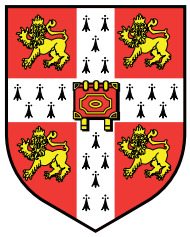


UNIVERSITY OF CAMBRIDGE

INSTITUTE OF ASTRONOMY

A THESIS SUBMITTED FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

PROBING THE INTERGALACTIC MEDIUM WITH HIGH-REDSHIFT QUASARS

ALEXANDER PETER CALVERLEY

DARWIN COLLEGE

Submitted to the Board of Graduate Studies

19th October 2011

UNDER THE SUPERVISION OF

DR. GEORGE D. BECKER

“Science is what we understand well enough to explain to a computer. Everything else we do is art.”

- DONALD KNUTH

DECLARATION

I hereby declare that my thesis entitled *Probing the Intergalactic Medium With High-Redshift Quasars* is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other University. I further state that no part of my thesis has already been or is being concurrently submitted for any such degree, diploma or other qualification. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

I note that Chapters 1 and 2 are intended as reviews, and as such contain little, if any, original work. They contain a number of images and plots extracted from other published works, all of which are clearly cited in the appropriate caption.

Those parts of this thesis which have been published or accepted for publication are as follows:

- Much of the work contained in Chapters 4 and 5 formed the basis of the work published as **Calverley, A. C.**, Becker, G. D., Haehnelt, M. G., and Bolton, J. S. (2011) 'Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect', *Monthly Notices of the Royal Astronomical Society* **412**, 2543-2562 and was completed in collaboration with the named authors.

This thesis contains fewer than 60,000 words.

Alex Calverley
Cambridge, October 19, 2011

ACKNOWLEDGEMENTS

This thesis would not exist without the continuing help and support of those around me, and there are many people who deserve a mention. The greatest thanks are reserved for my supervisor, George Becker, who chose to take me on as his student, and who has been a great source of ideas and inspiration during these past three years. I am also relatively rare amongst grad students in the IoA, in that I have had regular contact with my secondary supervisor, Martin Haehnelt. He first became involved during the writing of my paper in my second year, and has kept a careful eye on my work ever since. In many ways, the fact my thesis has been completed on time is largely down to his constant encouragement (and relentless deadlines!).

Before I could choose a supervisor though, I needed a place in the IoA first. I will forever be grateful to Bob Carswell for sending me that email offering me a place, and to the examiners who looked at my work during each year of my PhD (Max Pettini, Ian McCarthy, Richard McMahon, Anthony Challinor and Andrew Fox) and decided to let me stay! Similarly, I must thank Paul Hewett for showing himself to be *the* quasar expert, and for always being willing to answer my questions on both basic astrophysics, and topics tangential to my research.

Of course, the IoA is more than just a building full of researchers. Working behind the scenes are a dedicated team of support staff whose excellent work keeps the department working with incredible efficiency. Particular mentions must go to Siân Owen, who is quite simply the source of all knowledge, and Margaret Harding, whom I have to thank for all her sage advice when planning my observing and conference trips. I must also mention the Helpdesk staff (Hardip, Andy B, Andy D and Sue) who have always come to my rescue when my inability to use the command-line has caused my computer to throw a tantrum.

I also give thanks to STFC, who have funded my PhD studentship and given me the opportunity to spend three years studying astronomy. They, together with Darwin College, have also covered the costs of my adventures abroad, for conferences and observing. Via them I have visited both India and Hawai'i during my PhD, and these were experiences I will never forget.

My time at the IoA has been greatly enhanced by my friendships with other grad students. The most notable of these are my year group (Adrian, Amy, Becky, Dom, James M, James O, Jon, Ryan, Stephanie, Warrick and Yin-Zhe) with whom I have some great shared memories, ranging from games of 'foilball' to hiding under a particular professor's desk. A special mention must go to the subset known as the 'West's Crew' (Becky, James M and Jon), who ensured that I could spend my lunchtimes discussing current affairs away from the IoA, instead of eating alone at my desk. On a day to day basis though, it has been my officemates in H27, past and present (Caitlin, Simeon, Julie, Maciek, Adriano, Kim, Stephen and Warrick), who have kept me smiling, even when my code refuses to work. A special thank you goes out to Caitlin for teaching me some of her IDL skills (plus giving me her entire IDL library), and to Warrick for putting up with my endless requests for command-line help, questions on grammar, and showing me how to use Mathematica.

During my final year, I became much more involved in the outreach activities that take place at the IoA. These are all organised by Carolin Crawford, so I am very grateful to her for letting me take part.

Indeed, if it wasn't for Carolin, chances are I would never have heard about the job that I ultimately took, so I will always be thankful to her for forwarding on that email.

My world has not just revolved around the IoA during my time in Cambridge. I have met some great people in my college, particularly in my first year, who have been instrumental in helping me adjust to life in Cambridge. Similarly, student societies have helped me to balance my work with play. One of my greatest regrets in undergrad was never picking up rowing, so I'm very glad that Darwin College Boat Club let me make amends for that! Also, I am particularly grateful to the excellent characters I have met on the Cambridge music scene, and I will always feel honoured to have made music with such a talented band of musicians.

There have been several great science teachers during my life who have inspired me to pursue my dream of studying astronomy, although none more so than my MSci thesis supervisor, Alastair Edge. My project with him helped me realise that I enjoyed doing research, and with his encouragement I sent in my application to the IoA. I hope that through this thesis, I have lived up to the potential that he saw in me.

No acknowledgements would be complete without mentioning the tireless love and support of my parents. They have always been there for me, and I cannot thank them enough for everything they have done. They also formed one half of my diligent team of thesis proof-readers. The other half is someone very special to me, who has been at my side through it all. Sarah, thank you so much for all that you have brought to my life.

This thesis has been typeset in LaTeX using TeXnic Center and JabRef. Thanks to all the former IoA members who have contributed to the thesis template on which this is based, and particularly to Mark Booth from whom I received this version.

Alex Calverley
Cambridge, October 19, 2011

SUMMARY

Probing the Intergalactic Medium with High-Redshift Quasars

Alex Calverley

Clues about the timing of reionization and the nature of the ionizing sources responsible are imprinted in the ionization and thermal state of the IGM. In this thesis, I use high-resolution quasar spectra in conjunction with state-of-the-art hydrodynamical simulations to probe the IGM at high redshift, focusing on the ionization and thermal state of the gas.

After reionization, the ionization state of the IGM is set by the intensity of the ultraviolet background (UVB), quantified by the hydrogen photoionization rate, $\Gamma_{\rm bkg}$. At high redshifts this has been estimated by measuring the mean flux in the Lyα forest, and scaling $\Gamma_{\rm bkg}$ in simulations such that the simulated mean flux matches the observed value. In Chapter 3 I investigate whether the precision of these estimates can be improved by using the entire flux probability distribution function (PDF) instead of only the mean flux. Although I find it cannot improve the precision directly, the flux PDF can potentially be used to constrain other sources of error in observational estimates of $\Gamma_{\rm bkg}$, and so may increase the precision indirectly.

The ionizing output of a quasar will locally dominate over the UVB, and this leads to enhanced transmission bluewards of the quasar Lyα line, known as the proximity effect. In Chapter 4 I present the first measurements of $\Gamma_{\rm bkg}$ at $z > 5$ from the proximity effect. The UVB intensity declines smoothly with redshift over $4.6 < z < 6.4$, implying a smooth evolution in the mean free path of ionizing photons. This suggests that reionization ends at $z > 6.4$. There is a drop in $\Gamma_{\rm bkg}$ by roughly a factor of five, which corresponds to a drop in the ionizing emissivity by about a factor of two. Such a redshift evolution in the emissivity cannot continue to much higher redshift without reionization failing to complete, which suggests that reionization cannot have ended much higher than $z = 6.4$.

Estimates of $\Gamma_{\rm bkg}$ from the proximity effect and the mean flux are generally discrepant at $z \sim 2-4$, with those from the proximity effect systematically higher. This is generally attributed to effects of the quasar environment. I investigate the significance of several environmental biases on proximity effect measurements at $z \sim 5-6$ in Chapter 5. The biases are found to be small, and so the proximity effect is expected to give relatively unbiased estimates of $\Gamma_{\rm bkg}$ at $z > 5$, in contrast to lower redshifts.

Photoionization heats the gas in the IGM, and so the thermal history of the IGM provides important constraints on reionization. The thermal state of the IGM is reflected in the level of small-scale structure in the Lyα forest. In Chapter 6 I quantify the small-scale structure using two independent statistics, the curvature and the peakiness, and convert these into a temperature by comparing with simulations. These are the first measurements of the temperature in the general IGM at $z > 5$. Both statistics show an increase in the temperature by a factor of roughly two from $z = 4.4$ to 5.6. This rise is sensitive, however, to any smoothing of the gas density distribution due to the thermal history spanning reionization. I find that this should only be a small effect, as otherwise the corrected

temperatures at $z \sim 4-5$ are implausibly low. The temperature evolution therefore suggests a late reionization. The temperatures at $z \geq 4.8$ are well fit by an adiabatic cooling curve, for which reasonable peak temperatures at the end of reionization are reached at $6 \lesssim z \lesssim 7$. The temperatures at $z \sim 4-5$ are consistent with reionization being carried out by Pop II stars.

In conclusion, the ionization and thermal state of the IGM at $z \sim 5-6$ suggest a late hydrogen reionization, driven by star-forming galaxies and ending around $6.5 \lesssim z \lesssim 7$. This is consistent with other recent lines of observational evidence, and supports theoretical models that infer a late reionization from the observed star formation rate history.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

The beginning is the most important part of the work.

The Republic

PLATO

1 Introduction

EVER since the first human gazed at the stars, we have sought to know how the Universe works; to understand its beginning and evolution. Previously these endeavours were the realm of philosophers and theologians, yet in the past 100 years we have made greater strides towards answering them than in any previous century. Huge advances in observational technology have meant that we can now see objects at almost unimaginable distances, equivalent to looking back over immense periods of time, and even glimpse the smoldering remains of the creation of the Universe. Similarly, the exponential increase in computing power has allowed us to start to simulate many astronomical phenomena, from the formation of a galaxy to the formation of a planet. From this privileged position of knowledge, it is worth looking back on a century of astronomical progress. One hundred years ago there was still a debate as to whether the Messier objects (Messier, 1781) were within the Milky Way or not. Slipher (1913) noticed a shift in the emission lines of the spectra from the Andromeda Nebula[1], which implied that it was heading towards the Solar System at a speed of $\sim 300\ \mathrm{km\,s^{-1}}$, the fastest speed ever recorded for an astronomical object. Additional observations of other nebulae found NGC 4594 (now known as the Sombrero Galaxy) travelling away from the Solar System at $\sim 1100\ \mathrm{km\,s^{-1}}$, and that the average recessional velocity of spiral nebulae[2] was about 25 times the average stellar velocity (Slipher, 1915). This shift of the spectral lines to redder wavelengths due to the object's recessional velocity was coined 'red-shift' by Walter S. Adams in 1908 (although over time the

[1] From the Latin for 'cloud', nebula was originally the general name for any extended astronomical object.

[2] A distinct subset of nebulae defined as having an apparent spiral structure. Another subset, named for similar observational reasons, were planetary nebulae.

hyphen was dropped, e.g. in Zwicky, 1929). The redshift, z, of an object is defined as

$$z = \frac{\lambda_{\rm observed}}{\lambda_{\rm emitted}} - 1 = \sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}} - 1 \approx \frac{v}{c} \tag{1.1}$$

where v is the recessional velocity and c is the speed of light. The final approximation is only valid when $v \ll c$. Based on Slipher's measurement, NGC 4594 therefore has a redshift $z \sim 0.003$. Further redshift measurements for other nebulae would imply increasingly large v, and the astrophysical significance of redshift was widely debated. Finally, Hubble (1925) showed that, based on observations of Cepheid stars[3], at least one of these objects, NGC 6822, was $\sim 700,000$ light years away, and thus must be of extragalactic origin[4]. Far from being the exception, many of the Messier objects were reclassified from nebulae to galaxies, with a new, higher redshift object discovered every few years. Hubble (1929) demonstrated a link between v and the distance, d, to the galaxy, such that

$$v = H_0 d \tag{1.2}$$

where H_0 is the "Hubble constant". By comparison with the previous equation it is therefore clear that large redshifts correspond to large distances, and, due to the finite speed of light, we are seeing these galaxies as they looked in the past. Today the current redshift record holders are from the Hubble Ultra Deep Field[5], which has yielded UDFy-38135539 at $z = 8.5549 \pm 0.0002$ (Lehnert et al., 2010, although this spectroscopic redshift is somewhat controversial), and UDFj-39546284 with a photometric redshift of $z \sim 10.3$ (Bouwens et al., 2011a). These objects take us to within 500 Myr of the beginning of the Universe, and represent the frontiers of current observational technology. In the rest of this Chapter I will describe what we know about the high-redshift ($z \gtrsim 5$) Universe, starting with a historical summary of the current cosmological model and the formation of the first stars and galaxies. I will then turn to the impact of the ionizing photons from these galaxies on the surrounding intergalactic medium (IGM), which is a major theme of this thesis.

1.1 The standard cosmological model

Edwin Hubble's formulation of his law in Equation 1.2 was arguably the first evidence that the Universe is expanding. However, before publishing his law there had already been several theoretical motivations for an expanding Universe. The field equations of general relativity (Einstein, 1915) implied a Universe that could not be static, and so Einstein (1917) introduced a "cosmological constant", Λ, that would counteract the attractive effects of gravity. From these equations, Friedmann (1922) derived a

[3] Cepheid stars vary in brightness, and follow a well established period-luminosity relation. From an observed apparent magnitude and a derived absolute magnitude based on the period of the brightness variation, a distance to the star can be determined

[4] The Milky Way was thought to be only $\sim 30,000$ light years across

[5] http://www.nasa.gov/vision/universe/starsgalaxies/hubble_UDF.html

solution for an isotropic and homogeneous Universe[6]. Out of this solution came two equations that describe the dynamics of an expanding Universe,

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \tag{1.3}$$

known as the Friedmann equation, and

$$\dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \tag{1.4}$$

known as the Friedmann acceleration equation[7]. In both of these, Λ is the cosmological constant, $H \equiv \dot{a}/a$ is the Hubble parameter, G is Newton's gravitational constant, ρ is the density of the Universe, k/a^2 is the spatial curvature, p is the pressure of the Universe, and a is the 'scale factor'. The latter is a dimensionless quantity that relates the proper (i.e. physical) distance between two objects moving with the expansion of the Universe to the distance at some reference time, normally taken to be the present day. Mathematically, if $d(t)$ is the proper distance between two objects at time t and d_0 is the distance now, then

$$d(t) = d_0 a(t) \quad \Rightarrow \quad \dot{d}(t) = d_0 \dot{a}(t) \text{ and } d_0 = \frac{d(t)}{a(t)}\,, \tag{1.5}$$

and by combining these with the above definition of the Hubble parameter

$$\dot{d}(t) = \frac{\dot{a}(t)}{a(t)} d(t) = H d(t)\,. \tag{1.6}$$

Consequently, Hubble's Law (Equation 1.2) is a prediction of the Friedmann equations, although Friedmann himself never explicitly did this derivation. His solution to the Einstein field equations was derived independently by Lemaître (1927), who also explicitly predicted Hubble's Law, as well as laying the first conceptual foundations of what would come to be known as the 'Big Bang Theory'[8]. In his paper, Lemaître proposed that since the Universe is now expanding, at some point in the past everything must have been much closer, hotter and denser than it is today. He would go on to suggest that at the creation of the Universe there was a 'primeval atom' (e.g. Lemaître, 1931). The idea of an expanding Universe grew in popularity, until in the 1940s there were only two distinct cosmologies being researched - the 'Big Bang' model of Universal expansion (e.g. Gamow, 1948), and the 'Steady State' model of continuous creation (Bondi and Gold, 1948; Hoyle, 1948). Two major predictions of the 'Big Bang' model were the primordial ratios of the light elements formed during Big Bang nucleosynthesis (Alpher et al., 1948), and that the radiation from that initial fireball would today have a blackbody spectrum with a temperature of ~ 5 K (Alpher and Herman, 1948). The debate was fi-

[6] Known as the Cosmological Principle, it is usually stated formally as 'Viewed on a sufficiently large scale, the properties of the Universe are the same for all observers.'

[7] The dot notation is used to represent derivatives with respect to time.

[8] The metric used by both Friedmann and Lemaître to solve the field equations and derive the Friedmann equations was based purely on geometrical considerations of mathematically representing a homogeneous and isotropic Universe. However, Robertson (1935) and Walker (1935) went on to show that that metric is the only allowed one on a spacetime that is homogeneous and isotropic. Consequently, the metric is often referred to as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric.

Parameter	Value	Description
H_0	70.2 ± 1.4	Hubble constant, in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$
Ω_m	0.275 ± 0.017	Contribution to Ω by matter
Ω_c	0.229 ± 0.015	Contribution to Ω by cold dark matter
Ω_b	0.0458 ± 0.0016	Contribution to Ω by baryons
Ω_Λ	0.725 ± 0.016	Contribution to Ω by dark energy
Ω	1.000 ± 0.033	Present-day density parameter
σ_8	0.816 ± 0.024	Amplitude of the mass density fluctuations
n_s	0.968 ± 0.012	Slope of the initial power spectrum
τ	0.088 ± 0.014	Thomson optical depth to CMB photons
t_0	13.76 ± 0.11	Present-day age of the Universe (Gyr)

Table 1.1: A selection of parameters taken from the mean 7-year WMAP results, combined with data on the sound horizon size for baryon acoustic oscillations (BAO) from the Two-Degree Field Galaxy Redshift Survey (2dFGRS) and the Sloan Digital Sky Survey Data Release 7 (SDSS DR7; Percival et al., 2010), and independent measurements of the Hubble constant from supernovae data (Riess et al., 2009). All figures taken from Table 1 in Komatsu et al. (2011).

nally resolved when Penzias and Wilson (1965) detected a homogeneous and isotropic radiation field with a blackbody spectrum and a temperature of 3.0 ± 1.0 K, leaving the Big Bang Theory (BBT) as the dominant cosmological theory.

Since the 1960s we have made considerable progress towards understanding the details of an expanding Universe. Several of the outstanding problems of the original formulation of the BBT have been dealt with through additions to the theory (e.g. inflation, as proposed by Guth, 1981). Improved data from satellites and ground-based measurements designed to measure the nearly homogeneous radiation field, known as the Cosmic Microwave Background (CMB), have also led to giant leaps in our understanding. The Cosmic Background Explorer (COBE) probe confirmed that the CMB spectrum was a nearly perfect blackbody with $T \sim 2.7$ K (Mather et al., 1990), and also detected for the first time tiny fluctuations in the temperature of the CMB, of order $\delta T/T \sim 10^{-5}$ (Smoot et al., 1992). Subsequent measurements with the more sensitive Wilkinson Microwave Anisotropy Probe (WMAP; Spergel et al., 2003) have led to a wealth of knowledge, and will be added to in the future with Planck (due to give results early 2013). The constraints on several key cosmological parameters from 7 years of WMAP observations are summarised in Table 1.1.

The recent work with the CMB has provided unprecedented insight into the bulk contents of the Universe. Zwicky (1933) measured the mass of the Coma cluster and found it to be much higher than that expected purely from the stars, and so was the first to suggest that there must be significant quantities of "missing" matter. Further evidence came from flat galactic rotation curves (e.g. Rubin et al., 1980), but it was only with the CMB that the universal ratio of 'dark matter'[9] to baryonic matter could be determined. Subsequent simulations of galaxy formation showed that the large-scale structure could be accurately reproduced if dark matter consisted of massive particles moving at speeds $< 100\ \mathrm{km\,s^{-1}}$, called 'Cold Dark Matter' (e.g. Peebles, 1982; Bond and Szalay, 1983; Blumenthal et al., 1984). Dark matter represents about 85% of the total matter in the Universe, but it is not the dom-

[9] It is called 'dark' as it does not emit any light.

inant mass-energy component. Surprising findings from studies of distant supernovae have shown that the expansion of the Universe is accelerating (Riess et al., 1998; Perlmutter et al., 1999), suggesting the existence of 'dark energy'. CMB measurements reveal that this is about 73% of the contents of the Universe, and that it acts like a cosmological constant, Λ, in the Friedmann equations. Studies of the CMB have led to one other unexpected conclusion. From Equation 1.3 one can calculate for a basic FLRW Universe (i.e. $\Lambda = 0$) the critical density, $\rho_{\rm crit}$, for a flat ($k = 0$) geometry as

$$\rho_{\rm crit} = \frac{3H^2}{8\pi G} \,. \tag{1.7}$$

The density parameter, Ω, can then be defined as

$$\Omega \equiv \frac{\rho}{\rho_{\rm crit}} = \frac{8\pi G}{3H^2}\rho \,. \tag{1.8}$$

Often the flat geometrical case was used in cosmological models, since it was the most simple; however, the CMB showed that when the total mass-energy of the Universe is summed up, including the contributions from dark matter and dark energy, then $\rho \approxeq \rho_{\rm crit}$. Consequently $\Omega \approxeq 1$ and so the Universe is (to within experimental error) flat.

This geometrically-flat, dark energy-dominated cosmological model, where the majority of the mass is in cold dark matter (CDM), is known as the ΛCDM model of Big Bang cosmology, and is often referred to as the 'Standard Cosmological Model'. As outlined in this section, it is the product of many decades of research, and remains the most widely accepted explanation of the existence and structure of the CMB, the large-scale structure of the Universe, the observed ratios of the light elements in primordial gas and very old stars, and the accelerating expansion of the Universe. All the results in this thesis are presented within the ΛCDM cosmological framework. The rest of this section looks at the history of the Universe in the context of this model, and is summarised in Fig. 1.1.

