not actually require the noise power spectrum to be known on its own. However, in §4.4 we consider the optimal matched filter, which does have this requirement. Throughout this study, we assume perfect knowledge of the underlying power spectra.

Before applying the Wiener filter to our corrupted simulations, it is useful to estimate the expected signal-to-noise ratio of the filtered maps analytically, using simulated signal power spectra and the noise power spectrum of Equation 4.5. The expected signal-to-noise ratio of the Wiener-filtered field is $\mathcal{S}_{\text{wf}} = \tilde{\sigma}_S / \tilde{\sigma}_N$, where $\tilde{\sigma}_{S(N)}^2$ is the filtered signal (noise) variance. The signal and noise variance can in turn be calculated as integrals over their respective power spectra,

$$\tilde{\sigma}_{S(N)}^2 = \int \frac{d^3k}{(2\pi)^3} |W(k, \mu)|^2 P_{S(N)}(k, \mu). \quad (4.8)$$

4.3 Prospects for Imaging

Here we use $P_{S(N)}$ to denote the power spectrum of the signal (noise). One can also consider the impact of foreground cleaning here by downweighting modes where the foreground power is large compared to the signal power. In order to consider the dependence of the signal-to-noise ratio on the stage of reionization, we consider simulation outputs in which the volume-averaged ionization fraction is $\langle x_i \rangle = 0.51, 0.68, 0.79$ and 0.89 . We consider each of these models at our fiducial redshift of $z_{\text{fid}} = 6.9$.¹ Presently, we don't consider still earlier stages of reionization since the prospects for imaging with the MWA-500 are especially poor for lower ionized fractions.

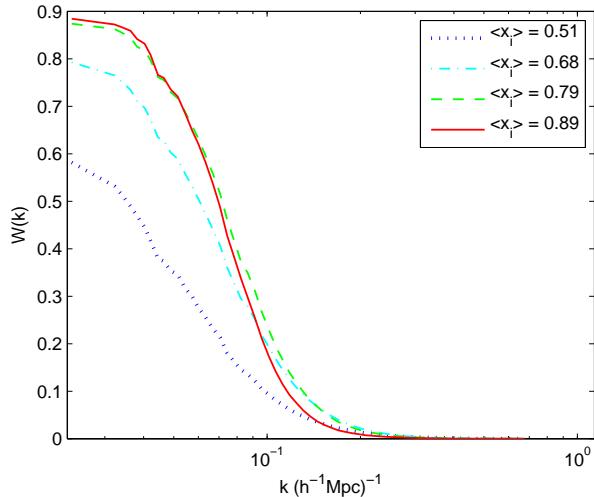


Figure 4.1: Fourier profile of the Wiener filter, $W(k)$. The filter is averaged over line-of-sight angle and the results are shown at $z_{\text{fid}} = 6.9$ for simulated models with $\langle x_i \rangle = 0.51$ (blue dotted), $\langle x_i \rangle = 0.68$ (cyan dot-dashed), $\langle x_i \rangle = 0.79$ (green dashed), and $\langle x_i \rangle = 0.89$ (red solid).

¹In practice, the simulated ionization fields for ionized fractions lower (higher) than our fiducial value ($\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$) come from slightly higher (lower) redshift simulation outputs. We generate the 21 cm signal and noise as though each data cube were in fact at $z_{\text{fid}} = 6.9$. This is appropriate to the extent that the statistical properties of the ionized regions are mainly determined by the volume-averaged ionized fraction, and are relatively insensitive to the precise redshift at which a given ionized fraction is reached (see McQuinn et al. 32 and Furlanetto et al. 21.)

4.3 Prospects for Imaging

The resulting Wiener filters for the different values of $\langle x_i \rangle$ are shown in Figure 4.1, after integrating over angle μ . In this figure, foreground cleaning has been accounted for by subtracting a running mean along the line of sight, as described in §4.2.4. It is helpful to note, from Equation 4.7, that the filter is equal to 1/2 for modes where the signal and noise power are equal. The figure suggests that a small range of k -modes with $k \lesssim 0.1 h \text{ Mpc}^{-1}$ will have signal-to-noise ratio larger than unity for all four ionized fractions considered, although imaging is less promising for the smaller ionized fractions. If the ionized regions are larger than in our fiducial model – as expected if, for example, rarer yet more efficient and more clustered sources dominate reionization (e.g., McQuinn et al. 32, ?) – then the prospects for imaging may improve somewhat. Performing the integrals in Equation 4.8, while incorporating foreground cleaning, we find that the total signal-to-noise ratio expected for the MWA-500 is $\delta_{\text{wf}} = 0.52, 0.79, 1, \text{ and } 1.2$ for $\langle x_i \rangle = 0.51, 0.68, 0.79, 0.89$, respectively.

4.3.2 Application to a Simulated 21 cm Signal

With the analytic signal-to-noise ratio estimates as a guide, we apply the Wiener filter to our mock noisy redshifted 21 cm data. The results of these calculations, for a particular slice through the simulation volume, are shown in Figure 4.2. The side length ($1 h^{-1} \text{Gpc}$) of each slice is a factor of ~ 3 smaller than the transverse dimension of the MWA. One can assess how well the original signal is ‘recovered’ by comparing the top-left panel of the figure which shows the input signal with the bottom-left panel which shows the filtered noisy signal, after mimicking foreground removal. The two panels do not bear a striking resemblance since the average signal-to-noise ratio is only of order unity. Nonetheless, it is encouraging that many of the minima in the filtered noisy signal do indeed correspond to ionized regions in the input signal. Furthermore, we can compare the filtered noisy signal in the bottom-left panel with the top-right panel, which shows filtered pure noise. While these two panels do not look drastically different, they are easily distinguishable from each other given the increased contrast in the filtered noisy signal. In addition, we see that

4.3 Prospects for Imaging

the filtered noisy signal obtains signal-to-noise values exceeding $6 - \sigma$, while the statistical significance of the filtered noise does not exceed $\sim 5 - \sigma$. Quantitatively, $\sim 3\%$ ($\sim 0.03\%$) of the volume in the filtered noisy signal is occupied by pixels with statistical significance greater (in magnitude) than $3 - \sigma$ ($5 - \sigma$). This is expected given that the filtered data cube has an average signal-to-noise ratio of $\sigma_S/\sigma_N \approx 1$, as anticipated in the analytic calculation of §4.3.1.

Comparing the filtered noisy signal and the filtered pure noise, one can see that ionized regions in the underlying signal are diminished if they happen to be coincident with upward fluctuations in the noise, as expected. For example, the ionized region in the bottom-right corner of the unfiltered signal lies very close to a $\sim 3 - \sigma$ upward fluctuation in the filtered noise and, as a result, appears with weak statistical significance in the filtered noisy signal. Conversely, some of the most statistically significant regions in the filtered noisy signal occur when ionized regions overlap downward noise fluctuations. We can further compare the filtered noisy signal with the filtered noise-*less* signal, shown in the bottom right panel of Figure 4.2. The filtered noise-less signal is normalized by the standard deviation of the filtered noise so that it can be compared with the signal-to-noise slices in the other panels. This comparison reveals that high significance regions ($\gtrsim 5 - \sigma$) in the filtered noisy signal only line up well with the corresponding regions in the filtered noiseless signal if they are coincident with downward fluctuations in the noise. On its own, the filtered noiseless signal only attains statistical significances of $\lesssim 4\sigma_N$. Finally, Figure 4.3 illustrates the impact of foreground cleaning, performed here over a bandwidth of 16 MHz (§4.2.4). Foreground cleaning removes the long wavelength modes along the line of sight – which is along the vertical axis in the figure – and thereby compresses structures along the line of sight. However, the cleaning process only impacts the long wavelength line-of-sight modes which still leaves room to image other modes robustly.

Note that the slice thickness ($8 h^{-1}\text{Mpc}$) in Figure 4.2 and 4.3 is somewhat arbitrary. However, the Wiener filter smooths out structure on significantly larger scales than this

4.4 Prospects for Identifying Ionized Regions

(Figure 4.1), and so we expect similar results for other values of the slice thickness, provided the slice is thin compared to the cut-off scale of the filter. In practice, of course, one can make many independent maps similar to Figure 4.2 from the MWA-500 or similar surveys. Collectively, our results mostly confirm previous wisdom; the prospects for imaging with the MWA-500 are limited. Nonetheless, it appears that a signal-to-noise ratio of order unity is achievable, suggesting that the MWA-500 *can* make low resolution maps of the reionization process.

4.4 Prospects for Identifying Ionized Regions

We now shift our focus to discuss whether it may also be possible to identify interesting individual features in upcoming 21 cm data cubes. In particular, we aim to identify ionized regions in noisy 21 cm data sets and, furthermore, to estimate the spatial center and approximate size of each ionized bubble. For this purpose, we will use an optimal matched filter technique. As we discuss, individual ionized regions may be identifiable as *prominent minima in the filtered field*.

4.4.1 The Optimal Matched Filter

The optimal matched filter is suited for the case of a corrupted signal containing a known feature that one would like to extract. The filter acts in Fourier space by cross-correlating the corrupted signal with a template describing the known feature, while downweighting k -modes in the corrupted signal by the noise power. The resulting form of the filter in Fourier space, $M(k, \mu)$, is

$$M(k, \mu) = \frac{T(\mathbf{k})}{P_N(k, \mu)}, \quad (4.9)$$

where $T(\mathbf{k})$ is the Fourier profile of the known feature. The filter is optimal in the sense that it maximizes the signal-to-noise ratio in the filtered data cube at the location of the feature being extracted. While the Wiener filter requires an estimate of the signal and total (signal

4.4 Prospects for Identifying Ionized Regions

plus noise) power spectra, the matched filter requires a good estimate of the template profile, $T(\mathbf{k})$, and the noise power spectrum, $P_N(k, \mu)$. For our present application, we would like templates describing the ionized regions. An appropriate choice is not obvious; theoretical models predict that the ionization state of the gas during reionization has a complex, and somewhat uncertain, morphology, with ionized regions of a range of sizes and shapes (? , ?, McQuinn et al. 32). However, we find that the simplest conceivable choice of template filters, corresponding to completely ionized spherical bubbles of varying size, are nonetheless effective at identifying ionized regions with a more realistic and complex morphology. In this case, $T(\mathbf{k}; R)$ is just the Fourier transform of a spherical top-hat of radius R and is given by

$$T(\mathbf{k}; R_T) = \frac{V}{k^3 R_T^3} [-kR_T \cos kR_T + \sin kR_T], \quad (4.10)$$

with V denoting the volume of the spherical top-hat. Note that the precise normalization of the filter is unimportant since we are mainly interested in the signal-to-noise ratio here, in which case the overall normalization divides out.

4.4.2 Application to Isolated Spherical Ionized Regions with Noise

It is instructive to first consider an idealized test case that can be treated analytically before applying the matched filter to our full mock 21 cm data sets. In particular, we consider the case of an isolated, spherical, and highly ionized region placed at the origin and embedded in realistic noise. We assume that the neutral fraction exterior to the ionized region is uniform, with a mass-weighted neutral fraction of $\langle x_{\text{HI}}(1 + \delta_\rho) \rangle$. Ignoring foreground contamination for the moment, the 21 cm signal may be written as:

$$\delta T_b(\mathbf{x}) - \langle \delta T_b \rangle = \tilde{B}(\mathbf{x}; R_B) + \tilde{N}(\mathbf{x}), \quad (4.11)$$

where $B(\mathbf{x}; R_B)$ denotes our isolated bubble of radius R_B , and $N(\mathbf{x})$ denotes the thermal noise contribution to the signal. We have subtracted off the overall mean brightness temperature, $\langle \delta T_b \rangle$, since this will not be measured in an interferometric observation. The

4.4 Prospects for Identifying Ionized Regions

tildes indicate that the spatial average has been removed from each of the underlying signal and noise so that $\tilde{B}(\mathbf{x}; R_B)$ and $\tilde{N}(\mathbf{x})$ each have zero mean. In this case $\tilde{B}(\mathbf{x}; R_B)$ has an inverted spherical top-hat profile,

$$\tilde{B}(\mathbf{x}; R_B) = \begin{cases} -\langle x_{\text{HI}}(1 + \delta_\rho) \rangle & |\mathbf{x}| < R_B, \\ 0 & \text{otherwise.} \end{cases} \quad (4.12)$$

The Fourier transform of the isolated bubble is hence related to the Fourier transform of our template by $\tilde{B}(\mathbf{k}; R_B) = -\langle x_{\text{HI}}(1 + \delta_\rho) \rangle T(\mathbf{k}; R_B)$. Note that we express brightness temperatures in units of T_0 (see Equation 4.2), and so all quantities here are dimensionless.

It is straightforward to derive the expected signal-to-noise ratio at the center of the isolated ionized region, and thereby gauge the prospects for bubble detection with a matched filter technique. Let us assume that the radius, R_B , of our template filter is well matched to the true radius of the ionized region. This will maximize the expected signal-to-noise ratio. Neglecting foregrounds for the moment, and using the fact that the thermal noise has zero mean, we find that the signal-to-noise ratio at bubble center for the optimal matched filter is:

$$S(R_B) = -\langle x_{\text{HI}}(1 + \delta_\rho) \rangle \left[\int \frac{d^3k}{(2\pi)^3} \frac{T^2(k; R_B)}{P_N(k, \mu)} \right]^{1/2}. \quad (4.13)$$

For our sign convention, in which the template and ionized regions have opposite signs, this quantity is negative – ionized bubbles are regions of low 21 cm signal. The contribution of a Fourier mode to the signal to noise ratio depends on the relative size of $T^2(k; R_B)$ and $P_N(k, \mu)$: modes for which the template is much larger than the noise power contribute appreciably to $S(R_B)$ while modes dominated by the noise power are not useful. The signal-to-noise ratio depends on the neutral fraction: a larger exterior neutral fraction increases the contrast between an ionized bubble and the exterior, and hence boosts the detectability of the ionized region. We would like to calculate the expected signal-to-noise ratio for ionized regions of different sizes and for various volume-averaged ionization fractions. To do this, we need to connect the volume-averaged ionized fraction with the mass-averaged fraction,

4.4 Prospects for Identifying Ionized Regions

$\langle x_{\text{HI}}(1 + \delta_\rho) \rangle$, which enters into Equation 4.13. Here we should incorporate that large scale overdense regions are generally ionized before typical regions during reionization, i.e., the neutral fraction and overdensity fields are anti-correlated. Defining $\delta_x = (x_{\text{HI}} - \langle x_{\text{HI}} \rangle) / \langle x_{\text{HI}} \rangle$, we approximate $\langle \delta_x \delta_\rho \rangle$ as fixed at $\langle \delta_x \delta_\rho \rangle = -0.25$ throughout the reionization process (?).

The results of the signal to noise calculation are shown in Figure 4.4 for the MWA-500 and a LOFAR-type experiment. Here we consider only our fiducial redshift, $z_{\text{fid}} = 6.9$. The (absolute value of) $\mathcal{S}(R_B)$ is evidently a strongly increasing function of bubble size. This occurs because the thermal noise is a strong function of scale and only the rather large scale modes are measurable. It is encouraging that the expected signal-to-noise ratio exceeds five, $\mathcal{S}(R_B) \gtrsim 5$, for a range of radii and neutral fractions. This corresponds to a $5 - \sigma$ detection: ‘false’ bubbles at this significance from downward fluctuations in the noise are highly unlikely, with a fraction of only $\sim 3 \times 10^{-7}$ of pixels in the filtered noise having such a large (negative) significance on their own. For simplicity we neglect the impact of foreground cleaning in this figure: this will degrade the expected signal-to-noise ratios somewhat, as we will consider subsequently (see §4.4.3, §4.5.3).

In order to estimate the number of bubbles that can be detected from these curves, we need to consider how many bubbles there are of different sizes, i.e., we need to fold in an estimate of the bubble size distribution. In particular, while the contrast of an ionized region increases with the neutral fraction, large ionized bubbles become increasingly scarce for larger values of the neutral fraction. For instance, we can consider the model bubble size distributions in Figure 4 of (?). This figure indicates that bubbles of radius larger than $30 h^{-1}\text{Mpc}$ are exceedingly rare for neutral fractions larger than $\langle x_{\text{HI}} \rangle > 0.5$, with only the tail end of the distribution extending past $25 h^{-1}\text{Mpc}$. However, bubbles this size are relatively common later in reionization. Since Figure 4.4 indicates that only bubbles with $R \gtrsim 30 h^{-1}\text{Mpc}$ exceed $\mathcal{S}(R_B) \gtrsim 5$, this suggests that bubble detection is feasible for the MWA-500 after the Universe is more than $\sim 50\%$ ionized, but that it will be difficult to use this method at earlier stages of the reionization process. Also, note again that the the calculation here

4.4 Prospects for Identifying Ionized Regions

neglects the effects of foreground cleaning. However, we find that incorporating foreground cleaning only has a small effect on bubbles of this size ($\lesssim 30 h^{-1} \text{ Mpc}$, §4.5.3). Bubble detection will also be challenging once the Universe is less than $\sim 10 - 20\%$ neutral, owing mostly to the reduced contrast between the bubbles and typical regions. If the ionized bubbles at a given stage of the EoR are larger than in the model of (?), then the prospects for bubble detection will be enhanced. We refer the reader to (32) for a quantitative exploration of the bubble size distribution across plausible models for the ionizing sources.

Finally, it is interesting to consider a LOFAR-style interferometer, as discussed further in §4.6. This is shown as the red dot-dashed curve in Figure 4.4. The expected $S(R)$ exceeds that of the MWA-500 for small bubble radii, before flattening off at larger radii. This occurs because the LOFAR-style interferometer has more collecting area per baseline, but a larger minimum baseline. This makes it more sensitive to the smaller ionized regions, but less sensitive to larger ones.

While the signal-to-noise curves in this toy case provide a useful guide, we should keep in mind their limitations. First, it considers only the case of a single isolated ionized region. Next, we consider here only the signal to noise at the bubble center, while an ionized region will typically have a strong (negative) signal to noise over much of its volume. This can help significantly with detection. Finally, we consider only the *average* signal-to-noise ratio here. In practice, the signal-to-noise ratio in a filtered map may fluctuate significantly around this average, as we will see.

4.4.3 Application to a Simulated 21 cm Signal

With the estimates of the previous section as a rough guide, we now apply the matched filter to our noisy mock redshifted 21 cm data. In order to illustrate the results of passing our mock data through a matched filter, we start by examining simulated signal-to-noise fields for a single template radius of $R_T = 35 h^{-1} \text{Mpc}$. This template radius corresponds to the typical size of the ionized bubbles we believe we can detect (see Figure 4.4). A representative

4.4 Prospects for Identifying Ionized Regions

slice through the simulation is shown in Figure 4.5. The results look promising, with signal-to-noise ratios comparable to the values anticipated in the idealized calculation of Figure 4.4. Although the Wiener filter provides the best overall map, or data cube, one can still detect individual features at greater significance by applying a matched filter. Comparing with Figure 4.2, it is clear that the Wiener filter is passing more small scale structure than the matched filter shown here. This results in the signal-to-noise ratio being larger (in absolute value) for the matched filter than for the Wiener filter. In particular, we find values of the signal-to-noise ratio that are as low as ~ -10 in the matched filter data cube, a significant improvement over the global minimum of ~ -6 for the Wiener filter. Moreover, we can compare the filtered noisy signal in the bottom-left panel with the filtered pure noise field in the top-right panel. They differ by more than in the case of the Wiener filter. Indeed, the very low signal to noise ratio regions (shown in dark blue/purple in the bottom-left panel) line up fairly well with ionized regions in the top-left panel. This is especially apparent when comparing the filtered noisy signal to the filtered noiseless signal, shown in the bottom-right panel. For the slice shown, almost all of the significant features in the filtered noiseless signal are preserved in the noisy case. Figure 4.6 shows the impact of foreground cleaning: as in the case of the Wiener filter (Figure 4.3), this compresses structures along the line of sight and reduces the overall signal-to-noise ratio in the data cube. The signal-to-noise ratio is still significant enough, however, to robustly identify ionized regions.

As with the Wiener filter, and in what follows subsequently, we show slices of $8 h^{-1}\text{Mpc}$ thickness. This choice is arbitrary, but we expect similar results provided the slice thickness is small compared to the radius of the template filter. It is important to keep in mind, however, the full data cube will consist of many separate slices of this thickness. Also note that the transverse dimension of the MWA-500 is larger than that of our simulation box by a factor of ~ 3 , and so these slices represent only $\sim 1/9$ of the MWA field of view.

These results are promising, but they are for a single filtering scale, and so we can do significantly better by considering a range of template radii, and looking for extrema in

4.4 Prospects for Identifying Ionized Regions

the resulting signal-to-noise fields. In particular, we proceed to apply a sequence of filters with template radii up to $R_T \leq 75 h^{-1} \text{ Mpc}$ – see §4.4.5 for a justification of this maximum – across the simulation volume. We assign the minimum (most negative) signal-to-noise value obtained over the range of template radii to each simulation pixel and use this to construct a new field. The position of local minima in this field are chosen to be the centers of candidate bubbles, and each such bubble is assigned a radius according to the scale of the template filter that minimizes its signal-to-noise. We focus on *minimum* values since ionized regions are expected to appear as regions of low 21 cm signal. All candidate bubbles whose central signal to noise is lower than -5 are considered to be detected ionized regions.

We find it important to apply one additional criterion to robustly identify ionized regions. The criterion is that a low signal to noise region on scale R_B must additionally be *low in signal to noise at all smaller smoothing scales*, $R_T < R_B$. This guards against the possibility that a detected bubble will be centered on neutral material that is nevertheless surrounded by ionized hydrogen. A region like this will have a high (least negative) signal-to-noise when filtered on small scales and then dive down (gaining statistical significance) when filtered on scales containing the surrounding ionized material. We discard such spurious bubbles by requiring that the field is low on all smaller smoothing scales. The only downside to this procedure is that it occasionally discards true ionized regions whose center happens to coincide with a significant upward noise fluctuation. Overall, however, it improves the quality of detected bubbles (§4.4.4). This cut also requires a threshold choice; we reject candidate regions if their signal-to-noise ratio crosses above a threshold S_{\max} at any smoothing scale less than R_T . After trying several thresholds, we found the most effective choice to be $S_{\max} = -1\sigma$. In principle, one might use the full curve of signal-to-noise ratio versus template radius for each candidate bubble to help verify the detection and determine the properties of the bubble. In practice, we found that individual signal-to-noise curves are noisy and difficult to incorporate into our analysis and so we don't consider this possibility further in what follows.