1.1.1 The growth of structure

In the ΛCDM model the Universe begins in a very hot, dense state at the Big Bang. Whilst the very first instant is highly speculative, we have a better understanding from $\sim 10^{-11}$ seconds onwards, as the relevant energies are achievable in particle accelerators on Earth. Stable baryons have formed by ~ 3 minutes after the Big Bang, and then during the next ~ 17 minutes the temperatures are hot enough for nuclear fusion, and thus the first atomic nuclei are created. The brevity of this period means that the majority of nuclei formed are ^{1}H and ^{4}He, and trace amounts of ^{2}H (deuterium), ^{3}He, and metals[10] such as ^{6}Li and ^{7}Li (although nothing heavier than beryllium). Once Big Bang nucleosynthesis stops, these primordial elemental abundances are essentially frozen into the gas, and are an important prediction of the BBT. Although the nuclei have formed, the Universe is still too hot to form neutral atoms, and thus it consists of an optically thick, fully ionized plasma. This state continues until $z \sim 1090$ ($\sim 380,000$ years after the Big Bang), when the Universe has cooled enough for

[10] In astronomy, any element heavier than helium is called a "metal".

hydrogen atoms to form out of the plasma of nuclei and electrons, a process known as *recombination*[11]. In a very short space of time the whole Universe goes from being fully ionized to fully neutral. Consequently, there are now far fewer free electrons and protons for photons to Thomson scatter off. The photons therefore travel freely, and so the Universe is now transparent to radiation. This 'first light' is detected today as the CMB.

Following recombination comes a period known as the cosmic 'dark ages' (Rees, 1998), so named due to the absence of luminous sources, in which essentially all the baryons are contained in a diffuse neutral gas. Whilst this gas is highly homogeneous, it is not *perfectly* so. Tiny quantum fluctuations, greatly amplified by inflation and visible as the anisotropies in the CMB temperature, cause some areas to be very slightly overdense and some underdense with respect to the mean density. These are the building blocks of the first structures. Gravity causes the density contrast between over- and underdense regions to increase with time. Since it is governed only by gravity, the dark matter in the Universe forms a patchwork of dense haloes, filamentary structures, and large voids, referred to as the 'cosmic web' (Bond et al., 1996). However, the gas is still rather hot, and is unable to gather efficiently in the potential wells of the dark matter haloes. A dark matter halo's capability to trap a baryon and keep it in a stable circular orbit is governed by the virial theorem. For a potential with the form $V(r) \propto r^n$ where r is the inter-particle distance, a particle orbit is stable if

$$\langle KE \rangle = \frac{n}{2} \langle PE \rangle \ , \tag{1.9}$$

with $\langle KE \rangle$ the average kinetic energy, and $\langle PE \rangle$ the average potential energy. For a gravitational potential $n = -1$. Since the dark matter is "cold" and can be modelled as a pressureless fluid, it can collapse and become virialised at $z \sim 100$. For dark matter consisting of weakly interacting massive particles (WIMPs; the current leading model predicted by supersymmetry) these first dark matter haloes[12] would have masses similar to that of the Earth, $\sim 10^{-6} \mathrm{M}_\odot$ (Diemand et al., 2005). For the gas the situation is somewhat different. The virial temperature, $T_{\rm vir}$ of a halo of mass, M, is defined as

$$T_{\rm vir} = \frac{\mu m_{\rm p} G M}{2 k_{\rm B} r_{\rm vir}} \propto M^{2/3} \tag{1.10}$$

where μ is the mean molecular weight, $m_{\rm p}$ is the proton mass, G is the gravitational constant, $k_{\rm B}$ is the Boltzmann constant, and $r_{\rm vir}$ is the virial radius. As an overdense region begins to collapse, the gas is collisionally heated and becomes pressure supported, halting the collapse. It can then only continue when the haloes are massive enough that the virial temperature of the halo is higher than the temperature of the gas. This can happen for (relatively) low mass haloes if the gas can cool radiatively. Above $\sim 10^4$ K this is done via the line emission of atomic transitions in H and He, but to cool below this temperature requires the rotational and vibrational emission lines of molecular hydrogen. Finally, the radiative cooling is unable to cool the gas below the temperature of the CMB, which is adiabatically cooling as $T_{\rm CMB} \approx 2.7(1+z)$ K. The redshift at which these three factors coincide (the

[11] Despite the name, this is (working chronologically) actually only the first time that the nuclei and electrons have combined to form atoms.

[12] A dark matter halo is defined as being in virial equilibrium.

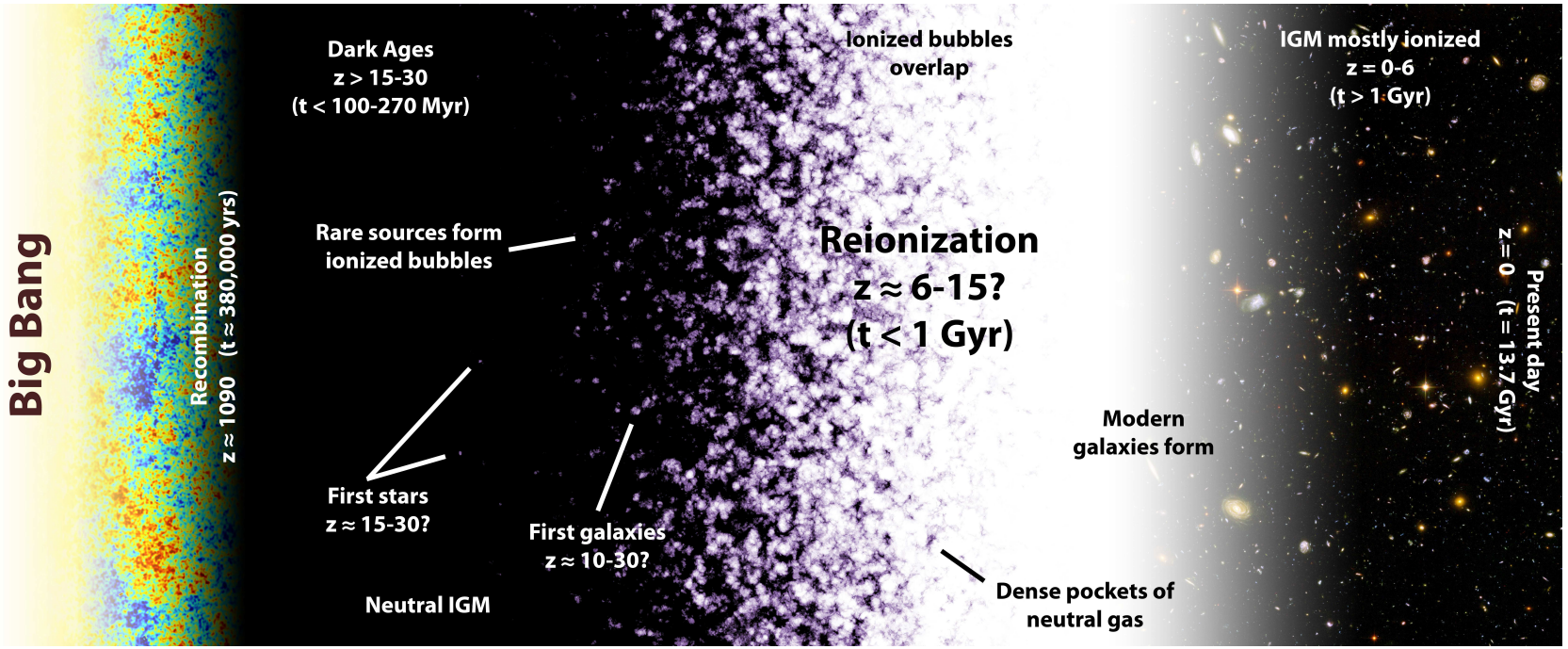


Figure 1.1: An outline of the history of the Universe. After the Big Bang, the gas in the Universe was fully ionized, until at recombination when the gas becomes fully neutral. This starts the cosmic Dark Ages, during which the gas cools to form the first stars and galaxies. They ionize the surrounding hydrogen gas, and the ionized bubbles slowly expand, until they overlap. This process wherein the IGM transitions from a fully neutral to highly ionized state is termed reionization. It ended about 1 Gyr after the Big Bang, and the Universe has been kept highly ionized since then by the ionizing radiation of star-forming galaxies and gas accretion onto supermassive black holes. This net radiation field is called the ultraviolet background. Image text adapted from Robertson et al. (2010).

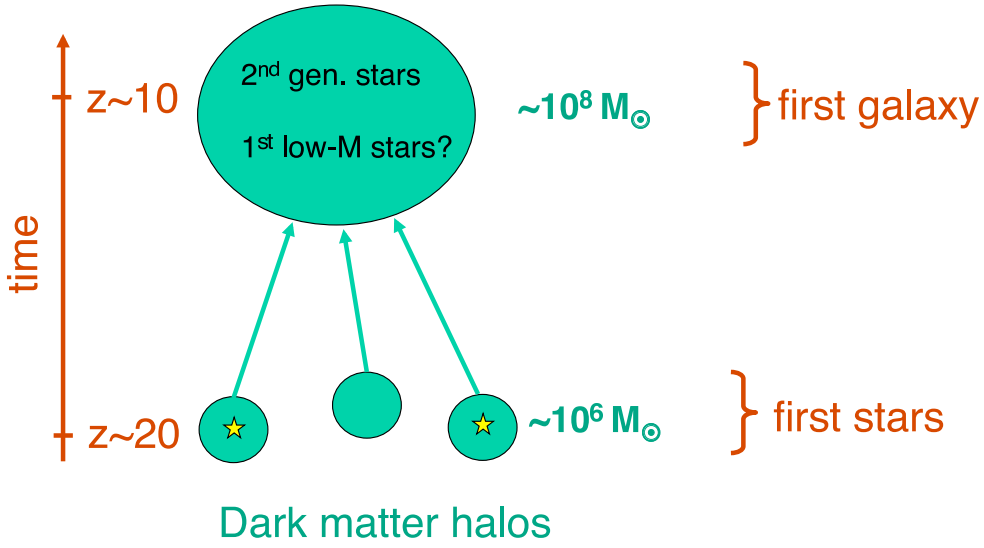


Figure 1.2: Figure showing the hierarchical growth of the first structures. The first stars form in dark matter haloes with mass $\sim 10^6 M_\odot$, however mergers of these first star groups leads to fragmented clumps of star forming regions within a dark matter halo of mass $\sim 10^8 M_\odot$, essentially one of the first galaxies. Figure taken from Bromm and Yoshida (2011).

increasing mass and $T_{\rm vir}$ of dark matter haloes, the increasing fraction of H_2 available for molecular cooling, and the decreasing CMB lower temperature limit) is generally expected to be in the region of $z \sim 15-30$ (e.g. Couchman and Rees, 1986). The first clouds able to successfully collapse are in dark matter 'minihaloes' with $M \sim 10^6 M_\odot$ ($T_{\rm vir} \sim 300$ K) (Tegmark et al., 1997), and these go on to form the first stars (see Section 1.2.1).

Galaxies form from essentially a scaled up version of this process, as extensively discussed and simulated in the literature (e.g. Press and Schechter, 1974; Silk, 1977; White and Rees, 1978; Frenk et al., 1988; White and Frenk, 1991; Cole et al., 2000; Springel et al., 2005). Within ΛCDM , structures grow via hierarchical formation, leading to increasingly massive haloes as time progresses. Those that grow to a mass $\sim 10^8 M_\odot$ ($T_{\rm vir} \sim 10^4$ K) are able to fulfill the Rees and Ostriker (1977) criterion that the gas cooling timescale, $t_{\rm cool}$, is less than the dynamical timescale, $t_{\rm dyn}$, with both shorter than the Hubble time, $t_H \simeq 1/H(z)$ (e.g. Bromm and Yoshida, 2011). Provided that efficient cooling continues, the gas will collapse under self-gravity and fragment into star forming regions of order $\sim 10^6 M_\odot$, as discussed above. The potential wells of the overall ensemble are deep enough to revirialise the gas affected by star formation in the minihaloes, and retain the gas heated by photoionization. As such, these structures could be considered the first dwarf galaxies. The relationship is shown in Fig. 1.2.

Not all baryons lie within the dark matter haloes. The large majority trace the underlying dark matter structure, and so can be found in dense filaments between galaxies, as well as in the voids. This gas is known as the intergalactic medium (IGM), and it traces the large-scale structure of the Universe (since the luminous objects are largely concentrated in galaxies which represent only a small fraction of the volume and mass of the Universe). The IGM also contains information about previous epochs of star formation, as it bears the imprint of feedback from metal enrichment, winds and ionizing photons. It is the last of these feedback process that will now be discussed.

1.1.2 Reionization

Shortly after the first sources of ionizing photons form (for example, the first stars) they begin to ionize the surrounding (neutral) gas. This is the start of another phase transition where the Universe, probably in an extended process, moves from being fully neutral to nearly fully ionized, as we see it today. This is known as the *epoch of reionization* (EoR) as it is the second time the Universe was essentially fully ionized (the first time being before recombination). There have been several excellent reviews during the last ten years on the topic of reionization and the relevant physics of the IGM (e.g. Barkana and Loeb, 2001; Loeb and Barkana, 2001; Ciardi and Ferrara, 2005; Fan et al., 2006a; Barkana and Loeb, 2007; Meiksin, 2009; Choudhury, 2009; Trac and Gnedin, 2009; Robertson et al., 2010). The currently adopted view of reionization will now be briefly summarised.

When an ionizing source turns on, it carves out an ionized region around itself (Shapiro, 1986; Donahue and Shull, 1987; Shapiro and Giroux, 1987; Meiksin and Madau, 1993) often referred to as a "HII bubble", and akin to the spheres of ionized gas found around large stars in the galaxy, known as "Strömgren spheres" (Strömgren, 1939, although these are in photoionization equilibrium). Whilst each of these bubbles is rather distinct, this part of reionization is known as 'pre-overlap'. So long as the sources continue to produce ionizing photons the ionization fronts expand, encompassing clusters of dwarf galaxies. Current theory suggests that ionized regions around individual galaxy clusters begin to overlap when they reach a characteristic size of ~ 10 proper Mpc (Furlanetto et al., 2004b; Wyithe and Loeb, 2004). This section of reionization is known as the 'overlap' phase, or 'percolation' (Fan et al., 2006a). From here the mean free path of ionizing photons increases rapidly, and is now limited by the mean distance between dense neutral clumps, known as Lyman limit systems, rather than the mean separation between sources. From an observational point of view, the completion of the overlap phase represents the end of the epoch of reionization, although there is still a very small fraction of neutral gas left in dense systems in the IGM[13]. These are slowly ionized over time by an increasing background of ionizing photons (see Section 1.4).

This description of reionization is known as an 'inside-out' picture, where high density regions are ionized first, and low density regions last. There has been some debate as to how important recombinations would be, as they will be much more significant in high density regions (Miralda-Escudé et al., 2000). Dense regions may therefore be the sources of the ionizing photons but also become neutral easily, leaving most of the ionized gas initially in the low density voids (known as the 'outside-in' picture). However, recent simulations and theoretical models prefer the inside-out topology as described above (Furlanetto et al., 2004b; Choudhury et al., 2009; Friedrich et al., 2011).

After Big Bang nucleosynthesis the main constituents of the Universe were hydrogen and helium gas and, due to the different energies for full ionization, each of these species has its own EoR. When discussing hydrogen, the process is the ionization of atomic neutral hydrogen, H, to atomic ionized hydrogen, H^+. In conventional astronomical notation these species are symbolised as HI and HII respectively[14]. The energy of this transition is 13.6 eV (1 eV = 1.60×10^{-19} J), equivalent to a pho-

[13] Although the IGM has only a tiny ionized fraction, the neutral fraction of the Universe is $\sim 5-10\%$ at $z = 3$ due to Damped Lyα systems (see Section 1.4.1).

[14] Technically, HI and HII are the relevant spectra of absorption and emission lines, rather than the ions themselves.

ton wavelength of 912 Å (1 Å = 10^{-10} m). The single ionization of HeI to HeII has an ionization energy (E = 24.6 eV) that is also within reach of HI ionizing sources (see Figure 1.3), and so it is expected to occur at the same time as hydrogen reionization. Thus, conventionally helium reionization refers to the ionization of HeII into doubly ionized helium, HeIII , (E = 54.4 eV), as only this epoch is likely to be distinct.

Reionization has a two-fold impact on the Universe. First, it changes the ionization balance of the gas, and thus the species available with which the gas can cool (hence affecting star formation). Secondly, photoionization heats the gas through the photoelectric effect, disrupting the collapse of smaller virialised haloes. Consequently, the reionization history of the Universe becomes a crucial factor in models of structure formation.

What follows in the rest of this Chapter is a summary of recent advances made in our understanding of the sources of reionization, a review of observational constraints on the reionization history from recent measurements made on a wide variety of phenomena, and finally the emergence of the metagalactic ultraviolet background and its relationship to the transmitted flux in quasar spectra.

1.2 Sources of reionization

The timing and duration of reionization strongly depends on the nature of the ionizing sources. There are several possible candidate sources, summarised in Fig 1.3, and each of which will now be discussed.

1.2.1 First stars

As described in detail in Section 1.1.1, tiny fluctuations in the density field implanted during inflation set the seeds for patches of neutral gas to gravitationally collapse and form the first stars (see Bromm and Larson, 2004, for a recent review on their formation and lifecycle). Baade (1944) was the first to suggest the idea of classifying stars into different stellar populations, and under his historical classification scheme these are known as Population III (Pop III) stars. They are thought to be the first luminous objects to emerge out of the 'dark ages', forming in dark matter minihaloes of mass $\sim 10^6 M_\odot$. Pop III stars would have had zero metallicity and are predicted to have formed at $z \gtrsim 15$ (e.g. Couchman and Rees, 1986; Haiman et al., 1996; Gnedin and Ostriker, 1997; Tegmark et al., 1997; Yoshida et al., 2003).

These first stars are likely to have been important ionizing sources for two key reasons. First, zero-metallicity stars produce nuclear energy through a different mechanism to metal-enriched stars (e.g. Marigo et al., 2001; Bromm et al., 2001). In the absence of metals, hydrogen is burnt into helium via the proton-proton chain (p-p chain). If, however, there is a tiny fraction of carbon in the core of the star then this can catalyse the burning of hydrogen via a process known as the CNO cycle[15]. The energy output of the p-p chain has a far weaker temperature dependence than the CNO cycle, so in

[15]So called as it uses additions of hydrogen nuclei to carbon to form first nitrogen, and then oxygen which decays back into carbon, emitting a helium nucleus in the process

order to produce the pressure required to balance the gravity of a star of a given mass, the central temperature needs to be higher. The opacity of stellar material is also lower at zero metallicity, allowing steeper temperature gradients and more compact configurations. As a consequence, a metal-free star should have a radius ~ 5 times smaller and a surface temperature ~ 2 times hotter than a metal-rich star of the same mass (Tumlinson et al., 2003). Secondly, the very first stars would have formed from a neutral gas consisting almost entirely of hydrogen and helium. As mentioned in Section 1.1.1, this means that the main cooling channel available to the gas is the collisional excitation of the rovibrational modes of H_2 molecules, down to temperatures of $\sim 100-200$ K. At these temperatures the mass of the collapsing cloud is $\sim 1000 M_\odot$, and due to the efficient accretion on the proto-stellar core, the masses of these first stars are expected to be $\gtrsim 100 M_\odot$ (e.g. Bromm et al., 2002; Marigo et al., 2003). The overall effect of these two factors is that Pop III stars are thought to be both massive (implying a hot surface temperature, regardless of population) and have a far hotter surface temperature than that found in other types of star. This would have made them excellent sources of early ionizing photons, and these stars can emit as many as 10^5 ionizing photons per baryon during their life (Tumlinson and Shull, 2000; Bromm et al., 2001; Schaerer, 2002).

There are several complications within this picture, however. The Strömgren spheres created around these first stars would have a substantial impact on subsequent star formation. They would heat the gas and photodissociate the H_2 around them, suppressing the collapse of similar objects. However, such high mass, high luminosity stars have very short lifetimes (~ few million years), so after their death the gas can cool and recombine. Since the gas has a high ionization fraction it is much easier to create a larger fraction of H_2 and HD (deuterated hydrogen) molecules. These allow the gas to cool more effectively, and thus fragment into smaller clouds and much lower mass (but still metal-free) stars. As such the massive stars unaffected by previous star formation are termed Pop III.1 and those that form out of the ionized gas left behind are termed Pop III.2 (McKee and Tan, 2008). Very recent simulations have cast doubt on the idea of Pop III.1 stars being as massive as widely believed (Tumlinson, 2006; Turk et al., 2009; Stacy et al., 2010; Clark et al., 2011), but the masses of Pop III stars are still broadly believed to be $\gtrsim 10 M_\odot$ (for a very recent and detailed review see Karlsson et al., 2011).

Although this radiative feedback is important, the most dominant effect is the chemical feedback of these stars. Since they are massive, they live short lives that often end as supernovae, polluting the initially pristine gas with metals. These provide an alternative and much more efficient cooling channel for the gas, using the fine-structure lines of [C II], [O I], [Si II] and, eventually, [Fe II]. Since the gas cooling is then essentially limited to $T_{\rm CMB}$, the fragments that become stars are much less massive. These less massive, metal-poor stars also burn via the CNO cycle at their core, and are known as Population II (Pop II) stars. The critical metallicity of the gas at which the star formation changes from the top-heavy initial mass function (IMF) of Pop III stars (characteristic mass $\gtrsim 10 M_\odot$) to the less massive Pop II IMF (characteristic mass $\sim 1 M_\odot$) is around $Z_{\rm crit} \sim 10^{-3.5} Z_\odot$ (Schneider et al., 2002; Bromm and Loeb, 2003), although this will vary with the details of the environment (Santoro and Shull, 2006)[16].

[16] Z denotes the mass fraction contributed by all metals.

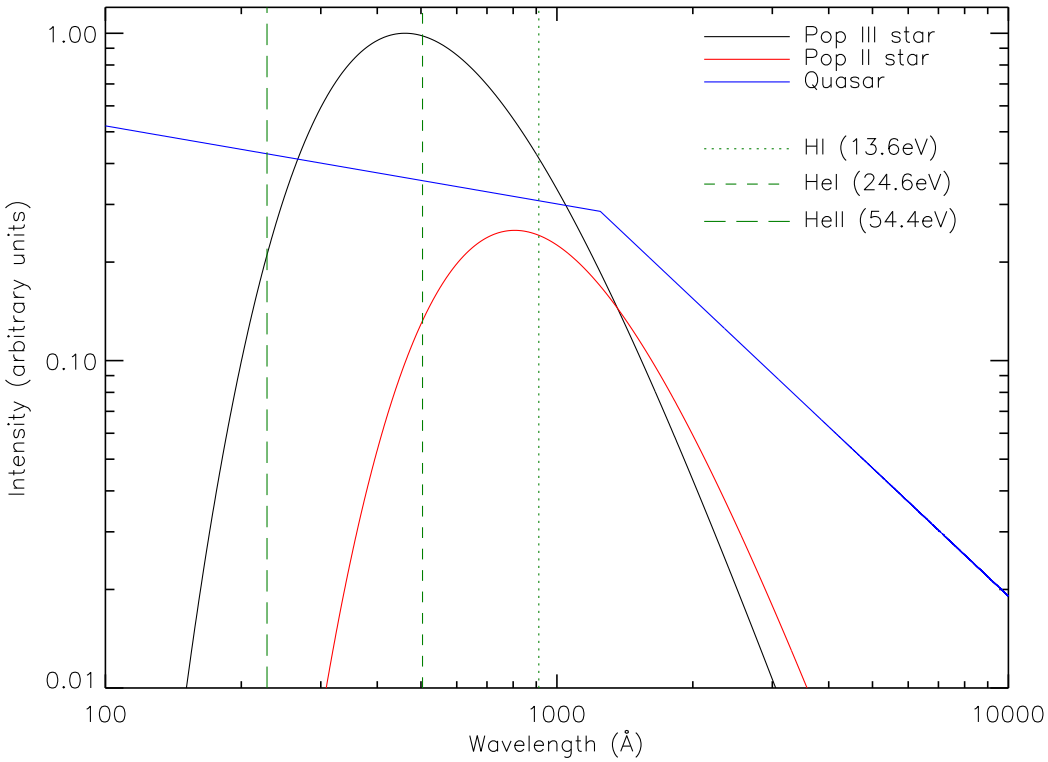


Figure 1.3: The continuum spectra of the three most probable sources of ionizing photons to carry out reionization. The black line is the continuum of a $15M_\odot$ Pop III star (zero metallicity). It radiates as a blackbody, with an effective temperature of 63,000 K. The red line is the continuum of a Pop II star of the same mass, with metallicity $Z = 10^{-3}Z_\odot$. Due to the presence of metals it has a lower effective temperature of 36,000 K. Both of these temperatures are taken from Tumlinson and Shull (2000). The blue line is the continuum spectra of a typical quasar. It is represented as a broken power law, based upon the composite from Telfer et al. (2002). Both Pop III and Pop II stars can ionize hydrogen (dotted green line; 912 Å) and singly ionize helium (dashed green line; 504 Å). Pop III stars are also able to doubly ionize helium (long dashed green line; 228 Å), however a $15M_\odot$ star will only be on the main sequence for $\sim 10^7$yr (Tumlinson et al., 2003), and so it is likely that helium was not fully reionized until quasars with their hard spectra were sufficiently numerous to do so, at around $z \sim 3$.

Fig. 1.3 compares the continuum shapes of a Pop III.2 star ($Z = 0$) and a Pop II star with $Z = 10^{-3}Z_\odot$. To zeroth order, the spectrum of a Pop III star is a blackbody (Bromm et al., 2001). Both stars have $M = 15M_\odot$, but the Pop III.2 has an effective surface temperature of 6.3×10^4 K, compared to 3.6×10^4 K for the Pop II star (Tumlinson and Shull, 2000). The lower temperature of the Pop II star means it has a substantially lower specific radiative intensity at short wavelengths (corresponding to high photon energies) than the Pop III star, and that the mode of the distribution is at longer wavelengths (lower photon energies)[17]. Despite this, both types of star are comfortably able to produce large numbers of photons with enough energy to singly ionize both HI and HeI. Pop III stars are even hot enough to produce a non-negligible amount of photons capable of fully ionizing helium, although a $15M_\odot$ star will only have a hydrogen-burning lifetime of 1.1×10^7 years (Tumlinson et al.,

[17]The relationship between the temperature of a blackbody and the wavelength of maximum intensity can be very simply described as $\lambda_{\max} T = 2.898 \times 10^{-3}$ m K. This is known as Wien's displacement law, first published by Wien (1894). Incidentally, Wilhelm Wien received the Nobel Prize for Physics for this 100 years ago, in 1911.

2003). Their large masses and concomitant short lifetimes, combined with their chemical feedback into the IGM, puts a strong time limit on the length of time Pop III stars can be influential. As such, although Pop III stars could have started the process of H I reionization (and perhaps He II reionization), the transition to Pop II stars probably happens before the ionization process is complete (although see Cen, 2010), and thus other sources will play a more significant role.

1.2.2 Star-forming galaxies

The first galaxies can be defined as being the first objects with potential wells able to retain the gas expelled by the supernovae of the first stars (see the excellent reviews in Bromm and Yoshida, 2011; Johnson, 2011). As described in Section 1.1.1, the sites for the first galaxies are likely to be dark matter haloes of mass $\sim 10^8 \mathrm{M}_\odot$ (Wise and Abel, 2007; Greif et al., 2008; Bromm et al., 2009). The violent ends to the brief lives of Pop III stars fill the surrounding gas with metals, and within only a few generations the metallicity is greater than a critical value, $Z_{\rm crit}$ (Wise and Abel, 2008), meaning that subsequent rounds of star formation form Population II stars. These are the stars that dominate in these early galaxies by $z \sim 10$ (Johnson et al., 2008, although there may well still be rare pockets of pristine gas forming low mass Pop III.2 stars). Consequently, even though Pop III stars produce (on average) 60% more H I ionizing photons than a Pop II star (Tumlinson et al., 2003), the majority of stellar ionizing photons at these high redshifts will come from Pop II stars.

Whether or not the Pop II stars within young galaxies can reionize the Universe is entirely determined by the number of stellar ionizing photons that pervade the IGM. Following Madau et al. (1999), the critical value for the photon emission rate per unit cosmological volume, $\dot{N}_{\rm ion}$, in order to ionize the IGM by a given epoch is

$$\dot{N}_{\rm ion} = 10^{49.3} \frac{C_{\rm HII}}{3} \left(\frac{1+z}{7} \right)^3 \ \mathrm{s}^{-1}\ \mathrm{Mpc}^{-1}\,, \tag{1.11}$$

assuming WMAP-7 cosmological parameters. Here $C_{\rm HII} = \left\langle n_{\rm HII}^2 \right\rangle / \langle n_{\rm HII} \rangle^2$ is the effective clumping factor (e.g. Bolton and Haehnelt, 2007c), and is a measure of the inhomogeneity of the ionized hydrogen in the IGM. Reaching at least this number of photons using only Pop II stars requires, first, that sufficient numbers of ionizing photons are produced, and second, that a significant fraction escape into the IGM. At high redshifts, both the production rate and the escape fraction are poorly constrained observationally. Since all stars have a spectral continuum that is essentially a blackbody (see Fig. 1.3), the main sources of ionizing photons amongst the Pop II stars are the massive stars with high surface temperatures, known as O and B type stars (e.g. Miralda-Escudé and Rees, 1998, see Fig. 1.4). As with all high-mass stars, they have short lives and so these stars are only found in areas that have undergone recent star formation. Young, star-forming galaxies are therefore widely believed to be responsible for the reionization of hydrogen (e.g. Haiman and Holder, 2003; Venkatesan et al., 2003; Haiman and Bryan, 2006; Fan et al., 2006a; Bolton and Haehnelt, 2007c; Shull and Venkatesan, 2008). The second part of our ignorance is parametrised as the escape fraction, $f_{\rm esc}$, the number of ionizing photons escaping into the IGM from these star-forming galaxies. This is a difficult parameter

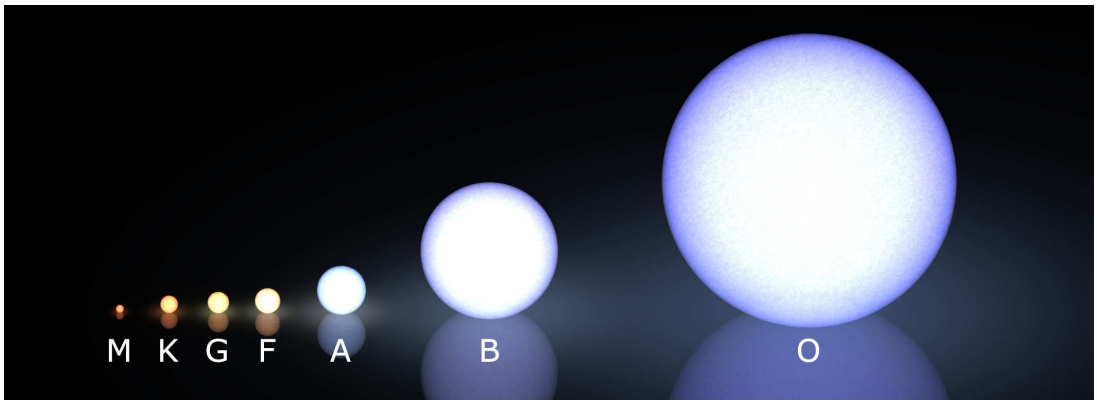


Figure 1.4: A pictorial representation of the sizes and colours of different spectral classes of stars, using the classification system of Morgan et al. (1943). The vast majority of stellar ionizing photons come from O and B type stars with surface temperatures $\gtrsim 2\times10^4$ K (masses $\gtrsim 10M_\odot$). For comparison, the Sun is a G type star. Image taken from *http://en.wikipedia.org/wiki/File:Morgan-Keenan_spectral_classification.png*

to measure even in our own galaxy (Bland-Hawthorn and Maloney, 1999), and is even more difficult in the high-redshift Universe, where there appears to be significant galaxy-to-galaxy variation (e.g. Shapley et al., 2006). Measurements at $z\sim3$ suggest $f_{\rm esc}\sim0.05-0.1$ (Boutsia et al., 2011; Nestor et al., 2011), and it is believed that $f_{\rm esc}$ should increase with redshift (although very high-redshift studies can normally only set weak upper limits, e.g. $f_{\rm esc}<0.9$ at $z\sim6.6$; Ono et al., 2010). The main theoretical motivation for $f_{\rm esc}$ increasing with redshift is that, in simulations, it seems to increase with decreasing halo mass (Wise and Cen, 2009; Razoumov and Sommer-Larsen, 2010; Srbinovsky and Wyithe, 2010; Yajima et al., 2011). Since dwarf galaxies increasingly dominate the global star formation at higher redshifts, the mean $f_{\rm esc}$ may be higher (although see Gnedin et al., 2008; Paardekooper et al., 2011).

Even with a higher $f_{\rm esc}$ at high redshift, there are still difficulties with providing enough ionizing photons. The discovery of very high-redshift galaxies in deep surveys, such as the Hubble Ultra Deep Field (HUDF), has led to discussion as to whether the population of galaxies responsible for reionization has been found (Stiavelli et al., 2004; Yan and Windhorst, 2004; Bouwens et al., 2004, 2006, 2008; Bunker et al., 2004, 2006; McLure et al., 2010, 2011; Oesch et al., 2010). These galaxies were discovered using the Lyman-break technique, whereby an object is undetected in filters above a certain wavelength. This is symptomatic of a generic feature in galaxy spectra where there is a very steep drop-off in flux at the Lyman break (in the rest frame UV). The reddest filter in which the object is not detected (the 'dropout' filter) can therefore be used to constrain its redshift. The highest redshift galaxies in the HUDF are only detected in the near-infrared, and Fig. 1.5 demonstrates the discovery of some of these galaxies at $z\sim7$ using this technique. Only the most massive and luminous galaxies, however, can be detected at $z\sim6-10$. These alone are unable to produce enough ionizing photons, even if $f_{\rm esc}=1$, and so there must be a sizable contribution from lower luminosity galaxies too faint to be seen in current surveys (e.g. Fernandez and Shull, 2011; Lorenzoni et al., 2011; Vanzella et al., 2011). The number of these galaxies is very challenging to constrain, and is observationally measured from the slope of the luminosity function. Very recent results suggest that the slope is sufficiently steep that there may be enough small galaxies to carry out HI reionization without assuming an $f_{\rm esc}$ very

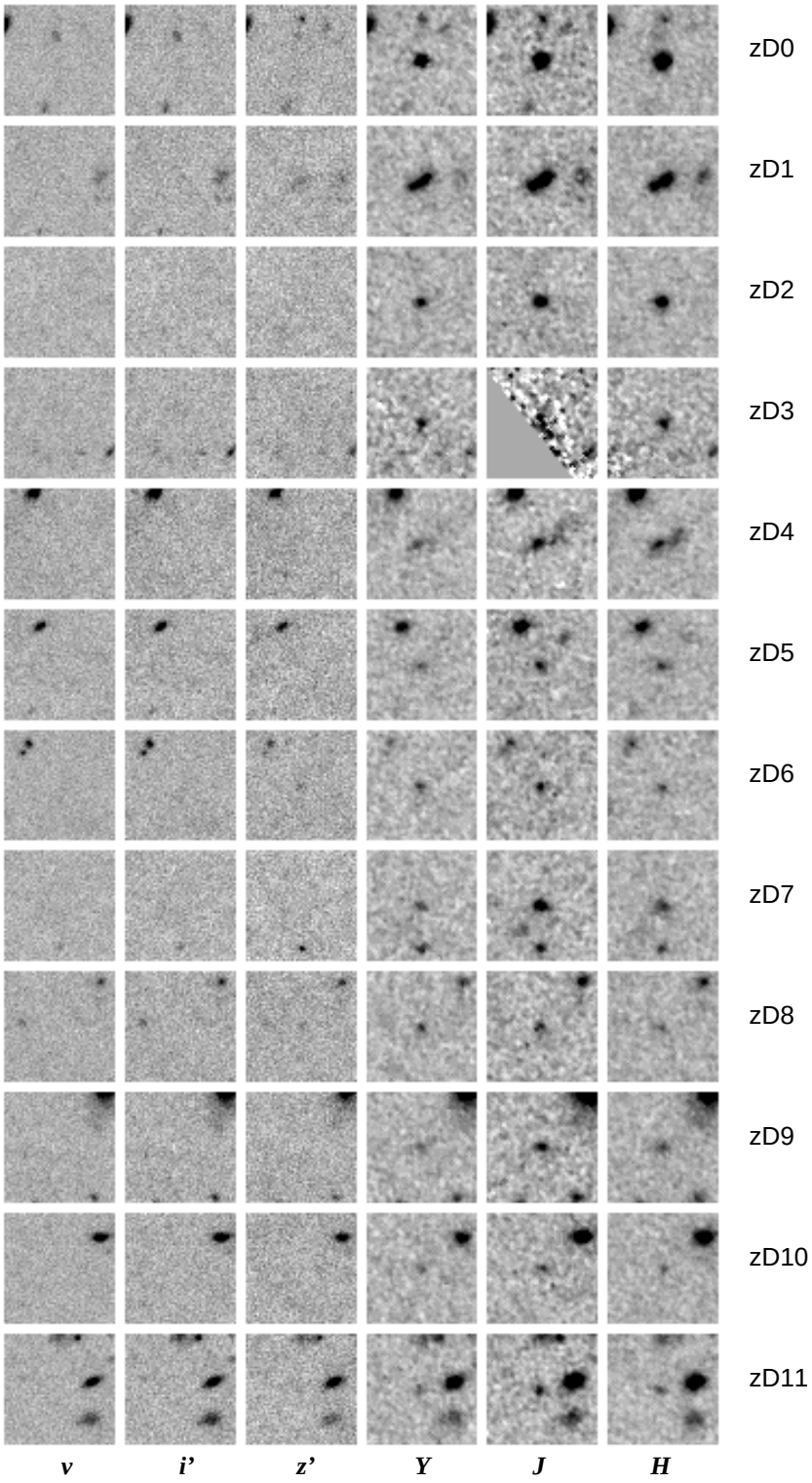


Figure 1.5: Galaxies discovered in the Hubble Ultra Deep Field using the Lyman-break technique. Each row shows the same patch of sky viewed through a variety of different filters ordered in increasing wavelength, going from the optical on the left (v) to the near-infrared on the right (H). The reddest filter the objects are not detected in is z' ($\lambda \sim 8500$ Å), so they are known as z'-dropouts. In order for the Lyman break to be redshifted to this observed wavelength, the galaxies need to be at $z \sim 7$. Figure taken from Bunker et al. (2010).

different from that measured at lower redshifts (Bouwens et al., 2011c). This is in agreement with recent simulations that show that small galaxies dominate the ionizing budget at high redshift and are able to complete reionization (Lacey et al., 2011; Raičević et al., 2011). Another important factor when considering the contribution of Pop II stars is that whilst most of their ionizing photons will be emitted during their lifetimes, there is also a non-negligible contribution from their supernovae. Accounting for this increases the stellar contribution to reionization of a star-forming galaxy by $\sim 10\%$ (Johnson and Khochfar, 2011).

Recently, Rauch et al. (2011) detected a galaxy at $z = 3.344$ that offers some insight into the physical picture of what may allow ionizing photons to escape from dwarf galaxies at high redshift. A merger between two gas rich galaxies will trigger a large burst of star formation, while at the same time providing a window through which the ionizing photons can escape, by disrupting part of the galaxy. As such, both the ionizing output (via a greater number of O and B type stars produced in the starburst) and the escape fraction of ionizing photons increases. Since mergers will have been much more common during the hierarchical growth rampant in the early Universe, such a mechanism could explain the prevalence of sufficient ionizing photons to complete reionization. There seems to have been a major period of galactic build-up at $z \sim 8-10$ (Oesch et al., 2011), and so the starbursts triggered by the associated mergers of these dwarf galaxies suggest an interesting possibility for an important source of ionizing photons responsible for the reionization of H I and He I .

1.2.3 Active Galactic Nuclei

Shortly after the Third Cambridge Catalogue of Radio Sources (3C; Edge et al., 1959) was released, several objects were shown to be of extragalactic origin. 3C 295 was found to be a radio-bright galaxy at $z = 0.46$ (Minkowski, 1960), making it the most distant object known at the time of its discovery. These objects became known as Active Galactic Nuclei (AGN), due to their radio activity. Some radio sources were also associated with what looked like stellar sources in the optical (Matthews and Sandage, 1963). When one of these, 3C 273, was found to have $z = 0.16$, it became clear that despite appearing to be a point source, it could not be a star (Schmidt, 1963). These sources were given the name 'quasi-stellar radio sources' (Greenstein and Schmidt, 1964), which was subsequently abbreviated to 'quasar'. Schmidt (1965) measured 3C 9 to have $z \sim 2$, emphasising that quasars were some of the most luminous objects in the Universe, making them visible at even very high redshifts. Concurrently, several other high redshift, point-like sources with no associated radio emission were discovered (Sandage, 1965). Both the radio-quiet and radio-loud objects were grouped together as 'Quasi-Stellar Objects' (QSOs)[18]. Investigation of the unusual spectra and their tremendous energy outputs led to the conclusion that these objects were powered by accretion onto a supermassive black hole (SMBH) (Lynden-Bell, 1969). For a review on the unification of all the different types of AGN see Antonucci (1993), and for a review on the creation of SMBHs at high redshift see Volonteri (2010).

Due to their exotic method of energy production, AGN have long been known to be excellent producers of ionizing photons (Arons and McCray, 1970; Rees and Setti, 1970). The spectra are domi-

[18] Whilst historically 'quasar' was the term used to refer to radio-loud QSOs, I will use the terms interchangeably.

nated by synchrotron emission, and thus can be described via a power-law[19], so they emit more of their energy as high-energy photons than even a very hot Pop III star (see Fig. 1.3). Similarly, the escape fraction from bright AGN, such as quasars with an intrinsically high luminosity, is thought to be close to unity (Loeb and Barkana, 2001; Fan et al., 2006a). Consequently, only a small percentage of the galaxies need to be in an AGN stage for them to dominate the ionizing budget at a given redshift. This is because the efficiency of ionizing photon production into the IGM from accretion onto a SMBH is comparable to that from the combined nucleosynthesis of stars in a galaxy (e.g. Meiksin, 2006). Having identified their potential, AGN were widely thought to be responsible for the reionization of HI (e.g. Arons and Wingert, 1972). However, the peak in the bolometric quasar luminosity function is at $z \sim 2$ and drops off sharply at higher redshifts (e.g. Hopkins et al., 2007). The luminosity function is steep at the bright end though, so faint AGN will be responsible for most of the ionizing photons. Surveys such as the Sloan Digital Sky Survey (SDSS; York et al., 2000) have sampled the luminosity function up to $z \sim 6$. Although only the bright end of the luminosity function is being sampled at these high redshifts, it seems to be getting flatter at $z > 4$ (Richards et al., 2006). This implies that there are not enough AGN at high redshift ($z > 6$) to reionize hydrogen, unless there is a substantial population of very low luminosity AGN (Meiksin, 2005; Bolton and Haehnelt, 2007c; Srbinovsky and Wyithe, 2007). In contrast, due to the high intensity of their spectra at short wavelengths[20], quasars are expected to be the dominant sources of the reionization of HeII (e.g. Ciardi and Ferrara, 2005; Bolton et al., 2006).