4.4 Prospects for Identifying Ionized Regions

We apply this algorithm to the mock redshifted 21 cm data and identify 220 ionized regions across the simulation volume (which is different than the MWA survey volume, as we will discuss subsequently). A representative example of a detected bubble is shown in Figure 4.7. The circle in the figure identifies the detected bubble size and the location of its center in both the filtered noisy signal (top-left and top-right panels), as well as in the input signal (bottom-left and bottom-right panels). The algorithm has convincingly identified an ionized region. The detected bubble overlaps a small fraction (%10) of neutral material in the input signal. Although this particular ionized bubble is well identified, most of the ionized regions in the signal will escape detection. This is because the significance levels of the detected bubbles are not that high, and an ionized region generally needs to be coincident with a downward fluctuation in the noise to pass our significance threshold. For example, consider the larger ionized region below and to the left of the detected region in the bottom-left panel of Figure 4.7. This region, while larger and therefore more detectable on average than the identified bubble, happens to coincide with a large upward noise fluctuation and hence fails to cross the significance threshold. While we can not identify all of the large ionized regions in the noisy mock data, we can robustly identify some regions; this may still be quite valuable.

It is also clear that the underlying ionized regions are manifestly non-spherical, creating an ambiguity as to what the appropriate ‘radius’ of the region is. Focusing on the bottom right panel in Figure 4.7, we could imagine the size being reasonably described by a radius $\sim \%50$ larger, so as to enclose more of the nearby ionized material. However, our method naturally favors radii causing little overlap with neutral material at these size scales. Therefore, an ionized region like the one shown in Figure 4.7 is more likely to be detected as several small ionized regions than one large one, although both characterizations seem reasonable.

Figure 4.8 gives a further example of how the algorithm identifies bubbles, and some of the ambiguities that can result. This figure shows an example of an irregular, yet contiguous,

4.4 Prospects for Identifying Ionized Regions

ionized region that is detected as more than one ionized bubble. Here we show spatial slices through the center of the middle sphere, marked with a solid circle, which happens to intersect neighboring ionized bubbles, whose cross sections are shown as dashed circles. Hence, our algorithm generally represents large, irregularly shaped, yet contiguous, regions as multiple ionized bubbles.

It is important to emphasize further the difference between the simulated results shown here and the idealized test case of the isolated bubble shown in the previous section. In particular, we consider here the application of matched filters to the 21 cm signal during a late phase of reionization in which many ionized regions, with a broad size distribution, fill the survey volume: the ionized regions *are not* isolated bubbles in a sea of partly neutral material. When applying a matched filter of template radius R_T around a point, the values of the field at many neighboring pixels impact the filtered field at the point in question. It is hence possible that a filtered pixel is affected by several distinct neighboring ionized regions. Indeed, this can result in even neutral regions having low signal-to-noise ratios provided they are surrounded by many nearby ionized regions. For instance, in the low noise limit, *any region with volume-averaged neutral fraction lower than the cosmic mean would pass our significance threshold*. To guard against this type of false detection, we implemented the requirement that a candidate bubble has low signal-to-noise for *all* template radii smaller than the detected radius. Another possibility might be to treat small ionized regions as an additional noise term in the filter. However, in practice, our attempts along these lines introduced an additional level of model dependence without significantly increasing the quality of the detected bubbles. Ultimately, it is important to keep in mind that the signal-to-noise values quoted here reflect only the likelihood that a value arises purely from noise, and so they are not strictly indicative of the quality of the detected bubbles.

4.4.4 Success of Detecting Ionized Regions

We hence turn to describe the characteristics of the detected ionized regions, and to quantify the method’s level of success in detecting ionized bubbles. To do this, we calculate the fractional overlap of each detected bubble with ionized material in the underlying signal. Additionally, we estimate how many ionized bubbles should be detectable across the entire MWA-500 survey volume.

The matched filter technique finds 220 bubbles across our simulation volume. However, the algorithm for determining bubble positions and sizes allows for bubbles to occupy overlapping areas, as shown in Figure 4.8. We find that $\sim 55\%$ of the detected bubbles have *some* overlap with another bubble, although only $\sim 15\%$ of the *total* volume occupied by detected bubbles is occupied by more than one. Regardless, $\sim 96\%$ of the detected ionized bubbles have an average ionized fraction larger than $x_i = 0.79$, which is the volume-averaged ionization fraction of the simulation box at the redshift under consideration. Furthermore, $\sim 42\%$ of the detected bubbles have an ionized fraction greater than $x_i = 0.9$. The lowest ionization fraction of a detected bubble is $x_{i,\min} = 0.77$, just slightly below the volume-averaged ionization fraction of the simulation. In total, we detect 9 bubbles whose ionized fractions are less than the average ionization fraction of the box. Inspection reveals that these regions happen to be coincident with significant ($\leq -3 - \sigma$) downward noise fluctuations.

In Figure 4.9, we plot the volume-averaged ionized fraction within each of our detected bubbles against the detected bubble radius. For comparison, we show the $1 - \sigma$ spread in the ionized fraction enclosed by *randomly* distributed spheres of the same size.¹ The spread in ionization of the randomly distributed spheres around the box average ionized fraction, $\langle x_i \rangle = 0.79$, decreases with increasing radius; this reflects the drop off in the power spectrum of the ionization field towards large scales. Most of the detected bubbles are significantly

¹The $1 - \sigma$ spread shown in the figure extends past $x_i = 1$, but this is only because the distribution of ionized fractions is not symmetric about the mean, i.e., the probability distribution function of the ionized fraction is non-Gaussian.

4.4 Prospects for Identifying Ionized Regions

more ionized than random regions, as expected, indicating a significant success level. There are a few poor detections which result mostly from downward noise fluctuations. There is a small overall decrease in the ionized fraction of detected bubbles larger than $R_B \gtrsim 40 h^{-1}\text{Mpc}$, suggesting that we may no longer be detecting individual ionized regions here. These regions may potentially be distinguished from isolated bubbles by examining the signal-to-noise ratio as a function of template radius closely, as we discuss in §4.4.5.

We can estimate the number of ionized regions detectable in the MWA-500 by scaling from our simulation volume to the MWA survey volume. At $z_{\text{fid}} = 6.9$, for an ionized fraction of $\langle x_i \rangle = 0.79$, we expect to find 140 bubbles in a $B = 6 \text{ MHz}$ chunk of the MWA, over its entire field of view of $\sim 770 \text{ deg}^2$. About 135 (60) of these detected bubbles are expected to have ionized fractions larger than 79% (90%). This estimate comes from simply scaling our simulation volume (which is deeper than the MWA bandwidth) to a 6 MHz portion of the MWA survey volume. Analyzing the MWA data over a 6 MHz chunk is meant to guard against redshift evolution: the full bandwidth of the survey is $B = 32 \text{ MHz}$ and so the prospects for bubble detection across the full survey are even better than this estimate suggests. The precise gain will be dependent on how rapidly the bubble size distribution evolves across the full survey bandwidth. One caveat with our estimate, however, is that $B = 6 \text{ MHz}$ corresponds to only $\sim 70 h^{-1}\text{Mpc}$. This is comparable to the size of our larger bubbles, and so analyzing chunks this small might weaken our ability to detect large bubbles. This effect is not incorporated in our scaling estimate, which simply takes the ratio of the MWA survey volume and our simulation volume. In practice, one can perform the bubble extraction for different analysis bandwidths to help ensure robust detections.

4.4.5 Range of Template Radius Considered

It is worth mentioning one further detail of our algorithm. In the previous section, we set the maximum template radius considered at $R_{\text{T,max}} = 75 h^{-1}\text{Mpc}$, without justification.

4.4 Prospects for Identifying Ionized Regions

In fact, we have a sensible and automated way for arriving at this choice. We discuss this procedure briefly here.

A good candidate ionized region should in fact obey three criteria. First, it should have a large (negative) signal-to-noise ratio, so that it is unlikely to result from a noise fluctuation. Second, the signal-to-noise ratio should be small for all trial radii smaller than the optimal template radius, as discussed in §4.4.3. Finally, the total signal must itself be small in an absolute sense. In the limit of low noise, anything less neutral than average would qualify as a bubble by the first criterion, and so this third criterion may then become important for robustly identifying bubbles. This low noise limit is relevant for the MWA-500 only on very large smoothing scales, where the noise averages down significantly.

Since this third criterion becomes important only on very large smoothing scales here, we use it only to set the maximum template radius considered. Without this third consideration, our algorithm generally identifies a few excessively large ionized bubbles, but this can be easily understood and avoided as follows. Consider, for the moment, the 21 cm brightness temperature field in the absence of noise and foregrounds. Let's further work in units of T_0 (Equation 4.2), and remove the average brightness temperature contrast across the data cube. In this case, the signal inside a highly ionized bubble is expected to be $-\langle x_{\text{HI}}(1 + \delta_\rho) \rangle$. If we now spherically average the field on scales smaller than the bubble, the value at bubble center will not change from this value, $-\langle x_{\text{HI}}(1 + \delta_\rho) \rangle$. Once the smoothing scale becomes larger than the bubble scale, however, surrounding neutral material will increase the value of the filtered field at bubble center. Hence, if the filtered field becomes everywhere larger than $-\langle x_{\text{HI}}(1 + \delta_\rho) \rangle$ on some smoothing scale, it is clear that no larger ionized bubbles exist within the data cube. This suggests that we can set the maximum template radius by requiring that the filtered noisy signal reaches sufficiently small values, at some locations across the data cube, for there to still plausibly be completely ionized regions. Since the presence of noise only increases the variance, this should provide a conservative estimate of the maximum size of the ionized regions. In practice, we need to chose

a threshold criterion without assuming prior knowledge of the neutral fraction. Here we set the maximum template radius to be the smallest smoothing scale at which the filtered noisy field everywhere exceeds $-\langle x_{\text{HI}}(1 + \delta_\rho) \rangle \geq -0.075$. This corresponds to the expected contrast at $\langle x_{\text{HI}} \rangle = 0.1$, assuming $\langle \delta_x \delta_\rho \rangle = -0.25$, and yields a maximum template radius of $R_{\text{T,max}} = 70 h^{-1} \text{ Mpc}$, $73 h^{-1} \text{ Mpc}$, $75 h^{-1} \text{ Mpc}$, and $75 h^{-1} \text{ Mpc}$ for $\langle x_i \rangle = 0.51, 0.68, 0.79$, and 0.89 , respectively. The precise threshold used here, -0.075 , is somewhat arbitrary but this choice is only being used to set the maximum template radius considered.¹

4.5 Variations on the Fiducial Model

So far, we have considered the prospects for bubble detection only in our fiducial model with $\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$ and only for the MWA-500. Here we consider first alternate models in which the Universe is more or less ionized at $z_{\text{fid}} = 6.9$ than in our fiducial case, and then consider how the sensitivity declines towards higher redshifts at fixed ionized fraction. In addition, we consider variations around our fiducial assumptions regarding the impact of foreground cleaning. Then we turn to consider the sensitivity of the MWA-128; this is meant to illustrate the prospects for bubble detection with the very first generation of redshifted 21 cm surveys, while the MWA-500 represents a second generation survey.

4.5.1 Ionized Fraction

In order to consider bubble finding at earlier and later stages of the EoR, we run our matched filter on simulation outputs with volume-averaged ionized fractions of $\langle x_i \rangle = 0.51, 0.68$, and 0.89 . As discussed in §4.3.1 these outputs are actually at slightly different redshifts, but we

¹This choice might appear to preclude the possibility of detecting bubbles at the end of reionization when $\langle x_{\text{HI}}(1 + \delta) \rangle \leq 0.075$. However, the threshold choice is only used to set the maximum template radius, and so ionized regions may still in principle be detected at these late stages of reionization. The ionized regions identified at the end of reionization are, however, generally less robust given the reduced contrast between fully ionized and average regions at these times (see §4.5.1).