1.2.4 Other sources

Whilst the first stars, high-redshift star-forming galaxies and AGN represent the favoured contributors to reionization, there have been studies into more exotic sources. Among them are 'mini-quasars' powered by intermediate-mass black holes with masses $200 - 1000 M_{\odot}$ (Oh, 2001; Venkatesan et al., 2001; Madau et al., 2004), quark-novae from the collapse of neutron stars to quark stars (Ouyed et al., 2009), globular star clusters (Ricotti, 2004; Schaerer and Charbonnel, 2011), early black holes formed from the first Pop III stars (Ricotti and Ostriker, 2004b), feedback from accreting black holes in high-mass X-ray binaries (Mirabel et al., 2011), dark matter particle annihilations or decays (Scott et al., 1991; Mapelli et al., 2006), and sterile neutrinos (Hansen and Haiman, 2004). Although there is currently no observational evidence for any of these, the possibility of sources other than galaxies reionizing hydrogen (and quasars reionizing helium) should not be ruled out.

1.3 Observational constraints on the Epoch of Reionization

The detailed interplay between the different sources of ionizing photons during the EoR and up to the present day leaves observational tracers that can greatly constrain hierarchical structure forma-

[19] Synchrotron emission is caused by charged particles (in this case electrons) moving through a magnetic field at speeds just under the speed of light. The two key properties of synchrotron radiation are its non-thermal power-law spectra, and its polarization. It was first detected coming from a jet out of the galaxy M87 by Burbidge (1956).

[20] Referred to in a astronomy as 'hard' spectra.

tion models (Trac et al., 2008; Cen et al., 2009). Since much of reionization seems to have taken place at redshifts that currently suffer from poor completeness in modern surveys, information gleaned about the sources of reionization will tell us about the galactic population statistics at these unseen epochs. This will help constrain the IMF of the first galaxies, as well as the escape fraction of ionizing photons and metals into the IGM. For the rest of this chapter, 'reionization' refers to the reionization of hydrogen (in parallel with HeI). This is because the reionization of helium is far better constrained from quasar absorption spectra (see Section 1.3.2) as finishing at $z \sim 2.7-3$ (e.g. Kriss et al., 2001), via optical depth measurements (e.g Madau and Meiksin, 1994; Worseck et al., 2011), and IGM temperature measurements (e.g. Becker et al., 2011a, see Chapter 6). This coincides with the rise of hard spectra sources such as quasars, which emit by themselves ~ 2.5 ionizing photons per HeII atom at $z = 3$, an ample number to complete helium reionization (Furlanetto and Oh, 2008a).

1.3.1 Polarization of the Cosmic Microwave Background

Perhaps the best current estimate of the redshift of reionization comes from the large angular scale polarization of the CMB. The CMB photons have been travelling through the Universe since recombination and hold information about the physical conditions of the Universe at that time. Reionization produces free electrons off which these photons Thomson-scatter, creating polarization. Therefore, the CMB polarization is sensitive to the presence of HII, and so the IGM Thomson optical depth provides information on when HII regions began to dominate the CMB sightlines (although note that a small fraction of residual electrons frozen out during recombination will affect this measurement slightly; Page et al., 2007). This polarization of the CMB is detected as an excess in the signal of the temperature-polarization correlation at large angular scales, away from that predicted by the primordial power spectrum of the CMB temperature anisotropies (see Fig. 1.6 and e.g. Kogut et al., 2003, for more details). A measurement of the Thomson optical depth, τ_e, can be converted into a reionization redshift, $z_{\rm reion}$, for a given reionization history. Assuming the Universe is reionized instantaneously, such that at $z > z_{\rm reion}$ it is neutral, and $z \leq z_{\rm reion}$ it is completely ionized, then in a flat ΛCDM cosmology τ_e is related to $z_{\rm reion}$ as[21]

$$\begin{aligned}\tau_e &= \frac{H_0 \Omega_b \sigma_T c}{4\pi G m_p \Omega_m} \frac{n_p}{n_b} \left[\left(1 - \Omega_m + \Omega_m (1+z_{\rm reion})^3\right)^{1/2} - 1 \right] \\ &\approx 0.0406 \frac{\Omega_b h}{\Omega_m} \left[\left(1 - \Omega_m + \Omega_m (1+z_{\rm reion})^3\right)^{1/2} - 1 \right]\end{aligned} \tag{1.12}$$

(e.g. Griffiths et al., 1999). The current best estimates from the Wilkinson Microwave Anisotropy Probe (WMAP)[22] seven-year data give $\tau_e = 0.088 \pm 0.015$, which corresponds to a reionization redshift of $z_{\rm reion} = 10.6 \pm 1.2$ (Komatsu et al., 2011). As already mentioned, this assumes an instantaneous reion-

[21] In this equation, H_0, Ω_b and Ω_m all have the same meaning as in Table 1.1, m_p is the mass of a proton, c is the speed of light, G is Newton's gravitational constant, σ_T is the electron Thomson scattering cross-section ($= 6.65 \times 10^{-25}$ cm^2), $h = H_0/100$ km s^{-1} Mpc^{-1}, and n_p and n_b are the proton number density and baryon number density respectively. Assuming a primordial mixture of hydrogen and helium based on Big Bang nucleosynthesis (24% of all baryons in ^{4}He nuclei, the rest in ^{1}H), then 88% of all baryons are protons, and so $n_p/n_b = 0.88$.

[22] http://map.gsfc.nasa.gov/

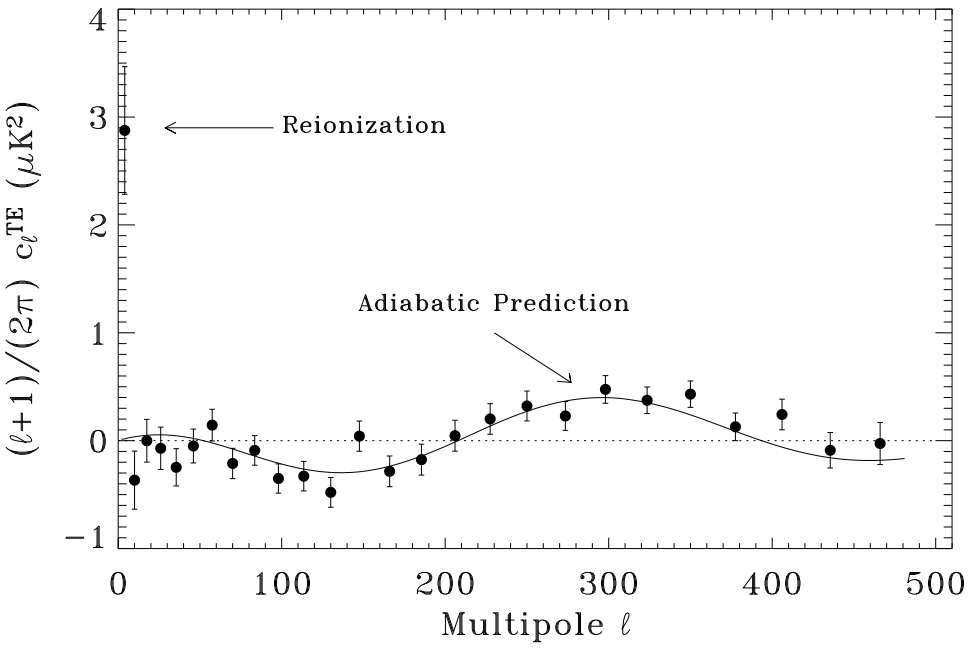


Figure 1.6: Polarization cross-power spectra for the WMAP first-year data, taken from Kogut et al. (2003). The solid line marks the prediction based upon the power spectrum of the temperature anisotropies. The dotted line represents the case if there were no temperature anisotropies. For $\ell < 10$ (corresponding to large angular scales) there is clear excess power over the predicted signal, indicating reionization.

ization event, which is unphysical (21 cm emission (see Section 1.3.4) has already constrained the duration of reionization to be $\Delta z > 0.06$; Bowman and Rogers, 2010), but provides a first approximate measurement of the redshift of reionization. Further refinements to this constraint will come in early 2013 from the data provided by the Planck[23] satellite, successfully launched in 2009.

1.3.2 High-redshift spectra

Quasar absorption spectra

Arguably the best data about the end of reionization comes from the spectra of high-redshift quasars. One of the most widely used spectral lines, due to its relative strength, is the Lyman alpha (Lyα) line of hydrogen. It corresponds to an electron moving between the $1\,^2S$ ground state and the first excited state, $2\,^2P$,

$$1\,^2S + \mathrm{Ly}\alpha \Longleftrightarrow 2\,^2P\,. \tag{1.13}$$

This transition is in the ultraviolet part of the spectrum, with a wavelength of 1216 Å. Gunn and Peterson (1965) calculated the optical depth, $\tau_{\mathrm{Ly}\alpha}$, of the IGM to Lyα photons as a function of the neutral frac-

[23] http://www.rssd.esa.int/index.php?project=Planck

tion of hydrogen, $x_{\rm HI} = (n_{\rm HI})/(n_{\rm H})$, as[24]

$$\tau_{{\rm Ly}\alpha} = \frac{\pi e^2 f_\alpha \lambda_\alpha n_{\rm HI}(z)}{m_{\rm e} c H(z)} \approx 3.46 \times 10^5 x_{\rm HI} \left(\frac{\Omega_{\rm b} h}{0.0333}\right) \left(\frac{\Omega_{\rm m}}{0.26}\right)^{-1/2} \left(\frac{1+z}{7}\right)^{3/2} . \quad (1.14)$$

From this equation it was clear that that the presence of neutral hydrogen between the source and the observer will cause a dip in the spectrum blueward of Lyα (a 'trough'). Significantly, however, only a very small neutral fraction (about one atom of H I per 10^4 atoms of H II) is required to saturate this area of the spectrum, forming a 'complete' trough. It was realised early on from observations of quasars at $z < 2$ that the lack of a complete Gunn-Peterson (GP) trough meant the low-redshift Universe was highly ionized (e.g. Arons and Wingert, 1972). By measuring the effective optical depth, $\tau_{\rm eff} \equiv -\ln\langle F\rangle$, where $\langle F\rangle$ is the mean normalised flux in the trough, one can then use Equation 1.14 to convert it into a neutral fraction. Further observations reaching $z \sim 6$ similarly failed to find complete GP troughs, implying an upper limit for $x_{\rm HI}$ of 10^{-4} (Fan et al., 2006a). This demonstrates that by $z \sim 6$ the Universe was essentially fully ionized.

Measurements using $\tau_{\rm eff}$ have historically been the most common method for calculating the neutral fraction at high redshift (e.g. Fan et al., 2006b; Bolton and Haehnelt, 2007c; Goto et al., 2011). The first complete GP trough was found by Becker et al. (2001) in the quasar SDSS J1030+0524 ($z = 6.28$). This implied a lower limit on $x_{\rm HI}$ of 10^{-3}, although measurements of the GP trough for the Lyβ transition suggested $x_{\rm HI}$ was perhaps as high as 0.1. Modelling the relationship between $\tau_{\rm eff}$ and z as a power law, Fan et al. (2006b) found that they required a separate power law for the data at $z < 5.7$ and $z > 5.7$. This has been taken as evidence that at $z \sim 6$ the IGM is within the overlap phase close to the end of reionization, where the mean free path of the ionizing photons increases very rapidly. This rapid rise in the optical depth of the spectra shortwards of Lyα is clear in Fig 1.7, taken from Fan et al. (2006c). It is worth noting, however, that the bright quasar SDSS J1148+5251 ($z = 6.42$) does *not* have a complete GP trough in either the Lyα or Lyβ transitions, indicating that this sightline is still highly ionized (Oh and Furlanetto, 2005; White et al., 2005). Significant sightline-to-sightline variation is expected from a clumpy IGM (Furlanetto and Oh, 2005), although this variation is not necessarily an indicator of an inhomogeneous reionization (Lidz et al., 2006b). The redshift of the completion of the overlap phase has been suggested to vary amongst different sightlines by $\delta z = 0.15$ (Wyithe and Loeb, 2004). Also, the significance of the presence of complete GP troughs has been questioned. Becker et al. (2007) suggested that by modelling the evolution of the transmitted flux using a lognormal distribution of optical depths (rather than the canonical density distribution of Miralda-Escudé et al., 2000) no break in the relationship between $\tau_{\rm eff}$ and z is found in the measurements of Fan et al. (2006c), implying overlap is at $z > 6.4$. Similarly, McQuinn et al. (2011) show that the properties of Lyman limit systems can explain the rapid evolution in the transmitted flux without invoking reionization.

[24] In this equation, e is the electric charge on an electron, $m_{\rm e}$ is the electron mass, c is the speed of light, f_α is the oscillator strength (= 0.4162), λ_α is the wavelength of the transition (= 1216 Å), and $n_{\rm HI}$ is the number density of atomic neutral hydrogen. $H(z)$ is the Hubble constant at redshift z, and in a flat ΛCDM Universe $H(z) = H_0\sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda} \approx H_0 \Omega_{\rm m}^{1/2}(1+z)^{3/2}$ at high redshift, as implicitly used here.

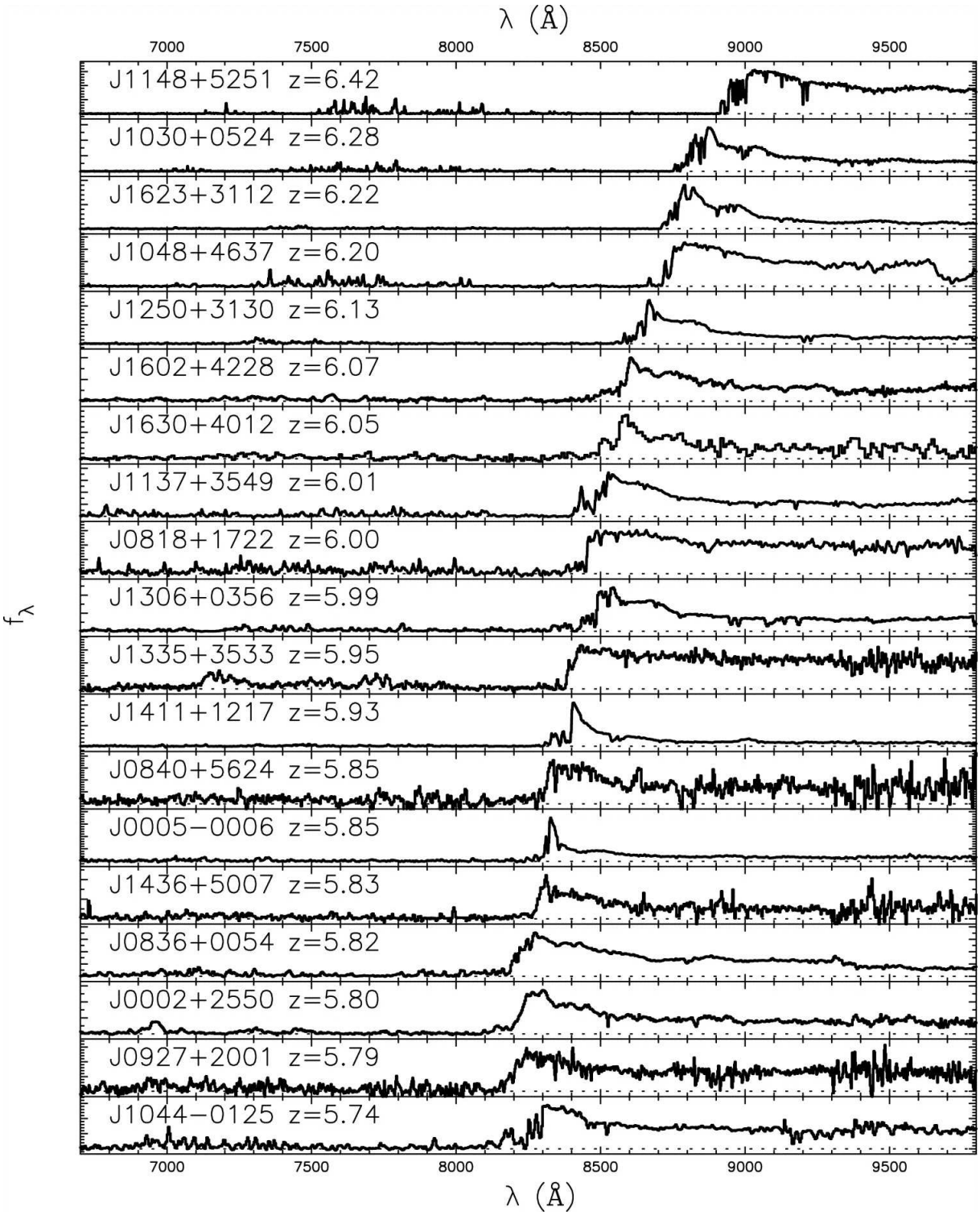


Figure 1.7: A reproduction of Fig 1. in Fan et al. (2006c) showing the 19 spectra in their sample. There is a clear reduction in transmitted flux in the Lyα forest towards higher redshift.

Other statistics have been used with quasar sightlines to help constrain the neutral fraction. The size evolution of the H II bubbles around quasars was used by Carilli et al. (2010) to suggest that the neutral fraction has increased by an order of magnitude from $z = 5.7$ to $z = 6.4$. Small sections of transmitted flux, similar in appearance to emission lines, found between long stretches of zero-flux absorption gaps (essentially GP troughs) are known as Lyα 'leaks' (Liu et al., 2007; Feng et al., 2008). They could be due to ionized gas in a largely neutral medium, or to a very low density void in the sightline. Determining their origin will therefore hold useful information about the neutral fraction at high redshift. Initial results suggest they are due to very low density voids, consistent with a highly ionized IGM at $z \sim 6$ (Liu et al., 2007). Gallerani et al. (2008a) used the statistics of the prevalence of dark gaps and transmission peaks in the spectra to get a robust upper limit of $x_{\rm HI} < 0.36$ at $z = 6.3$. Very recently, McGreer et al. (2011) claimed to have a nearly model independent estimate of the neutral fraction based on the covering fraction of dark pixels, with their most stringent constraint of $x_{\rm HI} \lesssim 0.5$ at $z = 6.1$.

Since these techniques all provide mainly upper limits, it is difficult to say much that is conclusive about the end of reionization. Recently, the reliability of $x_{\rm HI}$ measurements from quasar spectra has been questioned, since quasars sit within very biased regions of the IGM (see Chapter 5 for more details). The reionization history of the IGM around a quasar is expected to precede that of the general IGM by ~ 0.3 redshift units (Wyithe et al., 2008), and so some have argued that since much of our information on the end of reionization comes from these highly biased sightlines, it is hard to say with any certainty that reionization was even complete by $z \sim 5-6$ (Mesinger, 2010). Despite this, $z \sim 6$ is often quoted as a strict lower limit by which reionization must have ended, but the possibility remains that it ended at a higher (or lower) redshift.

As a final note, the very recent discovery of a bright quasar at $z = 7.085$ (Mortlock et al., 2011) provides a useful window into the very high-redshift Universe. The spectrum around the systemic Lyα redshift seems to exhibit evidence for a damping wing, suggesting a lower limit on the neutral fraction of $x_{\rm HI} > 0.1$ (Bolton et al., 2011; Mortlock et al., 2011). The feature could also be reproduced if a Lyman limit system lay within 5 proper Mpc of the quasar, however Bolton et al. (2011) found that this only happens in 5% of their simulated sightlines. This may be therefore be taken as tentative evidence that (at least along that particular sightline) reionization ended at $z \lesssim 7$.

Gamma-ray bursts

Gamma-ray bursts (GRBs) are brief ($\sim$ few second) flashes of very high energy emission that have an observed frequency of a few per day. Their existence was initially speculated on by Colgate (1968), although the first were only found by Klebesadel et al. (1973). Due to their brevity, it has been difficult to establish the nature of the sources, although their isotropic distribution on the sky, combined with the first confirmed host galaxy, showed that they were of extragalactic origin (Metzger et al., 1997). Some GRBs have been so luminous that they would still be detectable even at $z \sim 10$ (Kulkarni et al., 1998, 1999). Until recently, the most distant GRBs had redshifts that did not exceed those of other cosmological sources, such as $z = 6.3$ (Kawai et al., 2006) and $z = 6.7$ (Greiner et al., 2009). The GRBs

detected at redshift 8.3 (GRB 090423; Salvaterra et al., 2009; Tanvir et al., 2009) and possibly at 9.4 (GRB 090429B; Cucchiara et al., 2011) in 2009, however, make them some of the most distant objects known. For a brief amount of time, they are sufficiently bright to be detected over vast distances, and so their spectra provide a very useful probe of the early Universe (Miralda-Escudé, 1998). They are thought to be associated with the deaths of massive stars (Woosley and Bloom, 2006; Modjaz, 2011; Petitjean and Vergani, 2011), so their detection serves as a signpost of star formation in galaxies that are too faint to be imaged directly (Jakobsson et al., 2006). They have two key advantages over high-redshift quasars. The host galaxies seem to not have extreme properties (Chen et al., 2009; Berger, 2009), and presumably sit in a more 'average' part of the IGM. Secondly, unlike quasars, they have a very simple power law continuum, making continuum placement easier. They have been used successfully at lower redshift (e.g. Thöne et al., 2011) to measure directly the gas in the vicinity of the GRB, and provide clues about the host galaxy and its stellar component, with the hope to do similar analysis at $z > 6$. Totani et al. (2006) measured the shape of the damping wing for GRB 050904 at $z = 6.3$, and found that its shape could be entirely explained by a Damped Lyα (DLA) system at $z = 6.295$ indicated by several metal lines, without the need for any contribution from the IGM. The most probable neutral fraction was $x_{\rm HI} = 0.00$, with upper limits of < 0.17 and < 0.60 at 68% and 95% confidence levels, respectively. This suggests a highly ionized IGM, especially when taking into account the fact that constraints on the neutral fraction based on damping wings may be overestimates (Mesinger and Furlanetto, 2008a). By looking at the width of the GP trough in the same GRB spectrum, Gallerani et al. (2008b) concluded that $x_{\rm HI} = 6.4 \pm 0.3 \times 10^{-5}$ along that particular sightline, although McQuinn et al. (2008) claim a single GRB $x_{\rm HI}$ measurement can only constrain the global neutral fraction to $\Delta x_{\rm HI} \sim 0.3$.