4.5 Variations on the Fiducial Model

generate the 21 cm field as though they were at $z_{\text{fid}} = 6.9$. As far as bubble detection is concerned, varying the ionized fraction leads to two, generally competing, effects. First, the bubbles grow as reionization proceeds. This tends to boost detection, since it is only the large scale modes that are detectable over the thermal noise. Second, however, the *contrast* between an ionized region and a typical volume of the Universe is reduced as reionization proceeds. This makes bubble detection more difficult. Both of these effects are quantified in the idealized isolated bubble case in Figure 4.4. It is also clear that the ideal ionized fraction for bubble detection will be somewhat survey dependent. As already illustrated in Figure 4.4 and discussed further in §4.6, a LOFAR-type interferometer will perform better when the ionized regions are smaller.

We find that the matched filter is capable of detecting ionized regions for each of the ionized fractions studied. In Figure 4.10 we show histograms of the detected bubble size distributions for each ionized fraction. Since we preferentially detect large ionized regions, we don't expect these distributions to be representative of the true underlying bubble size distributions. For example, in Figure 4.10, the size distribution peaks around $\gtrsim 40 h^{-1} \text{ Mpc}$ for the case of $x_i = 0.79$, despite volume-weighted size distribution peaking around $\sim 30 h^{-1} \text{ Mpc}$ in Figure 4 of (?) at roughly the same ionized fraction. Nonetheless, the histograms illustrate a general shift from smaller to larger detected bubble radius as the ionized fraction increases. By applying the matched filter to several redshift bins, one can potentially observe precisely this evolution with the MWA-500. This would complement studies of the 21 cm power spectrum evolution over the same redshift range (e.g., ?). From the histograms, one can see that – of the models shown – the best ionized fraction for bubble detection is $\langle x_i \rangle = 0.79$. This is apparently near the sweet spot for the MWA-500 where the bubbles are large enough in the model for detection, but the contrast with typical regions is still sufficiently large.

The average ionized fraction within detected bubbles varies significantly across the different ionized fractions considered. Specifically, the percentage of detected bubbles that

are more than 90% ionized is 0%, 15%, 43%, and 91% for $\langle x_i \rangle = 0.51, 0.68, 0.79$, and 0.89, respectively. However, in each case the percentage of detected bubbles with ionized fraction larger than the (global) volume-averaged ionization fraction is fixed at $\gtrsim 95\%$. At first glance, one aspect of these results may appear to be in tension with the calculations of §4.4.2, where we estimated that bubble detection would be unsuccessful for neutral fractions larger than $\langle x_{\text{HI}} \rangle \gtrsim 0.5$. However, this estimate considered the detection of *isolated* bubbles. Inspection reveals that the detected bubbles at $\langle x_i \rangle = 0.51$ each correspond to clusters of smaller ionized regions. Evidently, these appear as a single larger ionized region after convolving with the template filter and downweighting the noisy short-wavelength modes. In practice, it may be possible to distinguish this case from that of an isolated bubble by analyzing the trend of signal-to-noise ratio versus trial template radius. The signal-to-noise ratio is expected to grow more rapidly with radius (before reaching the bubble scale) for a truly isolated bubble.

4.5.2 Timing of Reionization

We now consider how the prospects for bubble detection diminish if reionization occurs earlier and the observations are focused on the corresponding redshifts. In particular, we examine the case that our model with an ionized fraction of $\langle x_i \rangle = 0.79$ is observed at a higher redshift. We focus on this case since this ionized fraction appears close to optimal for bubble detection. Aiming for only a rough estimate here, we consider the prospects for detecting a $R_B = 40 h^{-1}\text{Mpc}$ bubble.

Although several different factors in the noise power spectrum of Equation 4.5 scale with redshift, the dominant scaling is with the sky temperature. The noise power scales as $P_N \propto T_{\text{sky}}^2$, and the sky temperature follows $T_{\text{sky}} \propto \nu^{-2.6} \propto (1+z)^{2.6}$. Therefore, we expect the signal-to-noise ratio of a detected bubble to fall off with increasing observation redshift roughly as

$$\mathcal{S}(z) = \mathcal{S}(z_{\text{fid}}) \left(\frac{1+z_{\text{fid}}}{1+z} \right)^{2.6}. \quad (4.14)$$

4.5 Variations on the Fiducial Model

This indicates the signal-to-noise ratio for a bubble detected with a signal-to-noise of $\mathcal{S}(z_{\text{fid}})$ at redshift $z_{\text{fid}} = 6.9$, if the bubble were instead observed at redshift z . A relatively large bubble with $R_B \approx 40 h^{-1}\text{Mpc}$ has a typical signal-to-noise ratio at bubble center of $\mathcal{S} \approx 4$ at our fiducial redshift. This value is found by incorporating foreground cleaning into the corresponding curve in Figure 4.4. According to Equation 4.14, the signal-to-noise value will be reduced to a significance of $\mathcal{S} \approx 2$ (1) at $z = 9.3$ (12.5). The bubble will, in fact, be more detectable than implied by this one number – the signal-to-noise ratio at bubble center – since an ionized region should have low signal-to-noise over much of its volume. From this, we conclude that bubble detection should be feasible with the MWA-500 if our fiducial ionized fraction occurs later than $z \lesssim 9$ or so, but that the prospects are rather limited in the case of significantly earlier reionization. A range of recent work in the literature, however, suggests that reionization is unlikely to complete so early. See, for example, Figure 9 from the recent study of (?) which combines Ly- α forest data (Fan et al. 17), measurements of the Thomson optical depth from WMAP (?), and measurements of the Lyman-break galaxy luminosity function (?). Hence, the prospects for bubble detection appear good for the MWA-500.

4.5.3 Effects of Foreground Cleaning

Next, we consider the impact of variations around our standard foreground cleaning model. As discussed previously (§4.2.4), our standard assumption is that the impact of foreground cleaning can be approximately mimicked by removing the running mean, over a bandwidth of $B = 16 \text{ MHz}$, across each line of sight. The optimal foreground cleaning strategy avoids ‘over-fitting’ by removing the smoothest possible function over the largest possible bandwidth, in order to preserve the underlying signal as much as possible. It also avoids ‘under-fitting’ by ensuring that foreground residuals do not excessively contaminate the signal. (?), for example, find that 21 cm foregrounds can be removed to one part in 10^5 or 10^6 by subtracting roughly three modes over $\sim 32 \text{ MHz}$ of bandwidth. This should have a fairly

similar impact to our fiducial cleaning model, but we would expect a bit more degradation in this case. A detailed investigation would add foreground contamination into our mock data cubes, and explore the impact of various cleaning algorithms directly.

Here, we instead check how our results change for slightly more and less aggressive foreground cleaning. In particular, we remove the running mean over each of $B = 8$ MHz and $B = 32$ MHz and rerun our bubble finding algorithm (at $z_{\text{fid}} = 6.9$, $\langle x_i \rangle = 0.79$). This has little impact on our results. In particular, the number of identified bubbles varies by less than 10 – 15% across the range of cleaning bandwidths considered. The quality of detected bubbles decreases slightly for the more aggressive cleaning model, and improves slightly in the most optimistic case. Specifically, for $B = 16$ MHz, 96% (43%) of bubbles have ionized fractions exceeding $x_i = 0.79$ (0.9); for $B = 8$ MHz the corresponding numbers are 90% (32%); and for $B = 32$ MHz the same numbers are 96% (62%). While these estimates are encouraging, a more detailed study is warranted. It may also be advantageous to estimate the power spectrum of the foregrounds, and incorporate this as an additional noise term in Equation 4.7 and Equation 4.9 for each of the Wiener filter and the matched filter, respectively.

4.5.4 128 Antenna Tile Configurations

So far our analysis has focused on the MWA-500, which is meant to represent a second generation 21 cm survey. In the near term, it is timely to consider the prospects for a 128 tile version of the MWA (the MWA-128) which is ramping up to take data in the very near future. This should be significantly less sensitive, since the number of baselines scales as the number of antenna tiles squared.

In order to generate thermal noise representative of the MWA-128, we start by considering a similar antenna distribution as for the MWA-500. In particular, we assume all of the antenna tiles are packed as closely as possible within a core of radius 8m and that the antenna distribution subsequently falls off as r^{-2} out to a maximum baseline of 1.5 km.

4.5 Variations on the Fiducial Model

After comparing the thermal noise power spectrum in this configuration with that in (?), we find that our noise power is larger by up to a factor of a few. This could possibly be due to our approximation of a smooth antenna distribution being less valid for the MWA-128, or to the fact that our analytic formula for the noise power spectrum does not incorporate a full treatment of rotation synthesis. In an effort to bracket somewhat the impact of the detailed antenna distribution, we further consider a configuration where all antenna tiles are packed as closely in a dense core of radius $\sim 25\text{m}$. This resembles the ‘super-core’ configuration considered in (?) for the power spectrum.

The results of applying the optimal matched filter for a single template radius of $35h^{-1}\text{Mpc}$ are shown in Figure 4.11 for the r^{-2} tile distribution. This shows that the sensitivity is much lower than for the MWA-500, as expected. It is much more difficult to distinguish the filtered noisy signal (bottom-left panel) from the filtered pure noise (top-left) panel here than in Figure 4.5. Most of the significant, dark blue regions in the filtered noisy signal correspond simply to low noise regions. However, applying the detection algorithm we do nonetheless detect 7 bubbles across a volume equivalent to a 6 MHz chunk of the MWA-128 survey. The success is generally lower than in the case of the MWA-500: here 75% of detected bubbles exceed the average ionized fraction of the box, while $\sim 42\%$ exceed $x_i = 0.9$. In the supercore configuration, we find slightly higher significance levels (up to $6.9 - \sigma$) but the identified regions generally correspond to several large clustered ionized regions, rather than a single ionized bubble. Altogether, the algorithm identifies 10 ionized regions across the MWA survey in the supercore configuration, but the identified regions have a lower overall quality than in the r^{-2} configuration.

Our conclusion is that bubble detection is only marginally possible with the MWA-128. While the results are unlikely to be very compelling, it is worth applying the matched filter to the first generation surveys as an initial test. Even a few weakly identified candidate bubbles would provide compelling targets for follow-up observations. Another possibility is to focus on *targeted searches* around known bright sources for the MWA-128 (e.g., ?, ?).

4.6 Favorable Antenna Configurations for Bubble Detection

The possibility of imaging or identifying ionized regions from second generation redshifted 21 cm surveys invites the question: how do we optimize future surveys for this goal? It seems unlikely that the optimal configuration for bubble detection is identical to that for measuring the power spectrum, although power spectrum measurements have mostly driven survey design considerations thus far. In the case of the power spectrum, one aims to minimize the error bar on power spectrum estimates in particular bins in wavenumber. The power spectrum error bar for each \mathbf{k} -mode contains a thermal noise term and a sample variance (sometimes called ‘cosmic variance’) term. Because of the sample variance contribution, the gain from reducing the thermal noise for a given \mathbf{k} -mode is limited: once the thermal noise is reduced sufficiently far below the sample variance, it is advantageous to instead measure a different \mathbf{k} -mode on the sky within the $|\mathbf{k}|$ bin of interest. As a result, grouping individual antennas into only small tiles to achieve a wide survey, generally reduces the statistical error bars on power spectrum measurements compared to antenna configurations with larger tiles that probe narrower fields of view. For imaging and bubble detection, one aims for the best possible signal to noise on *particular regions of the sky*. In other words, for good imaging one wants to reduce the thermal noise to *well* below the sample variance level. Grouping individual antennas into larger tiles, in order to devote more collecting area to a narrower field of view, may be better for this purpose.

In order to get some sense for these trade-offs, we consider here a LOFAR-style interferometer with the specifications listed in (32). Although the detailed specifications for LOFAR have evolved somewhat (e.g., ?), (as have the MWA specifications), this is nonetheless a helpful case to consider. In particular, our toy LOFAR-style interferometer has many fewer antenna tiles than the MWA-500 but a substantially larger collecting area per tile. Specifically, the interferometer considered has $N_a = 32$ antenna tiles, $A_e = 596\text{m}^2$ at our fiducial redshift (compared to $A_e = 11.25\text{m}^2$ for the MWA-500), $d_{\min} = 100\text{m}$, and $d_{\max} = 2 \text{ km}$. We assume that antenna tiles are packed as closely as possible, consistent

4.6 Favorable Antenna Configurations for Bubble Detection

with $d_{\min} = 100\text{m}$, inside a compact core and that the density subsequently falls off as r^{-2} , out to a maximum radius of $r_{\max} = 1000\text{m}$. These parameters are meant to broadly represent an upgraded version of the existing LOFAR array, analogous to our MWA-500 survey, which is an upgrade to the ongoing MWA-128 instrument. With these parameters, the LOFAR-style interferometer has more total collecting area than the MWA-500 setup by a factor of a few.

The results of applying a matched filter to a data cube with simulated LOFAR noise are shown in Figure 4.12. Here we zoom in to show a portion of our simulation box that matches the smaller field of view of this LOFAR-like instrument. From the figure it is evident that the filter removes large scale structures, a result of the relatively large minimum baseline of this interferometer. In addition, the maximum signal-to-noise achieved here is smaller than with the MWA-500 (it drops from 10 to 6.9). Nonetheless, many small-scale ionized regions in the unfiltered noise-less signal are well preserved in the filtered noisy signal. This is consistent with the idealized calculation of Figure 4.4, which showed that LOFAR should have a higher signal-to-noise detection of small ionized regions, but a reduced signal-to-noise otherwise. Because of this, the LOFAR-style configuration is more successful during earlier stages of reionization when the bubbles are still relatively small. In general, we find that the LOFAR-style configuration detects slightly fewer bubbles overall but with more success for $\langle x_i \rangle \lesssim 0.79$, while the MWA-500 has a greater level of success at later stages of the EoR.

This example suggests that the ideal configuration for bubble detection is likely intermediate between the MWA-style and LOFAR-style antenna configurations. It appears helpful to have more collecting area on fewer baselines than the MWA, but a smaller minimum baseline than in the LOFAR-style instrument is necessary to detect large bubbles. This deserves further study, however: for example, we have neglected calibration requirements and systematic concerns. These considerations will also certainly drive the experimental design. As a further concrete example of how systematic concerns could impact the design of future arrays, suppose foreground cleaning requires removing more large scale modes

4.7 Comparisons to Previous Work

than anticipated. In this case, it would make sense to focus efforts on smaller bubbles. This would shift the ideal configuration closer to a LOFAR-style instrument with a larger minimum baseline.

4.7 Comparisons to Previous Work

Previous work by (?) and (?) also considered the possibility of detecting ionized regions in noisy redshifted 21 cm data sets using a matched filter technique. The main difference between our study and this earlier work is that these previous authors considered the prospects for detecting a *specific* spherical ionized region of varying size, i.e., they considered the detectability of a spherical bubble at the origin, or offset slightly from the origin. These authors also considered the case where the bubble of interest was embedded in a variety of different ionization environments; the bubble under consideration was not always isolated. Altogether, their study is mostly similar to a targeted search, where one has a good prior regarding the likely location of an ionized region. It also provides a feasibility estimate for a more ambitious blind search. The main advantage of a targeted search is that, if a region is known *a priori* to be highly ionized, one need not worry about an entirely false detection from a downward noise fluctuation. This then allows a lower significance threshold for robust bubble detection, and may therefore be the most feasible approach for the MWA-128 and other first generation surveys.

Nonetheless, our work is a significant extension to the earlier work by (?) in that we conduct a blind search across an entire mock survey volume. A detailed comparison with their work is not straightforward given the difference between our approaches, but both studies have a similar bottom-line conclusion: ionized regions are detectable with surveys similar to the MWA-500.

4.8 Conclusion

We considered the prospects for making low-resolution images of the 21 cm sky and for direct, blind detection of ionized regions using first and second generation 21 cm surveys. We find that a 500-tile version of the MWA, the MWA-500, is potentially capable of detecting ionized regions. In our fiducial model, in which 79% of the volume of the Universe is ionized at $z_{\text{fid}} = 6.9$, the MWA-500 can find ~ 150 ionized regions in a $B = 6$ MHz chunk after $\sim 1,000$ hours of observing time. First generation surveys, such as the MWA-128, are substantially less sensitive. We find that the MWA-128 may, nonetheless, be able to detect a handful of ionized regions across its survey volume, with 7 expected in our fiducial model. The MWA-128 may be more effective at identifying ionized bubbles using targeted searches towards, for example, bright quasars (e.g. ?).

There are several possible future directions for this work. First, while we incorporate realistic levels of thermal noise and mimic the effect of foreground cleaning, it will be important to test the robustness of bubble detection with a more detailed model for foreground contamination, and to consider systematic effects from calibration errors and the MWA instrumental response. These considerations can also help in determining the optimal design for future surveys aimed at bubble detection. Our first efforts considering which configurations of antenna tiles are favorable for bubble detection, detailed in §4.6, suggest that an observing strategy intermediate to that of the MWA and LOFAR is favorable. It would also be interesting to consider the prospects for bubble identification across a larger range of reionization models than considered here. If the ionized regions at a given stage of reionization are, in fact, larger than in the models considered here, this should increase their detectability. On the other hand, if the ionized regions are smaller than in our present models, this would likely diminish detectability, at least for the MWA-500.

If blind bubble identification is indeed feasible in future 21 cm surveys, we believe this will open up several interesting avenues of investigation. First, direct identification of ionized regions can help to build confidence in early redshifted 21 cm detections. Next, if the

4.8 Conclusion

centers of ionized regions can be robustly identified, one may be able to use the brightness temperature contrast between the signal near the bubble's center and its surroundings to directly constrain the cosmic mean neutral fraction (e.g., ?). These authors also discuss how detected bubbles can be used to calibrate foreground cleaning (?). Finally, identifying ionized regions in redshifted 21 cm surveys allows one to commence follow-up observations, comparing galaxy properties inside detected bubbles with those in more typical regions. Typical regions and likely neutral regions can be identified as locations in the data cube with average and maximal signal-to-noise ratios, respectively, after applying the matched filter. Furthermore, if the edge of an ionized region can be identified precisely enough, one might imagine targeted searches for galaxies at the edge of bubbles, close to neighboring neutral gas. Spectroscopic observations of these galaxies might then help to reveal the damping wing redward of the Ly- α line (e.g., Miralda-Escude 36). This would provide yet another means for constraining the neutral fraction.

4.8 Conclusion

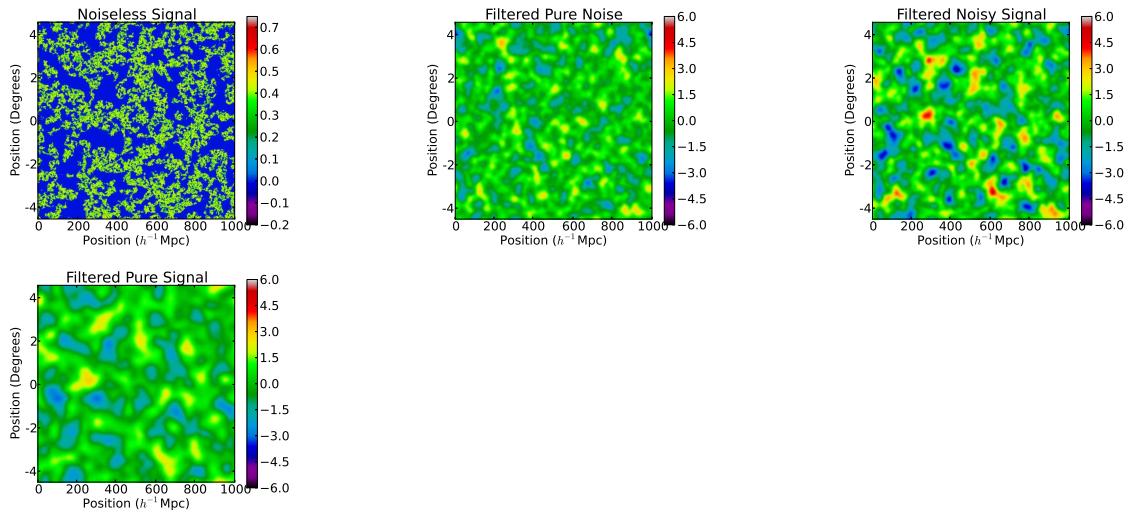


Figure 4.2: Application of the Wiener filter to simulated data. The results are for our fiducial model with $\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$. *Top-Left:* Spatial slice of the unfiltered and noise-less 21 cm brightness temperature contrast field (normalized by T_0). *Top-Right:* Simulated signal-to-noise field after applying the Wiener filter to a pure noise field. *Bottom-Left:* Simulated signal-to-noise field after applying the Wiener filter to the noisy signal. This can be compared with the uncorrupted input signal shown in the top-left panel and the noise realization in the top-right panel. *Bottom-Right:* Simulated signal-to-noise field after applying the Wiener filter to the noiseless signal. (The filtered noiseless signal shown here is normalized by the standard deviation of the noise to facilitate comparison with the other panels.) All panels show a square section of the MWA field of view transverse to the line of sight with sidelength $L = 1 h^{-1} \text{ Gpc}$. All slice thicknesses are $\sim 8 h^{-1} \text{ Mpc}$. Unless noted otherwise, the simulation slices in subsequent figures have these same dimensions.

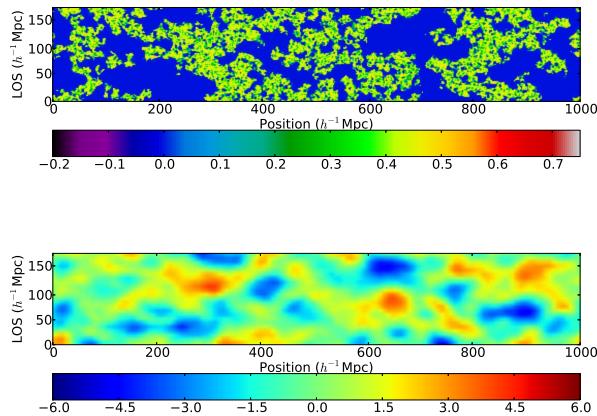


Figure 4.3: Impact of foreground cleaning on the Wiener-filtered field. The top slice is a perpendicular, zoomed-in view of the simulated, unfiltered, noise-less brightness temperature contrast. The bottom slice is the signal-to-noise of the same region after applying the Wiener filter to the noisy signal field. The vertical axis shows the line-of-sight direction, with its extent set to the distance scale for foreground removal, $L_{\text{fg}} = 185 h^{-1} \text{ Mpc}$. The horizontal axis shows a dimension transverse to the line of sight and extends $1 h^{-1} \text{ Gpc}$.