The usefulness of GRBs is not limited to their spectra. Number counts of high-redshift GRBs were used by Wyithe et al. (2010) as an indicator of the star formation rate density, and so to derive an escape fraction of ~ 0.05 from their host galaxies. It has also been suggested that GRBs can be linked to the death of Pop III stars (Bromm and Loeb, 2007; de Souza et al., 2011), perhaps giving valuable insight into the stellar IMF at high redshift. Regardless, the massive progenitors of GRBs are also likely to be important sources of ionizing photons (see Section 1.2.1 and Section 1.2.2) and perhaps provide a significant contribution to reionization at $z > 6$ (Bolton and Haehnelt, 2007c; Choudhury et al., 2008; Furlanetto and Mesinger, 2009; Stiavelli et al., 2009).

1.3.3 Lyman alpha emitters

Lyman alpha emitters (LAEs) are galaxies found via their strong Lyα emission lines. This emission is powered by the reprocessing of ionizing photons within the interstellar medium (ISM) of the galaxy (Partridge and Peebles, 1967). They are normally found in surveys using narrowband filters, which are tuned to coincide with a gap in the atmospheric emission spectrum. For very high redshift work, the commonly-used gaps correspond to $z = 5.7, 6.5-6.6$ and 7.7, so LAEs detected in these epochs bracket the possible end of the EoR, providing a probe of this interesting time in the Universe's history.

If the Universe is highly neutral at a given redshift then the visibility of Lyα from galaxies should

be substantially suppressed, by at least a factor of 3 (Malhotra and Rhoads, 2004). This effect should manifest itself in the luminosity function of LAEs. Several groups (e.g. Malhotra and Rhoads, 2004; Kashikawa et al., 2006; Dijkstra et al., 2007; Ouchi et al., 2010; Kashikawa et al., 2011; Nakamura et al., 2011) have measured the luminosity functions of LAEs at $z = 5.7$ and $z = 6.5 - 6.6$. They have consistently found that there are no dramatic differences between the two, arguing against a fully neutral IGM at $z = 6.6$. Very recently, Clément et al. (2011) failed to detect any LAEs at $z = 7.7$, and were able to use this non-detection to place robust constraints on the bright end of the luminosity function. They then compared it to those at $z = 6.5$ from the literature and found that they could rule out a non-evolution scenario at 85% confidence, with the data preferring a strong increase in the neutral fraction. Although the statistical robustness of their conclusion was dependent on which $z = 6.5$ luminosity function they used from the literature, this is tentative evidence that there may be a substantial change in the neutral fraction between $z = 6.5$ and $z = 7.7$ (although see Tilvi et al., 2010; Krug et al., 2011). Lehnert et al. (2010) did manage to detect an LAE at $z = 8.6$, although the strong Lyα emission is not enough in itself to suggest a highly ionized IGM or put statistically meaningful constraints on the neutral fraction (Dijkstra et al., 2011). Similarly, Mesinger and Furlanetto (2008a) use the six potential LAEs found at $z \sim 9$ in Stark et al. (2007) to constrain the luminosity function evolution and derive that $x_{\rm HI} \lesssim 0.7$, providing one of the highest redshift limits on the neutral fraction.

The clustering of LAEs can also be used to constrain the neutral fraction. A substantially neutral medium would enhance the apparent clustering of LAEs, as the most detectable would be those in galaxy clusters within a large H II bubble. Several groups (e.g. McQuinn et al., 2007; Ouchi et al., 2010) have used this to complement measurements from the luminosity function, and similarly conclude that the data implies a highly ionized IGM at $z = 6.5$. Finally, Ouchi et al. (2009) note that the distribution of their LAEs implies an ionized bubble of radius ~ 30 Mpc, and so using the analytical models of Furlanetto et al. (2006b) derive that $x_{\rm HI} \lesssim 0.2$ at $z = 7$. This is one of the most stringent constraints on the high-redshift neutral fraction.

1.3.4 Future 21 cm observations

The 21 cm line[25] (corresponding to a frequency of 1.4 GHz) is the radio emission due to the ground state hyperfine transition[26] of H I (for two very recent and detailed reviews see Furlanetto et al., 2006a; Morales and Wyithe, 2010). Its potential for measuring hydrogen in the IGM was first noted by Field (1959). The major observable of interest is the difference in the brightness temperature[27] of high-redshift patches of H I to the brightness temperature of the CMB. A cloud of H I may appear in either absorption or emission against the CMB, depending on the difference between the CMB temperature and the gas spin temperature. The spin temperature, $T_{\rm S}$, is linked to the ratio of the number of atoms in the excited hyperfine state, n_1, to the number in the ground hyperfine state, n_0, via Boltzmann

[25] This of course is in the frame of reference of the source - the observed wavelength will be $21(1+z)$ cm.

[26] It is caused by the electron flipping its spin relative to the nucleus.

[27] The 21 cm brightness temperature, $T_{\rm b}$, is defined as the temperature of a blackbody that produces the same intensity at 1.4 GHz as the measured 1.4 GHz intensity. Since the CMB has a blackbody spectrum, its brightness temperature is the same as its temperature.

statistics such that $n_1/n_0 \propto \exp(E_{10}/k_B T_S)$, where $E_{10} = 5.9 \times 10^{-6}$ eV is the energy between the two states, and k_B is the Boltzmann constant. The difference in brightness temperature between a sightline intersecting a HI cloud and a (hypothetical) sightline with a unobstructed view of the CMB can then be calculated as

$$\delta T_b \approx 27 x_{HI}(1+\delta)\left(\frac{1+z}{10}\right)^{1/2} \frac{T_S - T_{CMB}}{T_S} \text{ mK} \quad (1.15)$$

(e.g. Mellema et al., 2006), where $(1+\delta)$ is the fractional baryon overdensity. The value of T_S is determined by absorption of CMB photons, collisions with other particles, and scattering of UV photons, and its difference from the CMB temperature determines whether the 21 cm signal is observed in emission or absorption. Since δT_b is proportional to the abundance of neutral hydrogen, it is very difficult to measure in the highly ionized IGM of the low-redshift Universe (e.g. Ghosh et al., 2011). During and before reionization, however, there should be sufficient HI to produce a detectable signal. Indeed, during reionization T_S will roughly equal the kinetic gas temperature, so $T_S \gg T_{CMB}$, leaving δT_b essentially determined by the overdensity and the neutral fraction. Consequently, measuring the power spectrum of the HI 21 cm fluctuations is one of the more exciting techniques for investigating the EoR and distinguishing between different reionization histories (Carilli et al., 2004; Barkana and Loeb, 2005; Furlanetto, 2006). Perhaps the most interesting use is that spatially resolved (2-D) maps of a patch of sky over a range of frequencies (corresponding to different redshifts) offer the possibility of direct 3-D tomography of the high-redshift IGM (see Fig. 1.8, and e.g. Carilli, 2006; Geil and Wyithe, 2008). This will provide a test of the current view of reionization, where sources of ionizing photons carve out an ionized region of HII around themselves in the neutral IGM. These ionized 'bubbles' will ultimately merge in the percolation phase, and the consequent size evolution of these regions at different redshifts should be visible (Madau et al., 1997; Wyithe et al., 2005; Greif et al., 2009). Until recently, radio telescopes have only been sensitive enough to put upper limits on the amplitude of the 21 cm brightness temperature fluctuations (e.g. $\delta T_b < 70$ mK at $z \sim 9$; Paciga et al., 2011). However, radio telescopes currently being built, such as the Low Frequency Array (LOFAR)[28], will be able to measure the power spectrum of δT_b of any neutral hydrogen in the IGM directly. LOFAR can observe up to $z \sim 11.5$ at 115 MHz and so is expected to be able to probe the EoR. Future radio telescopes (e.g. the Square Kilometer Array[29]) should be sensitive at even higher redshifts and with adequate foreground subtraction may allow us to probe directly the neutral hydrogen in the cosmic dark ages at $z \sim 20$. Indeed, the statistics of 21 cm fluctuations offer us the opportunity to probe the pre-reionization IGM, the whole process of HII region overlap, and even the appearance of the diffuse IGM post-overlap (Wyithe and Loeb, 2008), making it a very exciting future source of observational constraints on reionization.

1.3.5 Summary

Due to the large error bars on z_{reion} from the CMB measurements, plus the relatively loose limits from other observations, several contrasting reionization histories are allowed. A selection are summarised

[28] http://www.lofar.org/
[29] http://www.skatelescope.org/

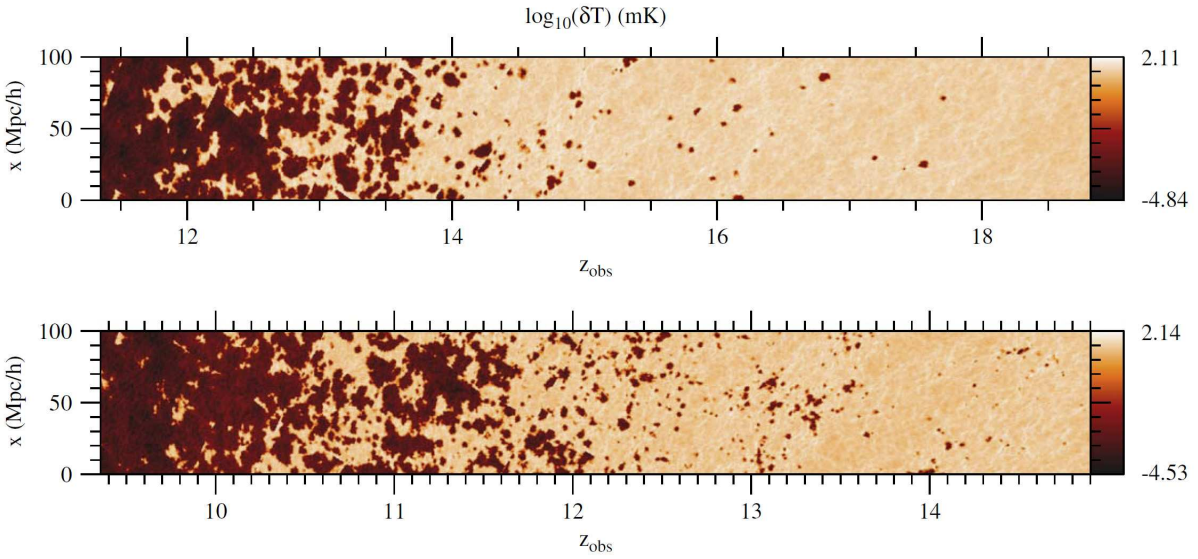


Figure 1.8: Slices through a simulation showing the evolving $\delta T_{\rm b}$ measured in an image-frequency volume for two different reionization histories. They illustrate the large-scale geometry of reionization seen at 21 cm and the significant local variations in reionization history. Figure taken from Mellema et al. (2006).

in Fig. 1.9, along with several measurements of the neutral fraction from the recent literature. Essentially all of our information about the end of reionization comes from quasar sightlines, whilst our highest-redshift constraints are from LAEs and the CMB.

Reionization histories can be classified according to when much of the ionization happens. If the majority happens at high redshifts then they are known as 'early', whilst if most happens at $z \sim 6-7$ they are known as 'late'. The reionization history of Choudhury et al. (2008) for an ionizing photon budget dominated by large haloes ($M(z=6) \gtrsim 10^9 M_\odot$; black dotted line in Fig. 1.9) is an example of a 'late' reionization history. This is due to the formation of such haloes happening rather late in the Universe's history. This model matches the observational constraints well at low redshift, but not at high redshift, implying that at those epochs the contribution to the ionizing budget from lower mass minihaloes (black solid line) is substantial. Haardt and Madau (2011) suggest a history based upon their model of the UV background (see Section 1.4) with overlap completing at $z = 6.7$, whilst Pritchard et al. (2010) find model-independent conclusions that reionization was mostly complete by $z \sim 8$. The models of Furlanetto (2006) are both 'early' reionization histories. They include only the contributions from Pop II and Pop III stars for both the case of no feedback (such as winds, ionizing photons and metals) from the stars into the IGM (green solid line), and for a limiting case of extremely strong feedback (green dashed line). In this latter scenario, between $z \sim 7-10$ the Universe starts to recombine and $x_{\rm HI}$ increases, before again decreasing. This type of reionization history is known as 'double', as there are two periods of reionization - one to start the process, and one to complete it.

The physics of what could cause a double reionization are intricately linked to the star-formation rate and its response to feedback. As described in Section 1.2.1, the ability to form a star from a cloud of gas is critically dependent on the gas' ability to cool, and the minimum temperature it reaches has implications for the stellar mass. Initially, the primordial gas cooling was limited to only a small frac-

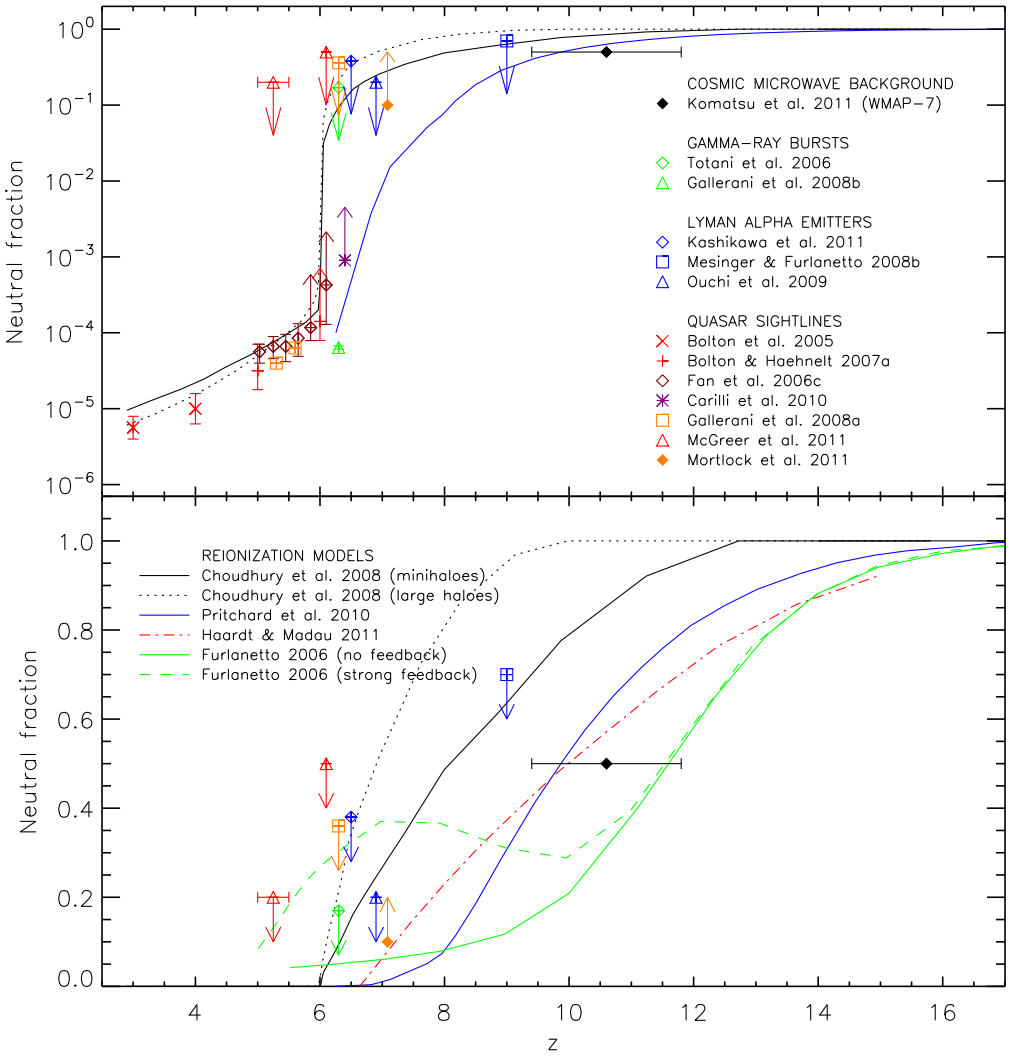


Figure 1.9: Summary of the observational constraints on the volume-averaged neutral fraction, as well as some recent model reionization histories. *Top panel:* Constraints from measurements of the CMB (Komatsu et al., 2011), GRBs (Totani et al., 2006; Gallerani et al., 2008b), LAEs (Kashikawa et al., 2011; Mesinger and Furlanetto, 2008b; Ouchi et al., 2009), and from quasar sightlines (Bolton et al., 2005; Bolton and Haehnelt, 2007c; Fan et al., 2006b; Carilli et al., 2010; Gallerani et al., 2008a; McGreer et al., 2011; Mortlock et al., 2011). Also shown are the reionization histories from Choudhury et al. (2008) when the ionizing photon budget is dominated by 'mini-haloes' ($M(z=6) \gtrsim 10^6 M_\odot$; black solid line) and by large haloes ($M(z=6) \gtrsim 10^9 M_\odot$; black dotted line), and the median reionization history of Pritchard et al. (2010) (blue solid line), parametrising over the observed ionizing efficiency ζ based on Lyα forest and WMAP-5 data. *Bottom panel:* Same as the top panel but with a linear y-axis and only displaying limits on the neutral fraction greater than 0.01. Additional to the top panel are the reionization histories of Haardt and Madau (2011) (red dash-dotted line) and two from Furlanetto (2006) (green solid and dashed lines). The former of those assumes no feedback from the ionizing sources on their surroundings and is a very extended process, failing to have finished by $z = 6$, whilst the latter assumes extreme feedback for a limiting case. In that situation the transition from the highly efficient ionizing Pop III stars to the rather less efficient Pop II stars means that the IGM starts to recombine before the Pop II stars eventually become numerous enough to complete reionization. This is the mechanism that leads to a 'double' reionization history.

tion of H_2 molecules, however in the relic H II regions of the first massive stars, sufficient numbers of H_2 and HD molecules should form and allow the gas to cool down to T_{CMB}. These first stars also polluted the IGM with metals, providing other efficient cooling pathways. So, on the one hand, the first, massive stars enable a second generation of smaller, cooler stars to form. On the other hand, the ionizing photons they produce dissociate H_2 and disrupt metal cooling (Aykutalp and Spaans, 2011), and the fast supernovae winds generated at their death heat and disrupt collapsing clumps of gas. If this feedback mechanism is fully efficient, then the result is that during the short lives of the first generation of Pop III stars, further star formation is suppressed, and after their death they are replaced with much smaller, cooler Pop II stars. The cooler Pop II stars cannot (initially) replace the ionizing emissivity of the Pop III stars, so the Universe begins to recombine, until the Pop II stars are sufficiently numerous to make up the shortfall. This negative feedback mechanism (Cen, 2003b; Sokasian et al., 2004; Ricotti and Ostriker, 2004a) would allow for the 'double' model as an example of one of the more complicated possible reionization histories. Even multiple epochs of ionization and neutralization have been proposed (Wyithe and Loeb, 2003b). However, it requires far fewer photons to keep a region ionized than it does to ionize it initially. Detailed numerical simulations have shown that the ionizing emissivity within the region would have to drop more than is plausible based on reasonable assumptions about the transition from Pop III to Pop II, making a nonmonotonic reionization highly unlikely (Furlanetto and Loeb, 2005; Furlanetto et al., 2008).

Although probably not strong enough to induce a global period of recombination within reionization, this negative feedback mechanism would work to extend the duration of reionization. High-speed supernova driven winds within high-redshift dwarf galaxies are unlikely to have much of an effect on the inflow of cold gas needed for star formation (Powell et al., 2011). The UV photons filling the H II regions created by the first quasars, however, would likely suppress the star formation in any enclosed galaxies (Bruns et al., 2011). Also, the non-zero relative velocity between the dark matter and baryonic fluid after recombination means that the formation of minihaloes (and hence star formation and local reionization history) will be delayed in some regions with respect to others (Tseliakhovich et al., 2011; Maio et al., 2011), resulting in patchy reionization (although see Stacy et al., 2011). The general consensus is that current models favour a gradual reionization over an extended period $6 < z < 14$ (e.g. Cen, 2003a; Wyithe and Loeb, 2003a; Haiman and Holder, 2003; Somerville et al., 2003; Gnedin, 2004; Furlanetto et al., 2004a; Choudhury and Ferrara, 2005; Fan et al., 2006a,c). Without better neutral fraction constraints at high redshift, however, current data does not firmly rule out other models. Additional limits may come from the local Universe (e.g. from the luminosity function of the satellite galaxies of the Milky Way; Lunnan et al., 2011), but the most revealing measurements will be made with future surveys. The James Webb Space Telescope[30] (JWST) is scheduled for launch in 2018 and will be able to detect galaxies out to $z \sim 14$ (Lacey et al., 2011; Wyithe and Loeb, 2011). With this capability, combined with potential first detection of 21 cm radiation during the EoR, this coming decade should see a substantial improvement in the observational constraints on the Universe's reionization history.

[30] http://www.jwst.nasa.gov/

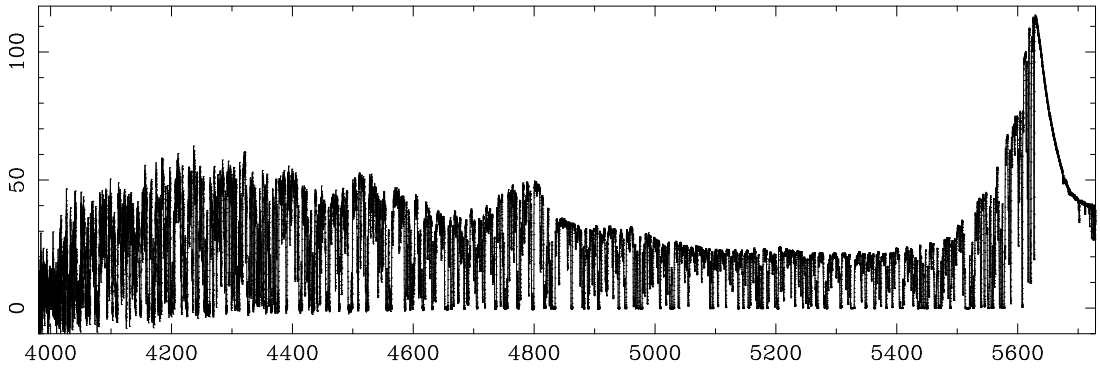


Figure 1.10: An example of a high-resolution spectrum showing the Lyα forest. This is for the quasar QSO1422+23 ($z_{\rm quasar} = 3.62$) and was taken with the HIRES instrument on Keck. In this example the flux shows a clear power law bluewards of Lyα up to the Lyman Limit at $\lambda \sim 4200$ Å. Figure taken from Rauch (1998).

1.4 The metagalactic UV background

Upon the completion of the 'overlap' stage of reionization, the Universe is essentially transparent to ionizing photons. The mean free path of these photons is now only limited by the few remaining small dense clumps of neutral gas, rather than the size of their source's H II region. As a consequence, the mean free path is now much longer than the separation between sources, and so ionizing photons pervade the Universe. This net background ionizing radiation field is referred to as the 'metagalactic' ultraviolet background (UVB), and is responsible for keeping the Universe reionized from the end of the EoR to the present day. The relative contributions from different ionizing sources, as well as filtering by the IGM itself (Bechtold et al., 1987), determine both the intensity and spectral shape of the UVB (e.g. Haardt and Madau, 1996). Thus by measuring the redshift evolution of the UVB, it is possible to constrain the redshift evolution of the source population (Haardt and Madau, 1996; Fardal et al., 1998; Haardt and Madau, 2001).

1.4.1 Observational constraints on the UVB

The decline of the quasar luminosity function at $z \gtrsim 3$ implies that star-forming galaxies are likely to be the dominant sources of the UVB from the end of overlap to $z \sim 3$, with quasars dominant at $z < 3$ (Bolton et al., 2006; Faucher-Giguère et al., 2008a). The precise contributions of quasars and galaxies at each redshift are not well understood, however. Uncertainty in the escape fraction of ionizing photons from galaxies makes it difficult to estimate the UVB by conducting a 'census' of star-forming galaxies, and so other methods are needed.

The majority of observational constraints on the UVB come from quasar absorption spectra. Blueward of the Lyα emission line, quasar spectra become overwhelmed with absorption lines in comparison to the area redward (see Fig 1.10). This was first seen in the high-resolution spectrum of a $z \sim 2$ quasar[31] by Lynds (1971). At the time there was much debate as to the source of these unidentified

[31] This is because the Lyα emission line (and that blueward of it) is in the rest-frame UV and so is only redshifted into

lines. Based on the work by Gunn and Peterson (1965) with the GP trough, Bahcall and Salpeter (1965) predicted the existence of narrow lines associated with Lyα absorption between the source and the Earth. This was the favoured explanation of Lynds (1971), and was soon the most accepted model (e.g. Arons, 1972). This feature in the spectrum of high-redshift sources is known as the 'Lyα forest' and is caused by absorption and scattering of Lyα photons by neutral hydrogen in the IGM along the line of sight to the quasar (e.g. Sargent et al., 1980; Weymann et al., 1981; Bi, 1993; Cen et al., 1994; Rauch, 1998).

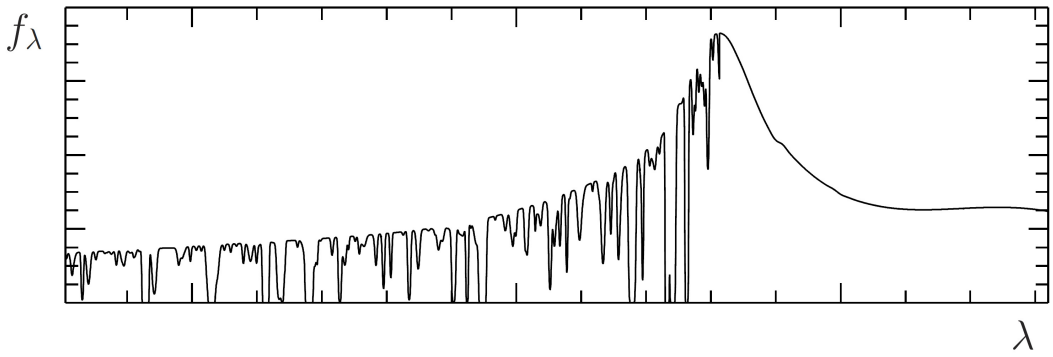
Absorption systems in the Lyα forest can be broadly classified into three categories, based upon their neutral hydrogen column density, $N_{\rm HI}$. The most common by far are Lyα forest systems, with $N_{\rm HI} < 1.6 \times 10^{17}$ cm^{-2}. They are optically thin ($\tau < 1$) to ionizing photons, meaning they are almost completely ionized, and are well fitted by Doppler line profiles. At higher column densities are Lyman limit systems (LLSs; Tytler, 1982), with $1.6 \times 10^{17} \leq N_{\rm HI} < 2 \times 10^{20}$ cm^{-2}. These are optically thick ($\tau > 1$) to hydrogen-ionizing photons. Despite their optical thickness though, LLSs are sufficiently penetrated by the UVB to be partially ionized. The final class contains the Damped Lyα systems (DLAs; Wolfe et al., 1986), with $N_{\rm HI} \geq 2 \times 10^{20}$ cm^{-2}. These are sufficiently optically thick that they show the radiation damping wings of the Lorentz profile, and so are well fit by Voigt line profiles. Due to their high optical depth, they are self-shielded from ionizing radiation, and so are essentially neutral.

The properties of the Lyα forest provide a useful probe of the large scale structure of the Universe at overdensities far lower than that probed by galaxies, and so can be used to derive several cosmological parameters (e.g. Viel, 2009). A key feature of the forest is that it becomes more opaque with increasing redshift. At $z \gtrsim 4$ absorption lines are so numerous that they can no longer be individually distinguished (an effect known as 'line blanketing'), merging into large troughs of little to no flux. By $z \sim 6$, almost no flux is transmitted through the IGM bluewards of Lyα (Fan et al., 2001; Becker et al., 2001). The number of absorption lines, and hence relative blackness of the forest, is directly linked to the underlying density field and to the amount of neutral hydrogen in the IGM. After the end of overlap, the latter is controlled by the UVB, so the average transmission (or, equivalently, effective optical depth) of the Lyα forest at a given redshift is directly connected to the UVB intensity. Two main methods have been used to measure the UVB by taking advantage of this feature: flux-decrement modelling, and the proximity effect.

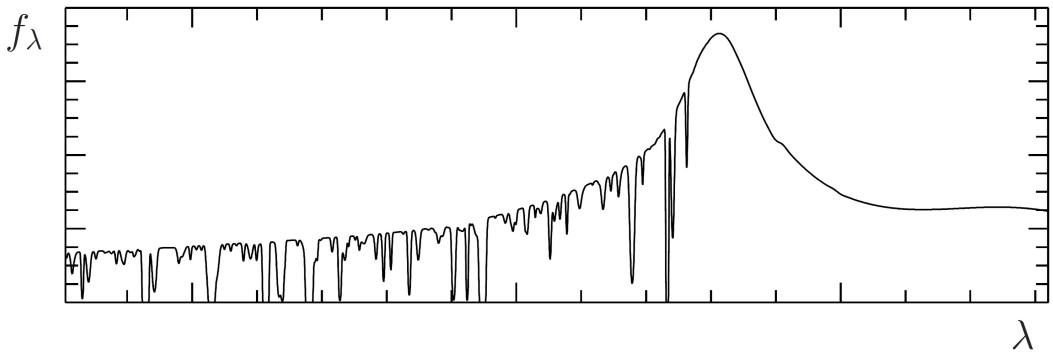
Measuring the UVB using flux decrements

The flux-decrement method consists of measuring the flux decrement, $D = \langle 1 - e^{-\tau} \rangle$, in a section of the Ly$\alpha$ forest away from the quasar redshift and comparing the results with those measured from synthetic spectra drawn from numerical simulations (Rauch et al., 1997; Songaila et al., 1999; McDonald and Miralda-Escudé, 2001; Meiksin and White, 2004; Tytler et al., 2004; Bolton et al., 2005; Kirkman et al., 2005; Jena et al., 2005). The output of the simulation is adjusted by altering the parameter $\mu \propto \Omega_{\rm b}^2 h^3 / \Gamma_{\rm bkg}$, where h is the fractional Hubble constant and $\Gamma_{\rm bkg}$ is the H I ionization rate due

the optical for $z \gtrsim 2$.



(a) Lyα forest due to only ionization by the UVB.



(b) Lyα forest due to ionization by both the UVB and the quasar.

Figure 1.11: A demonstration of the line-of-sight proximity effect in quasar spectra. *Top*: A spectrum of a quasar at $z \sim 3$ near the Lyα emission line. Blueward of the emission lies the Lyα forest. In this case, the ionization state of the IGM is set only by the UVB. *Bottom*: The same spectrum as above, but after the quasar has ionized the IGM in its vicinity. Close to the quasar the gas is more highly ionized, so there is less absorption. This is known as the proximity effect, and was first noticed in spectra by Carswell et al. (1982) as a decrease in the number of strong absorption features close to the quasar redshift. *Image credit: Lutz Wisotzki*

to the ionizing background, until the mean flux decrement agrees with the mean decrement observed in the real quasar spectrum. Assuming some known values of Ω_b and h, this method can therefore be used to derive a value for Γ_{bkg} at the redshift of the forest being sampled. This methodology is described in more detail in Chapter 3.

Measuring the UVB from the proximity effect

Quasars are highly luminous objects. Out to large distances (up to 5 proper Mpc for a bright quasar at $z \sim 2$; Kirkman and Tytler, 2008), therefore, their ionizing radiation will dominate over that of the UVB, causing the IGM in that region to be more highly ionized. Consequently, absorption lines near the redshift of the quasar will have a lower optical depth compared to lines in the forest. This is the line-of-sight proximity effect, and was first observed as a decrease in the number of absorption lines close to Lyα by Carswell et al. (1982), as demonstrated in Fig. 1.11. Subsequent studies using higher resolution spectra (Murdoch et al., 1986; Tytler, 1987; Carswell et al., 1987; Bajtlik et al., 1988; Scott et al., 2000) were able to use to the proximity effect to estimate the intensity of the UVB, using

refined versions of this 'line-counting' method. Unfortunately, the line-counting method suffers from various complicating factors. The main concern is the subjectivity with which a 'line' is identified and fitted in the spectra, a choice that is often complicated by blending, particularly at high redshift or low resolution.

An alternative approach for measuring the proximity effect is the 'flux-transmission' method. Here the observed effective optical depth close to the quasar is compared with that expected from the Lyα forest (Liske and Williger, 2001). Studies into the proximity effect in the literature are dominated by the line-counting method, however it is less sensitive to the expected proximity effect signal than the flux transmission method (Dall'Aglio et al., 2008). A variant on the flux transmission method is used, therefore, in the work presented in Chapters 4 and 5.

The transverse proximity effect

As a brief aside, it is worth pointing out that there are two proximity effects: the line-of-sight proximity effect mentioned above, and the transverse proximity effect (or foreground proximity effect). The latter of these relies on the ionizing region around the quasar being roughly spherical, and so for two quasars that are very close in angle on the sky but at different redshifts, $z_{\rm back}$ and $z_{\rm front}$, there should be an increase in Lyα forest transmission at $1216(1+z_{\rm front})$ Å in the spectra of the background quasar. Detection of such an effect would allow for a lower limit on the lifetime of a quasar given the time required for the ionizing radiation to cover the transverse distance, whilst non-detections allow for constraints on the level of anisotropy of a quasar proximity region (e.g. Lu and Yu, 2011). Many searches have been made for evidence of a transverse proximity effect without success (Liske and Williger, 2001; Schirber et al., 2004; Croft, 2004; Kirkman and Tytler, 2008). Gallerani et al. (2008a) tentatively claim to have found the first example of one by a $z = 5.65$ quasar in the spectra of a $z = 6.42$ quasar. The sightline within ~ 2 Mpc is enhanced with respect to the rest of the spectra, implying that the foreground quasar has been 'on' for $> 1.1 \times 10^7$ yrs. Despite this exception, Kirkman and Tytler (2008) see the scarcity of examples of the transverse proximity effect as evidence for higher gas densities near the quasar, an important consideration in proximity effect studies as discussed in Chapter 5. In the rest of this thesis the type of proximity effect being discussed will be the line-of-sight proximity effect.

1.5 Thesis outline

Throughout this thesis I will use high-redshift quasars as cosmic lighthouses, illuminating the IGM around and in front of them. Due to their tremendous luminosity they are bright enough to allow very high-resolution spectra to be taken, even though they are at $z > 5$. All of the work shown here will make use of these very high-resolution, high-redshift quasar absorption spectra. They will also be compared to a suite of state-of-the-art hydrodynamical simulations, with sufficiently high mass resolution to resolve the high-redshift Lyα forest. Together these provide a powerful probe of the properties of the IGM in the high-redshift Universe. The rest of this thesis is structured as follows:

- *Chapter 2*: I discuss the instruments utilised to acquire the spectra used in this thesis, and the techniques employed to reduce them. I also cover the details of the simulations of the IGM used, and describe how their outputs were modified to produce artificial quasar proximity regions.

- *Chapter 3*: I present a new flux-decrement method for measuring the UV background from Lyα forest spectra that applies a maximum-likelihood technique to the transmitted flux distribution. I then compare its effectiveness to the traditional flux-decrement method, which simply depends on the mean flux of the Lyα forest.

- *Chapter 4*: I measure the amplitude of the UV background using the proximity effect. These are the first proximity-effect measurements at $z > 4.5$, and are compared to measurements of the UVB via flux decrements in the literature. The work in this chapter forms part of the paper Calverley et al. (2011).

- *Chapter 5*: I investigate a number of environmental and systematic biases that normally afflict proximity-effect measurements, and gauge their significance at $z > 5$. Both the effect of halo mass and prevalence of dense clumps of neutral gas, known as Lyman limit systems, are thoroughly tested, as are several other second-order effects. The work in this chapter is the second part of the paper Calverley et al. (2011).

- *Chapter 6*: I present the first measurements of the temperature of the general IGM at $z > 5$, which are based on Lyα forest spectra. Two different statistics for measuring the temperature are used: the 'curvature', which has been used successfully at lower redshifts (Becker et al., 2011a), and a new statistic called the 'peakiness'. These provide independent measurements of the temperature, and their results are compared and discussed.

- *Chapter 7*: I summarise my conclusions from the work presented in this thesis, and discuss what constraints they place on the end of hydrogen reionization.

Computer Science is no more about computers than astronomy is about telescopes.

EDSGER W. DIJKSTRA

2

Real and synthetic spectra

TELESCOPES have been the backbone of observational astronomy ever since Galileo Galilei discovered the moons of Jupiter in 1610. A century ago, the largest telescope in the world was the Leviathan of Parsonstown, built by William Parsons, 3rd Earl of Rosse, with a mirror 1.83 m (6 feet) in diameter, and based in Ireland. Today, we have telescopes that observe across all parts of the electromagnetic spectrum, from radio waves to gamma rays, including some that observe from space. Currently, the largest optical telescope in the world[1] is the Gran Telescopio Canarias[2] (GTC), based on La Palma in the Canary islands, with resolving power equivalent to a 10.4 m diameter mirror. Over the next 10-15 years this record is set to be beaten several times as we enter the generation of 'Extremely Large Telescopes', including the Giant Magellan Telescope[3] (GMT; 24.5 m), the Thirty Meter Telescope[4] (TMT; 30 m), and the European Extremely Large Telescope[5] (E-ELT; 42 m). These telescopes are all set to greatly improve our understanding of the Universe, allowing us to see both the local Universe and the high-redshift Universe in unprecedented detail.

Astronomy is unusual amongst the sciences, as it is observational, rather than experimental, in nature. It is not possible to make a star in the laboratory, or watch a collision between galaxies in its entirety (due to the long timescales). However, the theory behind a star is ultimately governed by nuclear physics, and similarly a galaxy collision is controlled by Newtonian gravity. Importantly, both of these we *do* understand through experiment. Filling in the gap between the basic, well-known

[1] This is based upon the size of the primary mirror, and ignores telescopes used in optical interferometry.
[2] http://www.gtc.iac.es/en/
[3] http://www.gmto.org/index.html
[4] http://www.tmt.org/
[5] http://www.eso.org/public/teles-instr/e-elt.html

physics and the often complex, grand-scale observations are simulations. These provide the testing ground for a wide range of astronomical theories, and are used to produce predictions with which the observations are compared.

The last fifty years have seen a remarkable increase in computing power, and thus the phenomena that can be simulated. Crucially, this allows for numerical solutions to problems for which analytical solutions either do not exist, or are non-trivial to compute. One of the most important uses of simulations has been in solving what is known as the 'N-body problem'. This was first mathematically formulated by Newton in 1687, and concerns the behaviour of N particles interacting with each other through gravity. Each particle has a position and momentum in each dimension, leading to ten independent differential equations that must be integrated. This is analytically solvable for $N = 2$, but for $N \geq 3$ it becomes much less trivial. It was shown that an analytical solution exists for $N = 3$ (Sundman, 1913), and more generally for $N \geq 3$ (Wang, 1991) in the form of a convergent power series; however, these series converge so slowly that for practical purposes one must resort to approximate solutions by numerical integration of the differential equations. The first attempt to do this was by Holmberg (1941) for $N = 37$, where each 'particle' was a lightbulb, and the gravitational force each exerted on the other was measured via the light intensity (since flux follows an inverse square law, as does gravity). After each timestep the new positions and velocities of each particle would be recalculated, advanced by another timestep, and so on. The first digital attempt at an N-body simulation was by von Hoerner (1960) for $N = 16$. In a short amount of time, the increase in computing power meant this was improved to $N = 25$ (von Hoerner, 1963), and then $N = 100$ (Aarseth, 1963).

Rapid increases in the speed of computer processors over the following decades allowed N to increase substantially, improving the resolution and fidelity of the simulations (see Bertschinger, 1998, for a detailed review of the first four decades of N-body simulations). Today, one of the largest simulations is the Millennium Simulation (Springel et al., 2005), consisting of $N = 2160^3 \approx 10^{10}$ particles, each one of mass 1.2×10^9 $M_\odot$, in a cubic box with ($z = 0$) side length 2.23×10^9 light years ($500h^{-1}$ Mpc, where h is the fractional Hubble constant). This was one in a long series of simulations designed to study cosmological structure under the cold dark matter paradigm. Such simulations show a rich network of filaments and voids, as well as dark matter haloes, spread throughout the cosmological volume. Critically, the Millennium Simulation covered a large enough volume to contain a few very rare objects (such as rich galaxy clusters, and quasars), yet with enough particles to resolve dwarf galaxies as faint as the Small Magellanic Cloud. Such a large dynamic range enables us to make robust and precise theoretical predictions about the growth of structure, galaxy formation, and the signatures of dark matter and dark energy. This is essential for comparing with the accurate measurements coming from very large surveys, such as the SDSS. With the continuing relentless increase in computing power (a doubling in performance every two years, known as 'Moore's Law'), even larger simulations will be run in the future, such as Millennium XXL ($N = 3 \times 10^{11}$). These will help to numerically model much of what will be seen in the next generation of surveys of large scale structure, weak lensing, and other non-linear processes.