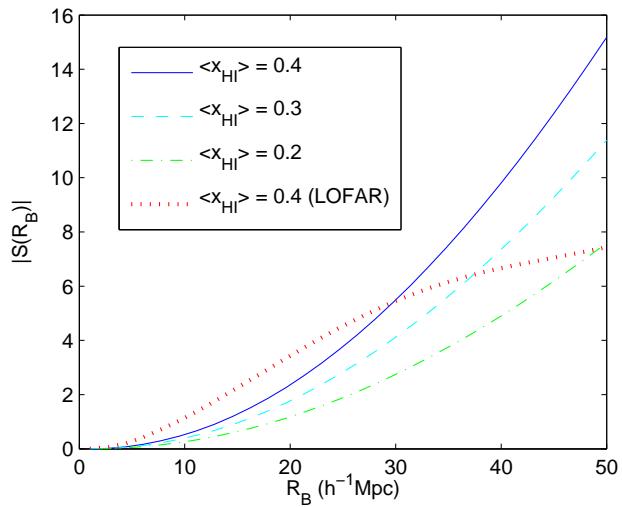


Figure 4.4: Expected signal-to-noise ratio at the center of isolated, spherical, ionized bubbles as a function of bubble radius after applying the optimal matched filter. The curves show the signal-to-noise ratio at $z_{\text{fid}} = 6.9$ for the MWA-500 at various neutral fractions: $\langle x_{\text{HI}} \rangle = 0.4$ (blue solid), 0.3 (cyan dashed), and 0.2 (green dot-dashed). For contrast, the red dotted curve indicates the expected signal-to-noise for an interferometer with a field of view and collecting area similar to a 32-tile LOFAR-like antenna array (at $\langle x_{\text{HI}} \rangle = 0.4$).

4.8 Conclusion

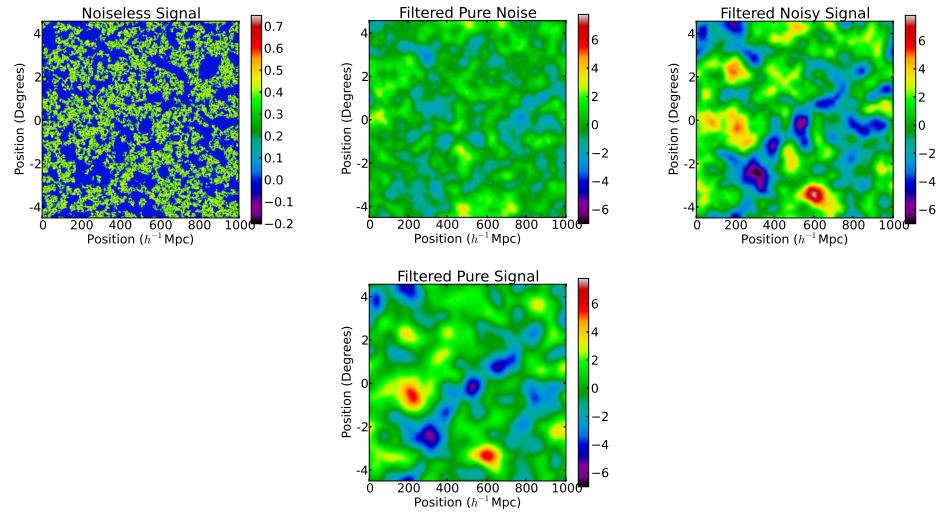


Figure 4.5: Application of the matched filter to simulated data and noise ($\langle x_i \rangle = 0.79$ at $z_{\text{fid}} = 6.9$). The template radius of the filter is $35 h^{-1} \text{Mpc}$, since this is a commonly detected bubble radius for our matched filter search. *Top-Left*: Spatial slice of the unfiltered and noiseless 21 cm brightness temperature contrast field. *Top-Right*: Simulated signal-to-noise field after applying the matched filter to a pure noise field. *Bottom-Left*: Simulated signal-to-noise field after applying the matched filter to the noisy signal. This can be compared directly to the top-left panel. *Bottom-Right*: Simulated signal-to-noise field after applying the matched filter to the noiseless signal. All panels are at the same spatial slice. See text for discussion on interpreting signal-to-noise values.

4.8 Conclusion

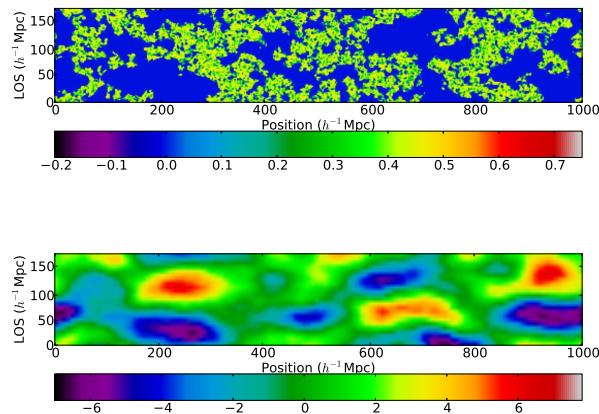


Figure 4.6: Impact of foreground cleaning on the matched-filtered field. This is similar to Figure 4.3, except that the results here are for a matched filter with a template radius of $R_T = 35 h^{-1} \text{ Mpc}$.

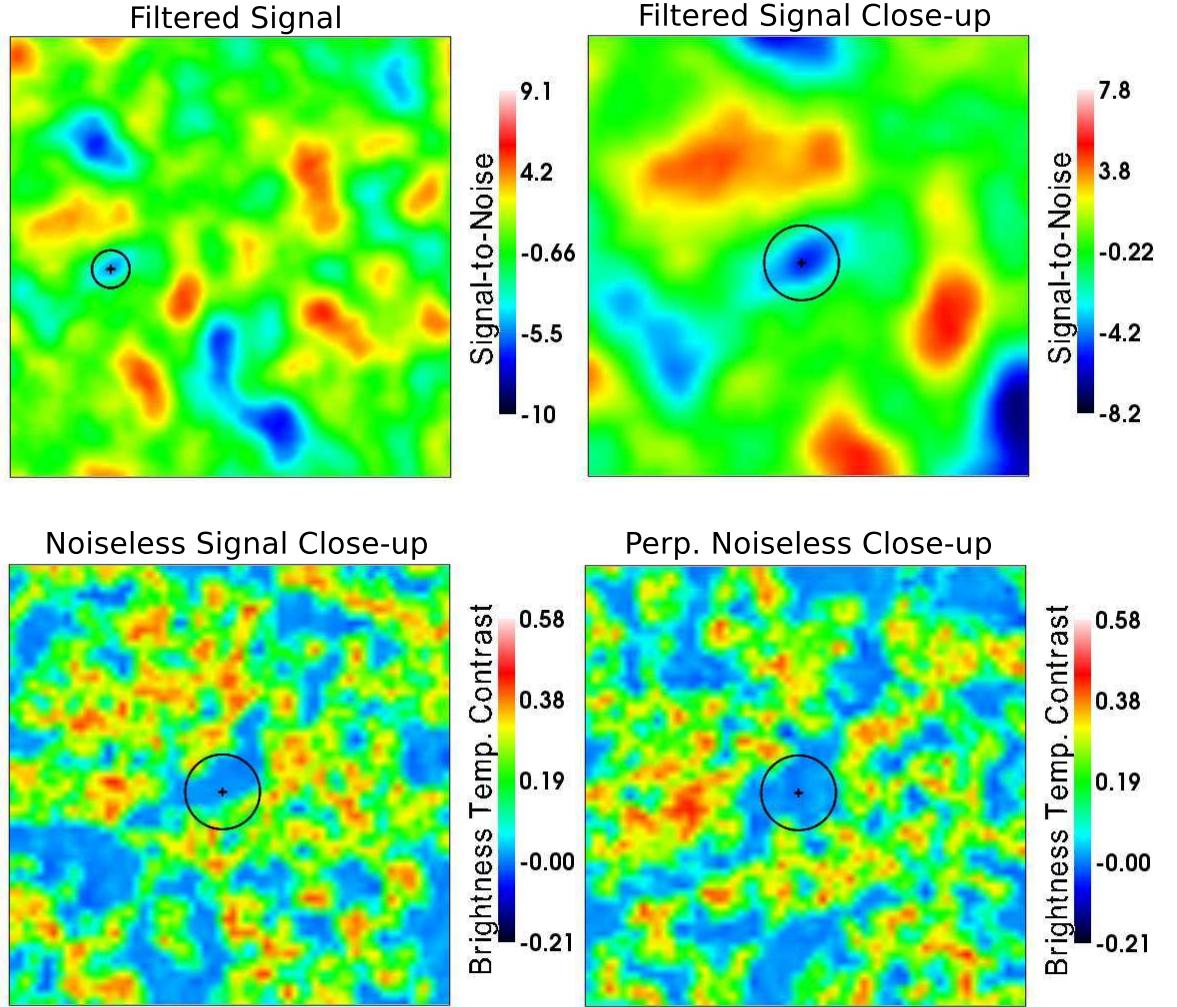


Figure 4.7: An example of a detected ionized region. *Top-left:* Signal-to-noise field after applying the matched filter to the noisy signal. The detected bubble is plotted on top of the corresponding region in the map. *Top-Right:* Zoomed-in view of the detected bubble in the matched-filtered map. *Bottom-Left:* Detected bubble superimposed on a zoomed-in view of the noise-less unfiltered 21 cm brightness temperature contrast map. *Bottom-Right:* A perpendicular zoomed-in view of the bubble depicted in the bottom-left panel. All matched-filtered maps use the template radius that minimizes the signal-to-noise at the center of the detected bubble. In the top-left case, the boxlength is $L = 1 h^{-1} \text{ Gpc}$, while in the zoomed-in slices it is $L \approx 500 h^{-1} \text{ Mpc}$.

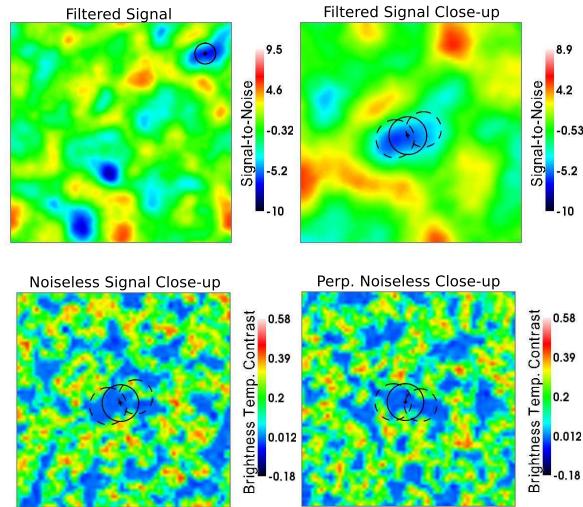


Figure 4.8: An example of an ionized region that our algorithm detects as several neighboring bubbles. *Top-left:* Signal-to-noise field after applying the matched filter to the noisy signal. The main detected bubble is plotted on top of the corresponding region in the map. *Top-Right:* Zoomed-in view of the main detected bubble in the matched filtered map (solid curve) along with two other nearby detected bubbles (dashed curve). *Bottom-Left:* The detected bubble superimposed on the zoomed-in, noise-less, unfiltered 21 cm brightness temperature contrast map. Again, the additional nearby detected bubbles are shown (dashed curve). *Bottom-Right:* A perpendicular view of the bubble depicted in the bottom-left panel, with the nearby detected bubbles visible. All matched-filtered maps use the template radius that maximizes the signal to noise at the center of the main detected bubble. The box length in the top-left figure is $L = 1 h^{-1} \text{ Gpc}$, while in the zoomed-in panels, the box length is $L = 550 h^{-1} \text{ Mpc}$.