The observational data used in this work come from some of the largest (6.5-m and 10-m) optical telescopes in the world. The data include spectra of 23 quasars at $z \gtrsim 4.5$, covering the range $\sim 3000-$

Name	z_q	Inst.	Dates	t_{exp} (hrs)	Ref.	Chapter
SDSS J1148+5251	6.42	HIRES	Jan 2005 - Feb 2005	14.2[a]	1	4, 6
SDSS J1030+0524	6.31	HIRES	Feb 2005	10.0	1	4, 6
SDSS J1623+3112	6.25	HIRES	Jun 2005	12.5	1	4, 6
SDSS J1048+4637	6.23	HIRES	Feb 2006	5.0	5	6
SDSS J0818+1722	6.02	HIRES	Feb 2006	8.3	2	4, 6
SDSS J1306+0356	6.02	MIKE	Feb 2007	6.7	4	4, 6
SDSS J0002+2550	5.82	HIRES	Jan 2005 - Jul 2008	14.2	1,4	4, 6
SDSS J0836+0054	5.81	HIRES	Jan 2005	12.5[a]	1	4, 6
SDSS J1044-0125	5.80	MIKE	Feb 2007 - Mar 2008	7.5	5	6
SDSS J0231-0728	5.41	HIRES	Jan 2005 - Feb 2005	10.0	1	4, 6
SDSS J1659+2709	5.33	HIRES	Sep 2007 - Jul 2008	11.7	3	4, 6
SDSS J0915+4924	5.20	HIRES	Feb 2005	10.0	1	4, 6
SDSS J1204-0021	5.09	HIRES	Jan 2005 - Feb 2005	6.7	1	4, 6
SDSS J0040-0915	4.98	MIKE	Oct 2007	8.3	3	6
SDSS J0011+1440	4.97	HIRES	Sep 2007	6.7	3	4, 6
SDSS J2225-0014	4.89	MIKE	Oct 2007	5.0	4	4, 6
SDSS J1616+0501	4.88	MIKE	Mar 2008	3.3	4	4, 6
SDSS J2147-0838	4.59	MIKE	Oct 2007	8.3	3	4, 6
BR 0353-3820	4.59	MIKE	Nov 2006 - Oct 2007	9.2	3	6
BR 0006-6208	4.52	MIKE	Nov 2006	1.7	6	6
BR 1033-0327	4.52	MIKE	Nov 2005	4.2	6	6
BR 0714-6449	4.49	MIKE	Feb 2007	6.7	3	6
BR 0418-5726	4.48	MIKE	Nov 2005 - Nov 2006	15.8	3	6

Table 2.1: The list of quasars included in this thesis. Columns give the quasar name and redshift, details of the observations, references to the paper in which they are first described, and the Chapter in which they are used.
[a] The present reductions include only data taken with the upgraded detector.
References: (1) Becker et al. (2006); (2) Becker et al. (2007); (3) Becker et al. (2011a); (4) Calverley et al. (2011); (5) Becker et al. (2011b); (6) This work.

10000 Å. Accompanying these are detailed, state-of-the-art simulations, which we use to generate artificial spectra for comparison. The highest resolution simulation has $N = 2 \times 512^3 \approx 10^{8.4}$ particles, each with a mass of 1.3×10^5 $M_\odot$. This resolution is necessary to accurately model the Lyα forest at $z \geq 5$. In the rest of this Chapter I will describe the acquisition and reduction of the data, as well as the generation and processing of the simulations.

2.1 Observations of high-redshift quasars

Spectra of a total of 23 quasars at $z_q \gtrsim 4.5$ were used in this thesis. Twelve were taken using the HIRES instrument on Keck, and eleven were taken with MIKE on Magellan. They are summarised in Table 2.1. All of the $z \geq 5.8$ quasars were discovered in the SDSS (the discovery papers are Fan et al., 2001, 2003, 2004, 2006c), and all SDSS quasars have flux-calibrated spectra, in either the SDSS database itself or the discovery papers. For the proximity effect work in Chapter 4, only those with flux-calibrated spectra could be used (for accurate luminosity determinations), limiting the sample to the SDSS quasars. In addition, SDSS J1048+4637 and SDSS J1044-0125 are broad absorption line (BAL) quasars,

Figure 2.1: Image of the twin Keck telescopes, on the summit of Mauna Kea on Hawaii. HIRES is mounted on Keck I, which is on the left. *Photo taken by the author, Oct 2010.*

meaning absorption would affect the proximity region, making them unsuitable for this analysis. Similarly, SDSS J0040-0915 is too noisy within the proximity region to provide meaningful data. All 23 quasars were used for the Lyα forest analysis in Chapter 6, as quasar luminosities were not required and the proximity region could be ignored. The telescopes and instruments used for obtaining the spectra will now be described.

2.1.1 The Keck telescopes and HIRES spectrograph

The Keck telescopes[6] sit just below the summit of the extinct volcano Mauna Kea, on the Big Island of Hawai'i. At 4,123 m (13,527 ft) above sea level, they provide an ideal place from which to observe. The site resides above a reasonable percentage of the atmosphere, which would otherwise distort images, and has little water vapour in the air, very dark skies, generally good weather, and an almost equatorial location (meaning long nights all year round). Keck I (on the left in Fig. 2.1) began observations in May 1993, whilst its twin, Keck II, was completed in October 1996. The primary mirror on both telescopes is composed of 36 hexagonal segments which act together as a single mirror. A computer controls each segment's position to an accuracy of 5 nm. The telescope aperture area is equivalent to that with a circular aperture 10-m in diameter, making the Keck telescopes some of the largest in the world.

The High Resolution Echelle Spectrometer[7] (HIRES) is positioned permanently at the right Nasmyth focus (see Fig. 2.2) of Keck I, and is fully described in Vogt et al. (1994). It is an echelle spectrograph, using two diffraction gratings perpendicular to each other. The first is relatively coarse (i.e. a small number of grooves per mm), and the angle of the grooves to the surface, known as the blaze, is high, so that high diffraction orders are used. Since the spacing between orders decreases towards

[6] http://www.keckobservatory.org/
[7] http://www2.keck.hawaii.edu/inst/hires/

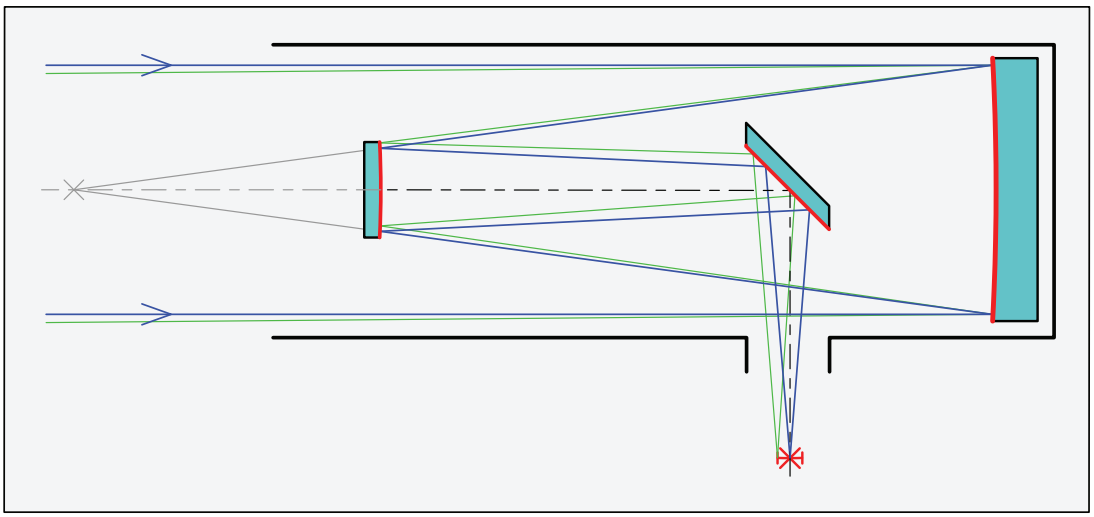


Figure 2.2: Light path in a Nasmyth telescope. The largest mirror (right) is known as the primary mirror, and is concave. The incoming light reflects off the primary mirror onto a convex secondary mirror (left), and then onto a flat tertiary mirror (centre) that reflects the light off to one side (in this case the left, in the frame of the telescope). The exit point of the beam is normally chosen to be on the altitude axis of the telescope, so that the instrument that records the light does not need to move up and down with the telescope. This makes the Nasmyth focus the ideal choice for heavy instruments like spectrographs that would otherwise upset the balance of the telescope. Image taken from *http://en.wikipedia.org/wiki/File:Nasmyth-Telescope.svg*

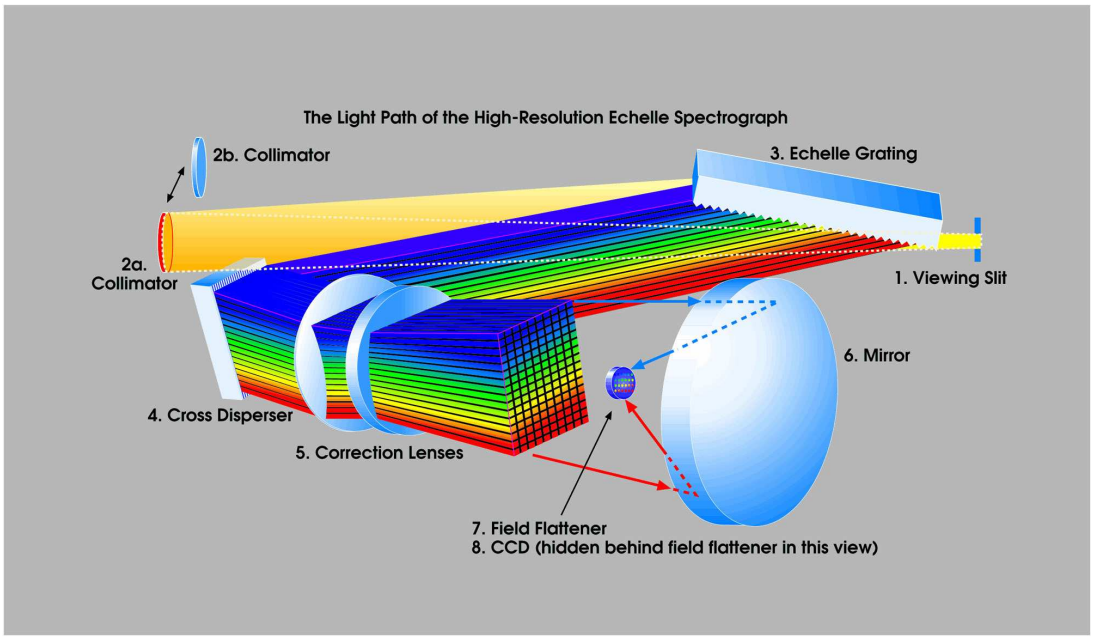


Figure 2.3: The light path through the High Resolution Echelle Spectrometer (HIRES). Image taken from *http://www.ucolick.org/~vogt/images/hires.jpg*

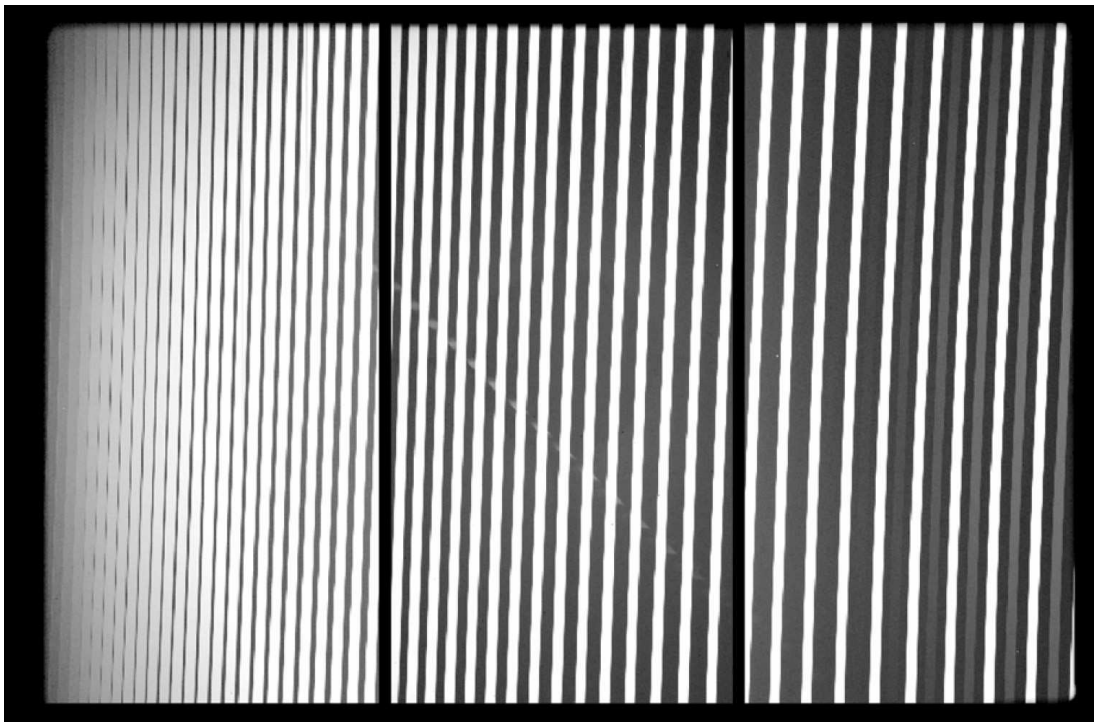


Figure 2.4: Example of how the spectra from HIRES is formatted, using a flat field (i.e. uniform illumination) source. There are 3 separate CCDs over which the light is spread, corresponding to blue (left), green (middle) and red (right). Each white band is a different spectral order, and the spacing between orders decreases with wavelength. Wavelength increases from bottom to top and from left to right on each chip. Image taken from the HIRES data guide, at *http://www2.keck.hawaii.edu/realpublic/inst/hires/docs/hires_data.pdf*.

higher (bluer) orders, overlap between the orders can occur. The second diffraction grating, known as the cross disperser, is blazed to separate out the orders. The final spectrum consists of stripes with different, but slightly overlapping, wavelength ranges that run across the imaging plane in an oblique pattern. Echelle spectrographs allow a large spectral range to be folded onto a small two-dimensional CCD[8], rather than a very long, linear array, leading to very efficient spectrographic observations. The light path within HIRES is shown in Fig. 2.3. In August 2004 the CCD detector was upgraded, and all HIRES observations shown here use only data taken with the new detector. HIRES uses three separate CCDs (one for the red orders, one for the green orders, and one for the blue orders). As an example, the raw output from the CCD for an internal quartz lamp (known as a flat field exposure) is shown in Fig. 2.4. Each CCD has 2048×4096 pixels, with a pixel size of 15 μm. Due to the high redshift of the objects, all the HIRES observations were carried out using the cross disperser optimised for the red spectral range. The objects were observed with a 0.86 arcsecond-wide slit, and an echelle grating with spectral resolution $R = 40\,000$, corresponding to a full width at half-maximum (FWHM) velocity resolution of 6.7 $\mathrm{km\,s^{-1}}$. The final pixel size of the extracted spectra is 2.1 $\mathrm{km\,s^{-1}}$, so each resolution element is adequately Nyquist sampled (≥ 2 pixels per resolution element).

[8] Charge-coupled device

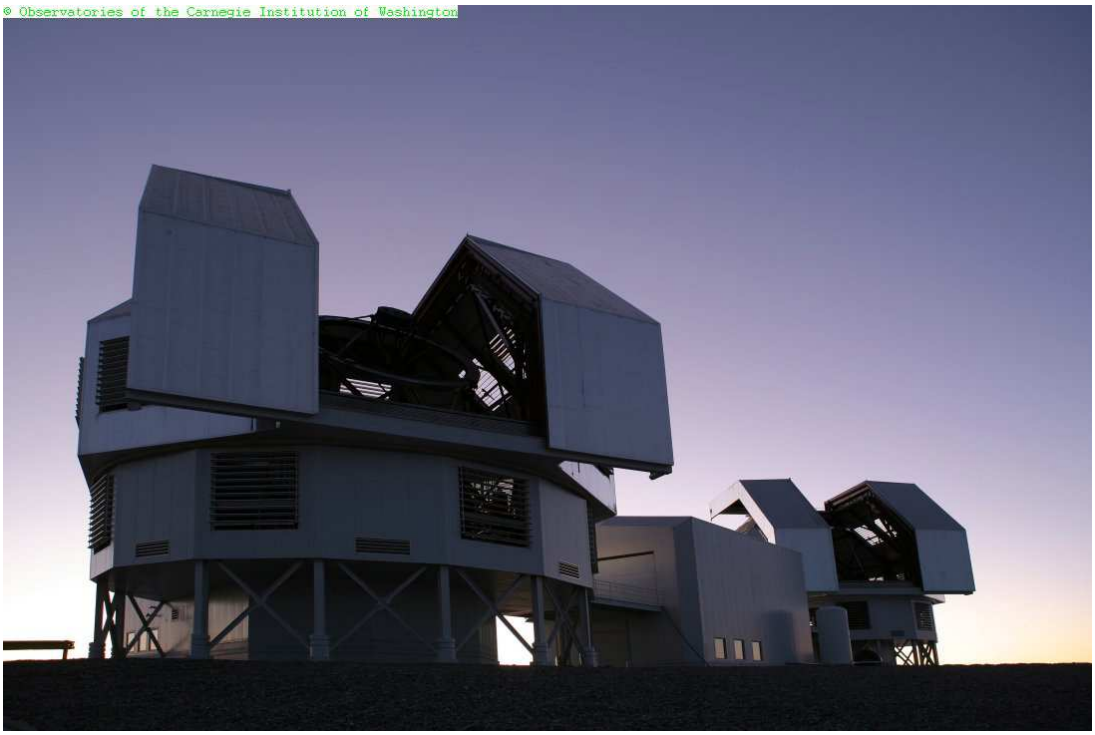


Figure 2.5: Image of the two Magellan telescopes, at the Las Campanas Observatory in Chile. MIKE is attached to Magellan-II Clay, which is on the right. *Photo taken by Matias Del Campo in March 2006, and available at http://www.baade-clay.org/content/photos. Used with permission of The Observatories of the Carnegie Institution for Science.*

2.1.2 The Magellan telescopes and MIKE spectrograph

The Magellan telescopes[9] are based at the Las Campanas Observatory[10], located at an altitude of 2380 m (7810 ft) in the Atacama Desert in Chile. Magellan I is known as the Baade telescope (on the left in Fig. 2.5) and began observations in September 2000, whilst Magellan II is known as the Clay telescope and started observing in September 2002. Both have a single circular primary mirror with a diameter of 6.5-m. An array of pistons underneath the mirror control the shape of its surface to within 50 nm, correcting for any distortions due to flexing under gravity.

The Magellan Inamori Kyocera Echelle spectrograph[11] (MIKE) is at the Nasmyth focus of the Magellan II (Clay) telescope, and was installed in November 2002. The instrument is described in detail in Bernstein et al. (2003). Like HIRES, MIKE is an echelle spectrograph, offering the same advantage of a large spectral range folded onto a small two-dimensional CCD. However, MIKE has two key differences. Firstly, the cross disperser is a prism rather than a grating, in order to allow a higher throughput. Secondly, light entering the instrument is passed through a dichroic, which separates it into a red channel and a blue channel, with the crossover wavelength at ~ 445 nm. Each channel has its own light path, with optical elements optimised for the high (blue) or low (red) orders. The channels each have a CCD with 2048×4096 pixels and a pixel size of 15 μm, allowing two spectra (red and blue) to be

[9] http://obs.carnegiescience.edu/Magellan/
[10] http://www.lco.cl/
[11] http://www.lco.cl/telescopes-information/magellan/instruments/mike

recorded independently. The objects were observed with the 1-arcsec slit, giving a spectral resolution of $R = 22\,000$ in the red channel, corresponding to a FWHM velocity resolution of 13.6 $\mathrm{km\,s^{-1}}$. The pixel size is 5.0 $\mathrm{km\,s^{-1}}$, so again there is adequate Nyquist sampling.

2.1.3 Details on the data reduction techniques

The spectra were reduced by G. Becker using a custom set of IDL routines that employ optimal sky subtraction techniques (Kelson, 2003), as detailed in Becker et al. (2006, 2007). The data reduction process will now be briefly outlined. The routines are:

1. The orders are traced using a bright source, e.g., a standard star, or an internal lamp through a pinhole. Since the orders will be slightly curved, a new coordinate grid is defined so that the x (spatial) coordinate on the CCD (horizontal in Fig. 2.4) and y (dispersion) coordinate (vertical) are translated to new coordinates x' and y', where the edges of the order have constant x' and the spectral lines have constant y'. This allows the sky to be modelled in the native spectra format, without 'straightening' the orders.

2. During an exposure, a voltage is applied across the CCD to avoid spurious negative counts, and so that it will respond linearly to incoming photons (each of which will produce a certain number of electrons). This means that even without any incident photons, a certain number of electrons will be counted in each pixel. This is known as the 'bias'. The bias is measured from a region of 'imaginary' pixels off the end of the CCD (equivalent to a zero-second exposure). It is measured for each row in the CCD, and is subtracted away from all pixels in that row.

3. The image is then flat fielded in order to remove any pixel-to-pixel variation in sensitivity. For this, a bright, uniform illumination source is required, for which an internal lamp is used. The science image (after subtracting the bias) is divided through by the resulting (non-uniform) flat-field image.

4. Arc lamp exposures are taken using a series of lamps containing elements with strong atomic lines (e.g. Thorium-Argon in HIRES), which allow for wavelength calibration. This is required for the transformation from $y' \rightarrow \lambda$. Also, the movement of the Earth around the sun induces an annual Doppler shift of ~ 30 $\mathrm{km\,s^{-1}}$ in the spectra, and so the observed wavelengths used are adjusted to be heliocentric in order to compensate for this effect. A conversion to vacuum wavelengths is also applied.

5. The sky subtraction is done with a 2-D fit to the sky, as detailed in Kelson (2003). A 2-D basis spline (b-spline) fit is carried out across the x' and y', ignoring the range in x' subtended by the object in the slit, to accurately model the number of counts due to the sky at the object x'. This fit is then subtracted from the 2-D frame.

6. To go from the 2-D CCD image to a 1-D spectrum, the object flux is extracted according to the optimal routine described in Horne (1986). The final flux for a given wavelength in a given order

is the weighted sum of all non-zero pixels, extracted simultaneously from all exposures of the object. There are two weights used: the first is based upon the shape of the object profile at a given wavelength, and the second is based on the variance, which is calculated analytically for each wavelength from the read noise, gain, sky flux and object flux.

7. So far, the process has produced a 1-D spectrum for each order, each with a slightly overlapping spectral range. To combine the orders into a single 1-D spectrum, a standard star with a simple spectrum (normally a white dwarf) is used to correct for the response function of each order.