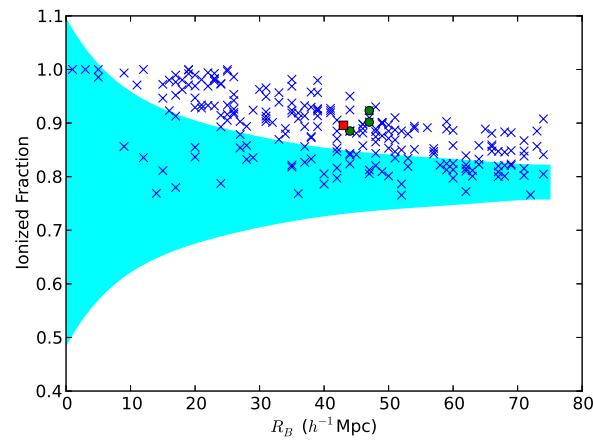


Figure 4.9: A measure of the bubble detection success rate. The points (\times) show the volume-averaged ionized fraction of detected bubbles versus their detected radius. For comparison, the cyan shaded region shows the $1-\sigma$ spread in the ionized fraction of *randomly* placed bubbles of the same radii. The bubble depicted in Fig. 4.7 is marked with a large red square, while the three bubbles shown in Fig. 4.8 are marked with large green circles.

4.8 Conclusion

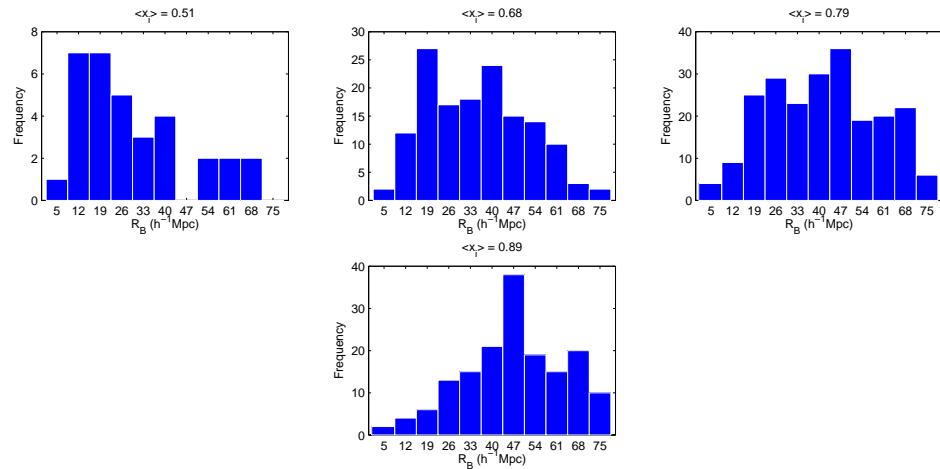


Figure 4.10: Size distributions of detected bubbles for varying (volume-averaged) ionization fractions. The histograms show the size distribution of (identified) ionized regions for simulation snapshots with volume-averaged ionized fractions of $\langle x_i \rangle = 0.51$ (top-left), 0.68 (top-right), 0.79 (bottom-left), and 0.89 (bottom-right). These figures demonstrate how the total number and size distribution of detected bubbles varies with ionized fraction.

4.8 Conclusion

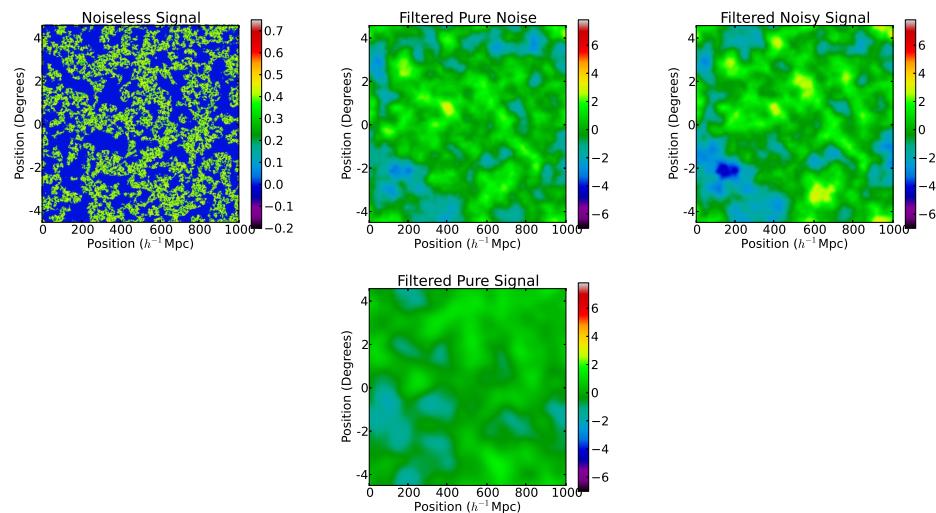


Figure 4.11: Bubble detection with the MWA-128. This figure is similar to Figure 4.5, except it is for the MWA-128 configuration rather than for the MWA-500.

4.8 Conclusion

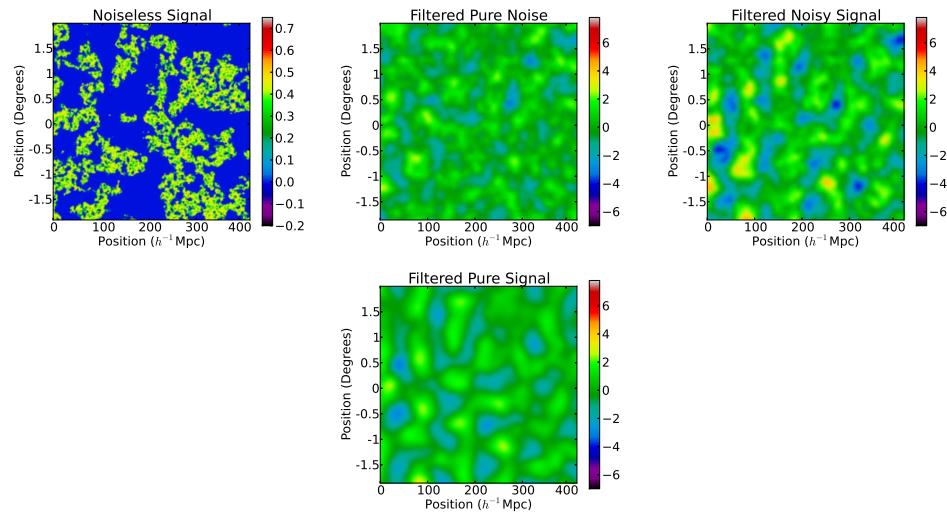


Figure 4.12: Bubble detection with a LOFAR-style interferometer. This figure is similar to Figure 4.5, except it is for the LOFAR configuration rather than the MWA-500. Additionally, all boxes in this figure have a side length of $426 h^{-1}$ Mpc, corresponding to the field-of-view of the LOFAR-style interferometer at $z = 6.9$.

Chapter 5

Conclusion

5.1 Wrapping up...

I rest my case. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus

5.1 Wrapping up...

mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy

5.1 Wrapping up...

vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Appendices

Appendix A

Some Appendix

We will inevitably have appendices...

Glossary

Roman Symbols

M Mass of object, page 190

Greek Symbols

τ Optical depth, page 190

Superscripts

$*$ Conjugate, page 190

Subscripts

\odot relating to the sun (Sol), page 190

Other Symbols

11HUGS 11 Mpc Halpha and Ultraviolet Galaxy Survey, page 190

Acronyms

Ly α Forest This describes the pattern of absorption lines seen blueward of the rest-frame Ly α line, typically in quasar spectra. These absorption lines can be due to significantly neutral gas in the diffuse IGM or due to dense ionized gas., page 2

GRB Gamma-Ray Burst. They are also really bright, OK?, page 2

Hjerting Function Also known as the *Voigt Function*, this function is commonly used in describing line profiles which incorporate Doppler broadening and the natural line width. This function is very relevant when studying the hydrogen damping wing., page 18

Lorentzian Distribution Probability distribution for the ratio of two standard-normal-distributed variables. This distribution also describes the intrinsic line profile for absorption lines., page 13

Optical Depth Optical depth, denoted by τ , is a quantity that describes that likelihood for a photon to be absorbed, usually by a gas. The fraction of photons that will pass through the gas unabsorbed is $e^{-\tau}$., page 5

Quasar It's bright, OK?, page 2

Redshift A quantity commonly used to refer to cosmic periods of time or distances. The redshift of an object or location in space is defined as the fractional increase in wavelength that a photon undergoes due to the expansion of the Universe while travelling from the object or location to us., page 2

Voigt Function Also known as the *Hjerting Function*, this function is commonly used in describing line profiles which incorporate Doppler broadening and the natural line width. This function is very relevant when studying the hydrogen damping wing., page 18

2MASS Two-Micron All Sky Survey, page 190

References

- [1] P.A.R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. 2013. 29, 90
- [2] G. D. Becker, P. C. Hewett, G. Worseck, and J. X. Prochaska. A refined measurement of the mean transmitted flux in the Ly α forest over $2 < z < 5$ using composite quasar spectra. *MNRAS*, 430:2067–2081, April 2013. doi: 10.1093/mnras/stt031. 35
- [3] George D. Becker and James S. Bolton. New Measurements of the Ionizing Ultraviolet Background over $2 < z < 5$ and Implications for Hydrogen Reionization. 2013. 30
- [4] George D. Becker, James S. Bolton, Piero Madau, Max Pettini, Emma V. Ryan-Weber, et al. Evidence of patchy hydrogen reionization from an extreme Ly α trough below redshift six. 2014. 27
- [5] Robert H. Becker et al. Evidence for reionization at $Z = 6$: Detection of a Gunn-Peterson trough in a $Z = 6.28$ quasar. *Astron.J.*, 122:2850, 2001. doi: 10.1086/324231. 25, 35, 49, 83
- [6] J. S. Bolton and M. G. Haehnelt. The observed ionization rate of the intergalactic medium and the ionizing emissivity at $z > 5$: evidence for a photon-starved and extended epoch of reionization. *MNRAS*, 382:325–341, November 2007. doi: 10.1111/j.1365-2966.2007.12372.x. 30
- [7] J. S. Bolton and M. G. Haehnelt. On the rapid demise of Ly α emitters at redshift $z \gtrsim 7$ due to the increasing incidence of optically thick absorption systems. *MNRAS*, 429: 1695–1704, February 2013. doi: 10.1093/mnras/sts455. 27
- [8] J. S. Bolton, G. D. Becker, J. S. B. Wyithe, M. G. Haehnelt, and W. L. W. Sargent. A first direct measurement of the intergalactic medium temperature around a quasar at $z = 6$. *MNRAS*, 406:612–625, July 2010. doi: 10.1111/j.1365-2966.2010.16701.x. xii, 18, 24
- [9] James S. Bolton, Martin G. Haehnelt, Stephen J. Warren, Paul C. Hewett, Daniel J. Mortlock, et al. How neutral is the intergalactic medium surrounding the redshift $z=7.085$ quasar ULAS J1120+0641? 2011. 12
- [10] R. Chornock, E. Berger, D.B. Fox, W. Fong, T. Laskar, et al. GRB 140515A at $z=6.33$: Constraints on the End of Reionization From a Gamma-ray Burst in a Low Hydrogen Column Density Environment. 2014. 12, 14
- [11] Ryan Chornock, Edo Berger, Derek B. Fox, Ragnhild Lunyan, Maria R. Drout, et al. GRB 130606A as a Probe of the Intergalactic Medium and the Interstellar Medium in a Star-forming Galaxy in the First Gyr After the Big Bang. *Astrophys.J.*, 774:26, 2013. doi: 10.1088/0004-637X/774/1/26. 12, 14
- [12] Ryan Cooke, Max Pettini, Regina A. Jorgenson, Michael T. Murphy, and Charles C. Steidel. Precision measures of the primordial abundance of deuterium. 2013. 34, 41
- [13] Asantha Cooray and Ravi K. Sheth. Halo models of large scale structure. *Phys.Rept.*, 372:1–129, 2002. doi: 10.1016/S0370-1573(02)00276-4. 66
- [14] Rupert A.C. Croft, David H. Weinberg, Mike Bolte, Scott Burles, Lars Hernquist, et al. Towards a precise measurement of matter clustering: Lyman alpha forest data at redshifts 2–4. *Astrophys.J.*, 581:20–52, 2002. doi: 10.1086/344099. 32, 84
- [15] S.G. Djorgovski, S.M. Castro, D. Stern, and A. Mahabal. On the threshold of the reionization epoch. 2001. 25
- [16] X. Fan, M. A. Strauss, R. H. Becker, R. L. White, J. E. Gunn, G. R. Knapp, G. T. Richards, D. P. Schneider, J. Brinkmann, and M. Fukugita. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars. *Astrophysical Journal*, 132:117–136, July 2006. doi: 10.1086/504836. xi, 9, 21
- [17] Xiao-Hui Fan, Michael A. Strauss, Robert H. Becker, Richard L. White, James E. Gunn, et al. Constraining the evolution of the ionizing background and the epoch of reionization with $z \sim 6$ quasars. 2. a sample of 19 quasars. *Astron.J.*, 132:117–136, 2006. doi: 10.1086/504836. 26, 28, 35, 37, 82, 83, 112, 137, 166
- [18] Xiaohui Fan, Vijay K Narayanan, Michael A Strauss, Richard L White, Robert H Becker, Laura Pentericci, and Hans-Walter Rix. Evolution of the ionizing background and the epoch of reionization from the spectra of $z \sim 6$ quasars. *The Astronomical Journal*, 123(3):1247, 2002. 6
- [19] Xiaohui Fan et al. A Survey of $z > 5.8$ quasars in the Sloan Digital Sky Survey I: Discovery of three new quasars and the spatial density of luminous quasars at $z > 6$. *Astron.J.*, 122:2833, 2001. doi: 10.1086/324111. 25

-
- [20] Claude-Andre Faucher-Giguere, A. Lidz, M. Zaldarriaga, and L. Hernquist. The Line-of-Sight Proximity Effect and the Mass of Quasar Host Halos. *Astrophys.J.*, 673:39, 2008. doi: 10.1086/521639. 6
- [21] Steven Furlanetto, Matias Zaldarriaga, and Lars Hernquist. The Growth of HII regions during reionization. *Astrophys.J.*, 613:1–15, 2004. doi: 10.1086/423025. 32, 86, 87, 139, 141, 148
- [22] Simona Gallerani, T. Roy Choudhury, and A. Ferrara. Constraining the reionization history with qso absorption spectra. *Mon.Not.Roy.Astron.Soc.*, 370:1401–1421, 2006. doi: 10.1111/j.1365-2966.2006.10553.x. 37
- [23] J. E. Gunn and B. A. Peterson. On the Density of Neutral Hydrogen in Intergalactic Space. *ApJ*, 142:1633–1641, November 1965. doi: 10.1086/148444. 25, 82
- [24] L. Hui and N. Y. Gnedin. Equation of state of the photoionized intergalactic medium. *MNRAS*, 292:27, November 1997. 8
- [25] M. Kuhlen and C.-A. Faucher-Giguère. Concordance models of reionization: implications for faint galaxies and escape fraction evolution. *MNRAS*, 423:862–876, June 2012. doi: 10.1111/j.1365-2966.2012.20924.x. 30
- [26] Hee-Won Lee. Asymmetric Absorption Profiles of Ly and Ly in Damped Ly Systems. *Astrophys.J.*, 772:123, 2013. doi: 10.1088/0004-637X/772/2/123. 34
- [27] Adam Lidz and Matthew Malloy. On Modeling and Measuring the Temperature of the $z \approx 5$ intergalactic medium. *Astrophys.J.*, 788:175, 2014. doi: 10.1088/0004-637X/788/2/175. 32, 33
- [28] Adam Lidz, Matthew McQuinn, and Matias Zaldarriaga. Quasar Proximity Zones and Patchy Reionization. *Astrophys.J.*, 670:39–59, 2007. doi: 10.1086/521974. 10, 15, 26, 82
- [29] Ian McGreer, Andrei Mesinger, and Valentina D’Odorico. Model-independent evidence in favour of an end to reionization by $z \approx 6$. *Mon.Not.Roy.Astron.Soc.*, 447(1):499–505, 2015. doi: 10.1093/mnras/stu2449. xi, 12, 22
- [30] Ian D. McGreer, Andrei Mesinger, and Xiaohui Fan. The first (nearly) model-independent constraint on the neutral hydrogen fraction at $z \approx 5$. 2011. xi, 9, 12, 22, 26, 27, 28, 37, 82
- [31] Matthew McQuinn, Lars Hernquist, Matias Zaldarriaga, and Suvendra Dutta. Studying Reionization with Ly-alpha Emitters. *Mon.Not.Roy.Astron.Soc.*, 381:75–96, 2007. 28, 29, 32, 84, 87, 94, 110
- [32] Matthew McQuinn, Adam Lidz, Oliver Zahn, Suvendra Dutta, Lars Hernquist, et al. The Morphology of HII Regions during Reionization. *Mon.Not.Roy.Astron.Soc.*, 377:1043–1063, 2007. doi: 10.1111/j.1365-2966.2007.11489.x. 32, 100, 139, 148, 149, 152, 155, 169
- [33] A. Mesinger. Was reionization complete by $z \approx 5$? *MNRAS*, 407:1328–1337, September 2010. doi: 10.1111/j.1365-2966.2010.16995.x. 28, 37
- [34] Andrei Mesinger. Was reionization complete by $z \approx 5$? 2009. 26, 82
- [35] Andrei Mesinger and Zoltan Haiman. Constraints on Reionization and Source Properties from the Absorption Spectra of $z > 6.2$ Quasars. *Astrophys.J.*, 660:923–932, 2007. doi: 10.1086/513688. 27, 28
- [36] Jordi Miralda-Escude. Reionization of the intergalactic medium and the damping wing of the Gunn-Peterson trough. *Astrophys.J.*, 501:15–22, 1998. doi: 10.1086/305799. 28, 173
- [37] Jordi Miralda-Escude and Martin J. Rees. Searching for the earliest galaxies using the Gunn-Peterson trough and the Lyman alpha emission line. 1997. 39
- [38] Jordi Miralda-Escude, Ren-yue Cen, Jeremiah P. Ostriker, and Michael Rauch. The Lyman alpha forest from gravitational collapse in the CDM + Lambda Model. *Astrophys.J.*, 471:582, 1996. doi: 10.1086/177992. 84
- [39] Daniel J. Mortlock, Stephen J. Warren, Bram P. Venemans, Mitesh Patel, Paul C. Hewett, et al. A luminous quasar at a redshift of $z = 7.085$. *Nature*, 474:616, 2011. 12, 14
- [40] L. Pentericci, E. Vanzella, A. Fontana, M. Castellano, T. Treu, et al. New observations of $z \approx 7$ galaxies: evidence for a patchy reionization. 2014. 27
- [41] Jason X. Prochaska, Stephane Herbert-Fort, and Arthur M. Wolfe. The sdss damped lya survey: data release 3. *Astrophys.J.*, 635:123–142, 2005. doi: 10.1086/497287. 60
- [42] Alireza Rahmati and Joop Schaye. Predictions for the relation between strong HI absorbers and galaxies at redshift 3. 2013. 59
- [43] Brant E. Robertson, Steven R. Furlanetto, Evan Schneider, Stephane Charlot, Richard S. Ellis, et al. New Constraints on Cosmic Reionization from the 2012 Hubble Ultra Deep Field Campaign. *Astrophys.J.*, 768:71, 2013. doi: 10.1088/0004-637X/768/1/71. 27, 30, 92, 97
- [44] R. Salvaterra. High redshift Gamma-Ray Bursts. 2015. 15

-
- [45] M. A. Schenker, D. P. Stark, R. S. Ellis, B. E. Robertson, J. S. Dunlop, R. J. McLure, J.-P. Kneib, and J. Richard. Keck Spectroscopy of Faint $3 < z < 8$ Lyman Break Galaxies: Evidence for a Declining Fraction of Emission Line Sources in the Redshift Range $6 < z < 8$. *ApJ*, 744:179, January 2012. doi: 10.1088/0004-637X/744/2/179. 27
- [46] J. Schroeder, A. Mesinger, and Z. Haiman. Evidence of Gunn-Peterson damping wings in high- z quasar spectra: strengthening the case for incomplete reionization at $z \sim 6 - 7$. *MNRAS*, 428:3058–3071, February 2013. doi: 10.1093/mnras/sts253. 27, 28
- [47] R. A. Simcoe, P. W. Sullivan, K. L. Cooksey, M. M. Kao, M. S. Matejek, and A. J. Burgasser. Extremely metal-poor gas at a redshift of 7. *Nature*, 492:79–82, December 2012. doi: 10.1038/nature11612. 14
- [48] Jessie Taylor and Adam Lidz. What do observations of the Lyman-alpha fraction tell us about reionization? 2013. 27
- [49] Tomonori Totani, Nobuyuki Kawai, George Kosugi, Kentaro Aoki, Toru Yamada, et al. Implications for the cosmic reionization from the optical afterglow spectrum of the gamma-ray burst 050904 at $z = 6.3$. *Publ.Astron.Soc.Jap.*, 58:485–498, 2006. 137
- [50] Matteo Viel, George D. Becker, James S. Bolton, and Martin G. Haehnelt. Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman- forest data. *Phys.Rev.*, D88(4):043502, 2013. doi: 10.1103/PhysRevD.88.043502. 5, 50, 82
- [51] Arthur M. Wolfe, Eric Gawiser, and Jason X. Prochaska. Damped lyman alpha systems. *Ann.Rev.Astron.Astrophys.*, 43:861–918, 2005. doi: 10.1146/annurev.astro.42.053102.133950. 59, 61
- [52] Gbor Worseck, J. Xavier Prochaska, John M. O’Meara, George D. Becker, Sara Ellison, et al. The Giant Gemini GMOS survey of $z > 4.4$ quasars - I. Measuring the mean free path across cosmic time. 2014. 31
- [53] Yidong Xu, Bin Yue, Meng Su, Zuhui Fan, and Xuelei Chen. An analytical model of the large neutral regions during the late stage of reionization. *Astrophys.J.*, 781:97, 2014. doi: 10.1088/0004-637X/781/2/97. 30
- [54] L. Zuo and E. S. Phinney. Absorption by Discrete Inter-galactic Clouds: Theory and Some Applications. *ApJ*, 418:28, November 1993. doi: 10.1086/173367. 31, 65