8. Not only is the sky bright at particular wavelengths (known as skylines), but molecules of O_2 and H_2O in the atmosphere cause absorption at particular wavelengths known as telluric absorption. The strength of the absorption is determined by the amount of atmosphere through which the object light passes and the amount of water vapour in the air, and so will change from exposure to exposure. To deal with this, a very high-resolution standard sky spectrum is used. The optical depth and resolution is adjusted to that of the data (after compensating for the heliocentric correction), and the data is divided through by the best matching telluric spectrum.

9. The final step is then to continuum normalise the spectra, which is done following the method described in Bolton et al. (2010). The spectrum is first divided through by a power law $F_\nu \propto \nu^{-0.5}$, normalised at 1280 $(1+z)$ Å, and the Lyα emission line is then fitted with a slowly varying spline. It is difficult to fit the continuum over the Lyα forest at high redshift due to the low flux levels; however, in the proximity region the transmitted flux maxima will be nearer to the continuum. We estimate the continuum to be within $\sim 20\%$ of the correct value over the region of interest for Chapter 4. The difficulty in fitting the continuum in the Lyα forest is discussed in more detail for the measurements made in Chapter 6.

All the spectra presented in this thesis were reduced in this way. The final reduced spectra (before continuum normalisation) are shown in their rest frame wavelengths in Fig. 2.6 and Fig. 2.7 for the HIRES and MIKE data, respectively.

2.2 Constructing synthetic Lyman alpha forest spectra

A total of 23 different cosmological hydrodynamical simulations were used in this thesis. These are divided into two suites, each summarised in Table 2.2. All of them were run by J. Bolton. The simulations were performed using a customized version of the code `GADGET-3`, an updated version of the publicly available code `GADGET-2` (Springel, 2005)[12]. `GADGET-3` computes gravitational forces with a hierarchical tree algorithm in combination with a particle-mesh (PM) scheme for long-range gravitational forces[13], and represents gas dynamics with the Lagrangian (particle-tracking) method of

[12] http://www.mpa-garching.mpg.de/gadget/

[13] The review by Bertschinger (1998) provides a good description of the different categories of N-body simulation codes.

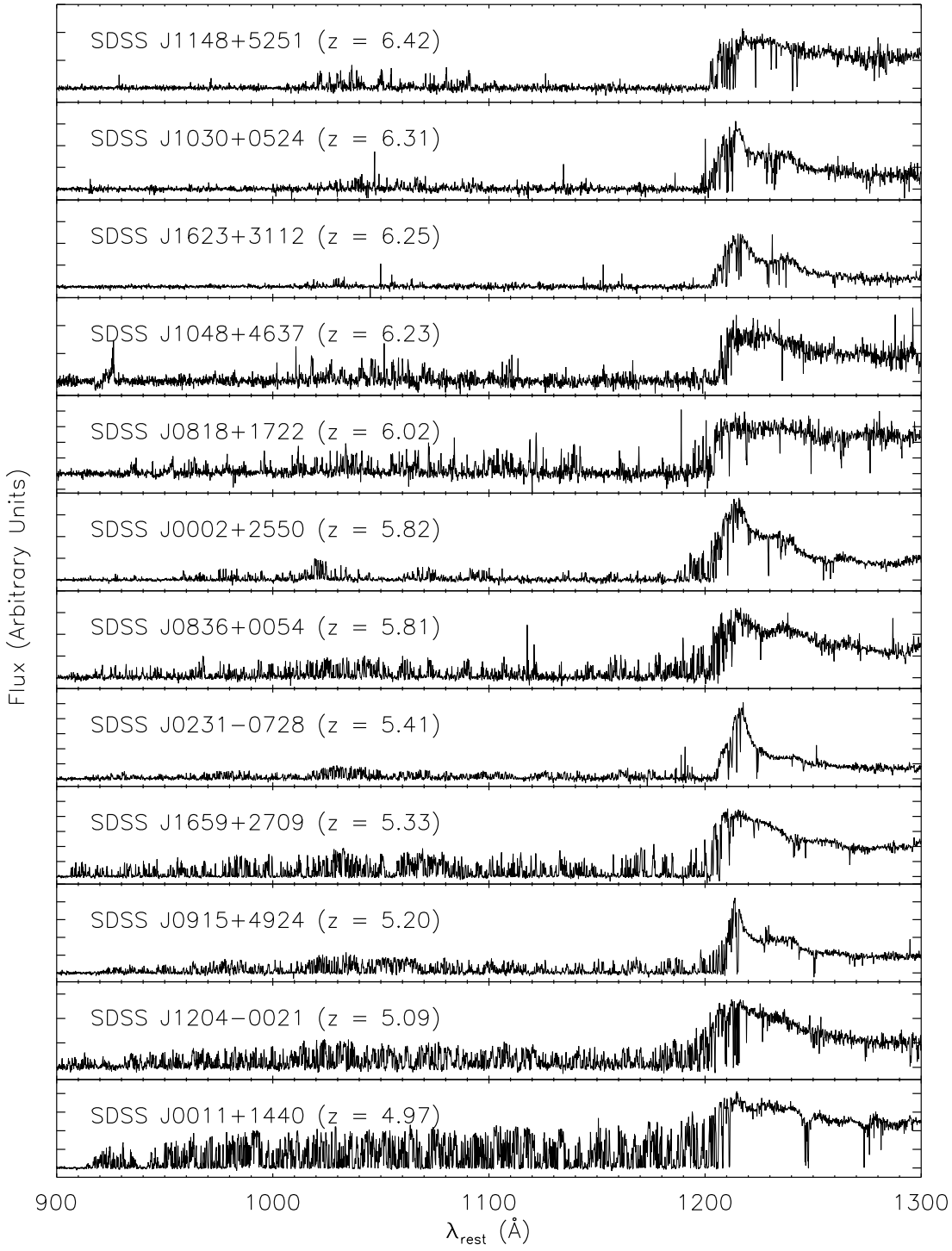


Figure 2.6: The HIRES spectra in their restframe wavelengths, showing the Lyα emission line at 1216 Å, and the Lyα forest extending down to the Lyman limit at 912 Å.

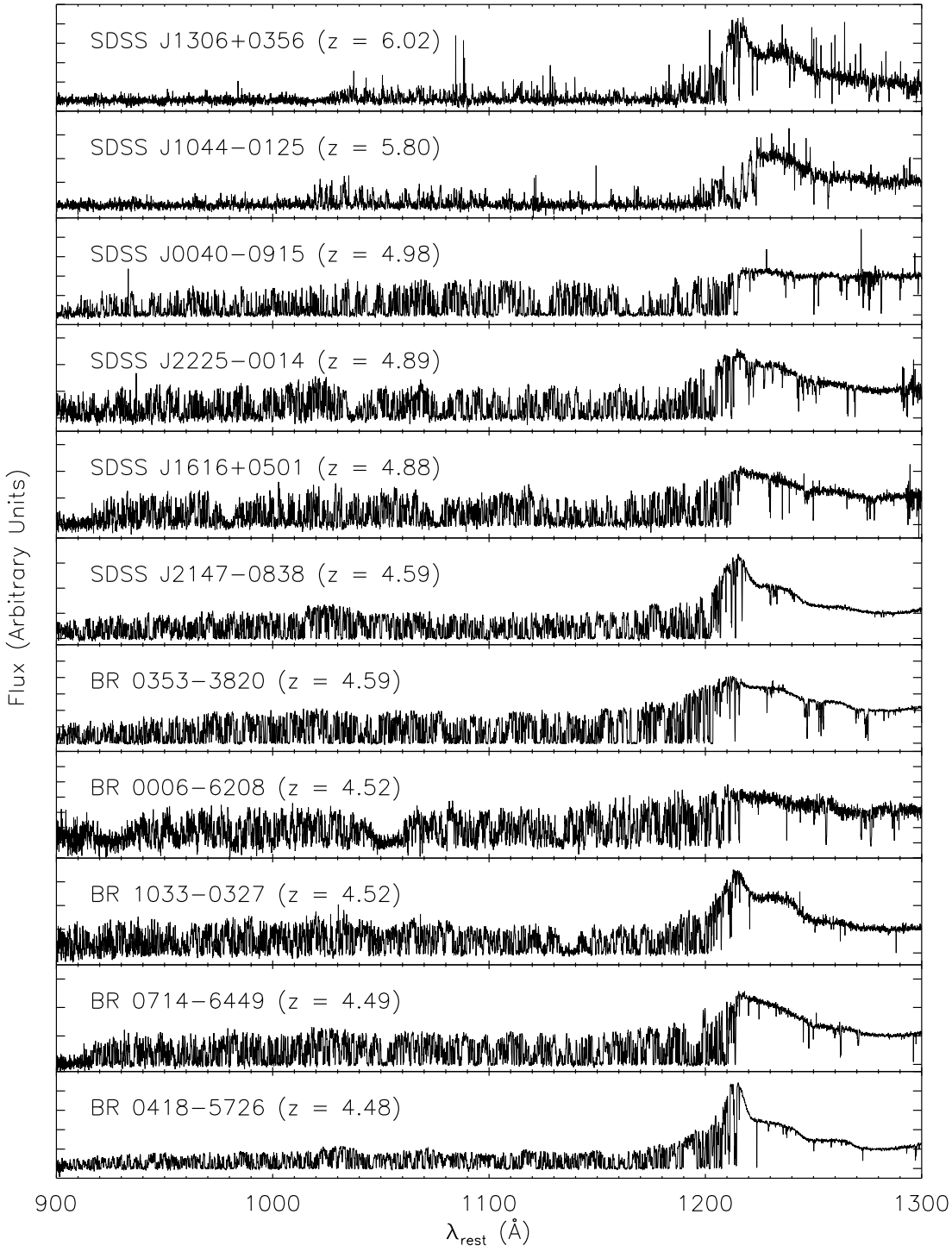


Figure 2.7: The MIKE spectra in their restframe wavelengths, showing the Lyα emission line at 1216 Å, and the Lyα forest extending down to the Lyman limit at 912 Å.

Model	L [h^{-1} Mpc]	N	$M_{\rm gas}$ [h^{-1} M$_\odot$]	Ref.
A	20	2×100^3	1.03×10^8	1
B	20	2×200^3	1.29×10^7	1
C	20	2×400^3	1.61×10^6	1
D	40	2×200^3	1.03×10^8	1
E	40	2×400^3	1.29×10^7	1
F	80	2×400^3	1.03×10^8	1

(a) Summary of the properties of the first suite of simulations, used in Chapters 3, 4 and 5.

Model	L [h^{-1} Mpc]	N	$M_{\rm gas}$ [h^{-1} M$_\odot$]	ζ	ξ	$T_0^{z=3}$ [K]	$\gamma^{z=3}$	Ref.
A15	10	2×512^3	9.2×10^4	0.30	0.00	5100	1.52	2
B15	10	2×512^3	9.2×10^4	0.80	0.00	9600	1.54	2
C15	10	2×512^3	9.2×10^4	1.45	0.00	14000	1.54	2
C10	10	2×512^3	9.2×10^4	1.45	-1.00	13700	1.02	2
D15	10	2×512^3	9.2×10^4	2.20	0.00	18200	1.55	2
D13	10	2×512^3	9.2×10^4	2.20	-0.45	18100	1.32	2
D10	10	2×512^3	9.2×10^4	2.20	-1.00	18000	1.03	2
D07	10	2×512^3	9.2×10^4	2.20	-1.60	17900	0.71	2
E15	10	2×512^3	9.2×10^4	3.10	0.00	22500	1.55	2
E10	10	2×512^3	9.2×10^4	3.10	-1.00	22200	1.04	2
F15	10	2×512^3	9.2×10^4	4.20	0.00	27000	1.55	2
G15	10	2×512^3	9.2×10^4	5.30	0.00	31000	1.55	2
T15S	10	2×512^3	9.2×10^4	Varied	0.00	18200	1.53	2
T15F	10	2×512^3	9.2×10^4	Varied	0.00	18600	1.54	2
TZ9	10	2×512^3	9.2×10^4	Varied	0.00	17800	1.56	3
TZ12	10	2×512^3	9.2×10^4	Varied	0.00	17500	1.57	3
TZ15	10	2×512^3	9.2×10^4	Varied	0.00	17500	1.57	3

(b) Summary of the properties of the second suite of simulations, used in Chapter 6.

Table 2.2: Summary of the the properties of the hydrodynamical simulations used in this thesis. Columns give the model name, the size of each box in comoving Mpc, the total particle number (gas + dark matter), the particle mass resolution, and references to the paper in which they are first described. For the second suite of simulations, there are additional columns for the constants ζ and ξ used in rescaling the Haardt and Madau (2001) photoheating rates as $\epsilon = \zeta\Delta^{\xi}\epsilon^{\rm HM01}$, and the parameters of the temperature-density relation, $T = T_0\Delta^{\gamma-1}$, at $z = 3$, measured from particles with $-0.5 \leq \log\Delta \leq 0.0$. The values of T_0 are given to three significant figures. References: (1) Bolton and Becker (2009); (2) Becker et al. (2011a); (3) This work.

smoothed particle hydrodynamics (SPH). The main advantage of using parallel TreePM codes is that they are fast, with the number of operations per timestep being of order $N \log N$ for N particles, compared with order N^2 for direct summation techniques (as used in the first N-body simulations). The main disadvantage is that they require a substantial amount of memory to store the tree nodes.

All the simulations were run in periodic cubic boxes of side length L, with L varying between $10h^{-1}$ and $80h^{-1}$ comoving[14] Mpc, and total particle number N varying from $10^{6.3}$ to $10^{8.4}$. The particles were split evenly between dark matter and gas. The small box sizes with large N were chosen to adequately resolve the density field within voids, since the voids determine the properties of the Lyα forest at $z \geq 5$ (Bolton and Becker, 2009). They were started at $z = 99$ with initial conditions from the CDM transfer function of Eisenstein and Hu (1999). Star formation is included by converting all particles with an overdensity $\Delta \equiv \rho / \langle \rho \rangle > 10^3$ and $T < 10^5$ K into collisionless stars. The baryons in the simulations are photoionized and heated by the UVB, which is taken from the model by Haardt and Madau (2001) and includes emission from both quasars and galaxies.

The first suite are a selection of those used in Bolton and Becker (2009), which were carried out on the COSMOS cosmology supercomputer[15] at DAMTP in Cambridge. The UVB is switched on at $z = 9$ and applied in the optically thin limit ($\tau < 1$) using a non-equilibrium ionization algorithm. It assumes the cosmological parameters $(\Omega_\mathrm{m}, \Omega_\Lambda, \Omega_\mathrm{b}h^2, h, \sigma_8, n_\mathrm{s}) = (0.26, 0.74, 0.024, 0.72, 0.85, 0.95)$, and that the IGM consists of primordial gas with a helium fraction (by mass) of $Y = 0.24$ (Olive and Skillman, 2004). Simulation outputs were at $z = (2, 3, 4, 5, 6)$.

The second suite consists of the simulations used in Becker et al. (2011a), with some additional models run especially for this work (TZ9, TZ12 and TZ15). These were carried out on the Darwin Supercomputer[16] of the University of Cambridge High Performance Computing Service. These simulations use the cosmological parameters $(\Omega_\mathrm{m}, \Omega_\Lambda, \Omega_\mathrm{b}h^2, h, \sigma_8, n_\mathrm{s}) = (0.26, 0.74, 0.023, 0.72, 0.80, 0.96)$ and also that $Y = 0.24$. The UVB is switched on at $z = 9$, except for models TZ12 and TZ15 where it is switched on at $z = 12$ and $z = 15$ respectively. It is applied in the optically thin limit, although unlike in the first suite, the gas is assumed to be in ionization equilibrium. Simulation outputs were at $1.831 \leq z \leq 6.010$, although only those at $z \geq 3.991$ were used in Chapter 6. Gas with $\Delta \lesssim 10$ follows a tight temperature-density relation of the form $T = T_0 \Delta^{\gamma - 1}$, where T_0 is the temperature at the mean density ($\Delta = 1$), and γ is the logarithmic slope of the relation (Hui and Gnedin, 1997). This relationship is shown for the simulations D15 and D07 at two different redshifts in Fig. 2.8. The red dashed line shows the relationship measured from the median temperatures at $\log\Delta = 0$ and $\log\Delta = -0.5$. The implicit Haardt and Madau (2001) photoheating rates of the simulations, ϵ^HM01, were rescaled such that $\epsilon = \zeta \Delta^\xi \epsilon^\mathrm{HM01}$, where ζ and ξ are the constants shown in Table 2.2, in order to achieve a variety of thermal histories with different values of $T_0(z)$ and $\gamma(z)$. The fiducial γ (i.e. $\xi = 0$) was ~ 1.5 at $z \sim 3$ (see top panels in Fig. 2.8), although flattened ($\gamma \sim 1.3$), isothermal ($\gamma \sim 1.0$), and inverted ($\gamma \sim 0.7$; bottom panels in Fig. 2.8) temperature-density relations were also investigated. Models beginning with 'T' were used to investigate different thermal histories from hydrogen reionization, using

[14] 'Comoving' means that the Hubble expansion has been factored out, such that $L_\mathrm{comoving} = (1 + z) L_\mathrm{physical}$

[15] http://www.damtp.cam.ac.uk/cosmos/

[16] http://www.hpc.cam.ac.uk/services/darwin.html

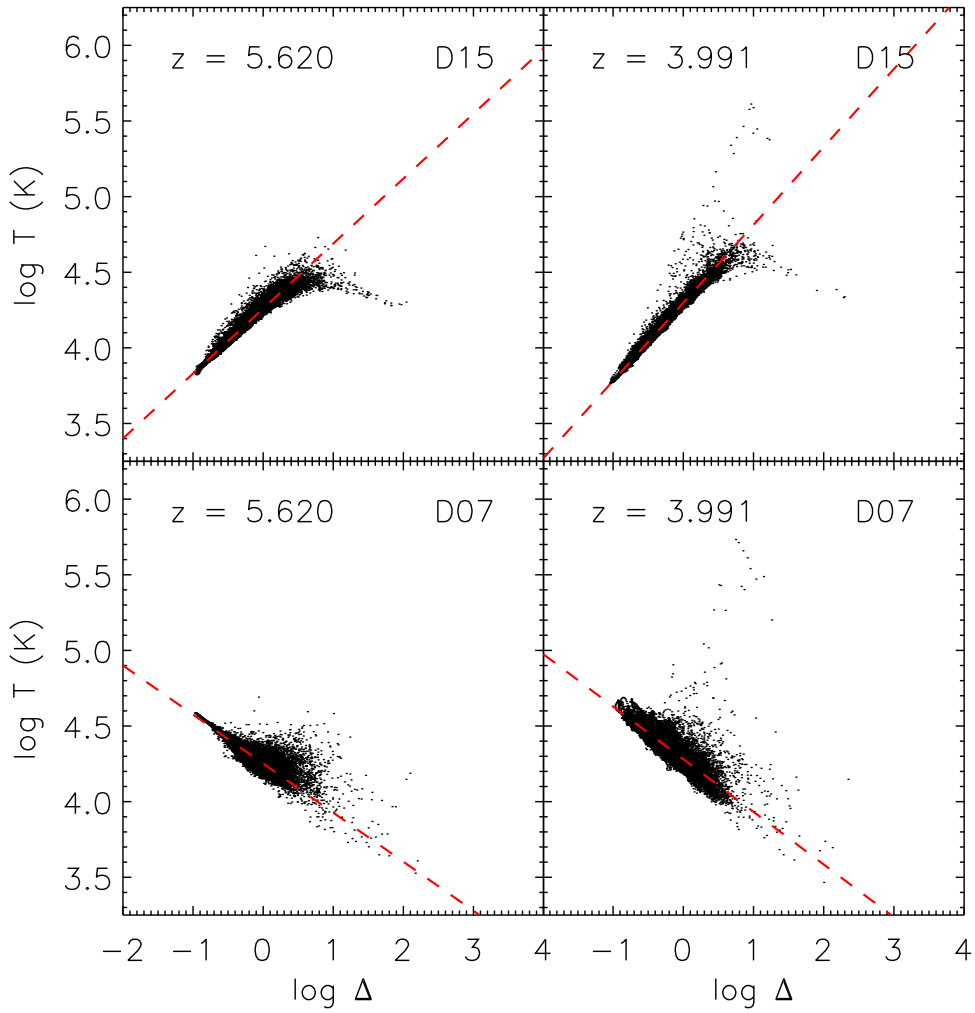


Figure 2.8: Plot of the temperature-density relation from two of the simulations from the second suite at two different redshifts. In the top panels $\gamma \sim 1.5$, whilst in the bottom panels $\gamma \sim 0.7$. The red dashed line is the power-law relationship of the form $T = T_0 \Delta^{\gamma-1}$ where γ is measured in the region $-0.5 \leq \log \Delta \leq 0.0$.

a density-independent scaling that varies with redshift (see Chapter 6 for more details).

Synthetic Lyα spectra were constructed from interpolated quantities of the relevant fluid variables from the distribution of SPH particles along sightlines drawn parallel to the box boundaries (e.g. Theuns et al., 1998). Each sightline has an associated density, Δ, peculiar velocity[17], $v_{\rm pec}$, neutral H I fraction, $x_{\rm HI}$, and temperature field. In all of the simulations, these fields were 1024 pixels long. The neutral hydrogen number density, $n_{\rm HI}$, is first computed as

$$n_{\rm HI} = \frac{x_{\rm HI}}{m_{\rm p}}(1-Y)\rho\,, \quad \text{where} \quad \rho = \Delta \frac{3H_0^2\Omega_{\rm b}}{8\pi G}(1+z)^3\,, \tag{2.1}$$

and then the H I optical depth at pixel i, $\tau(i)$, is calculated by summing up all the contributions from

[17] Random motions of the gas due to local gravitational potentials, and separate from the Hubble velocity, $v_{\rm H}$.

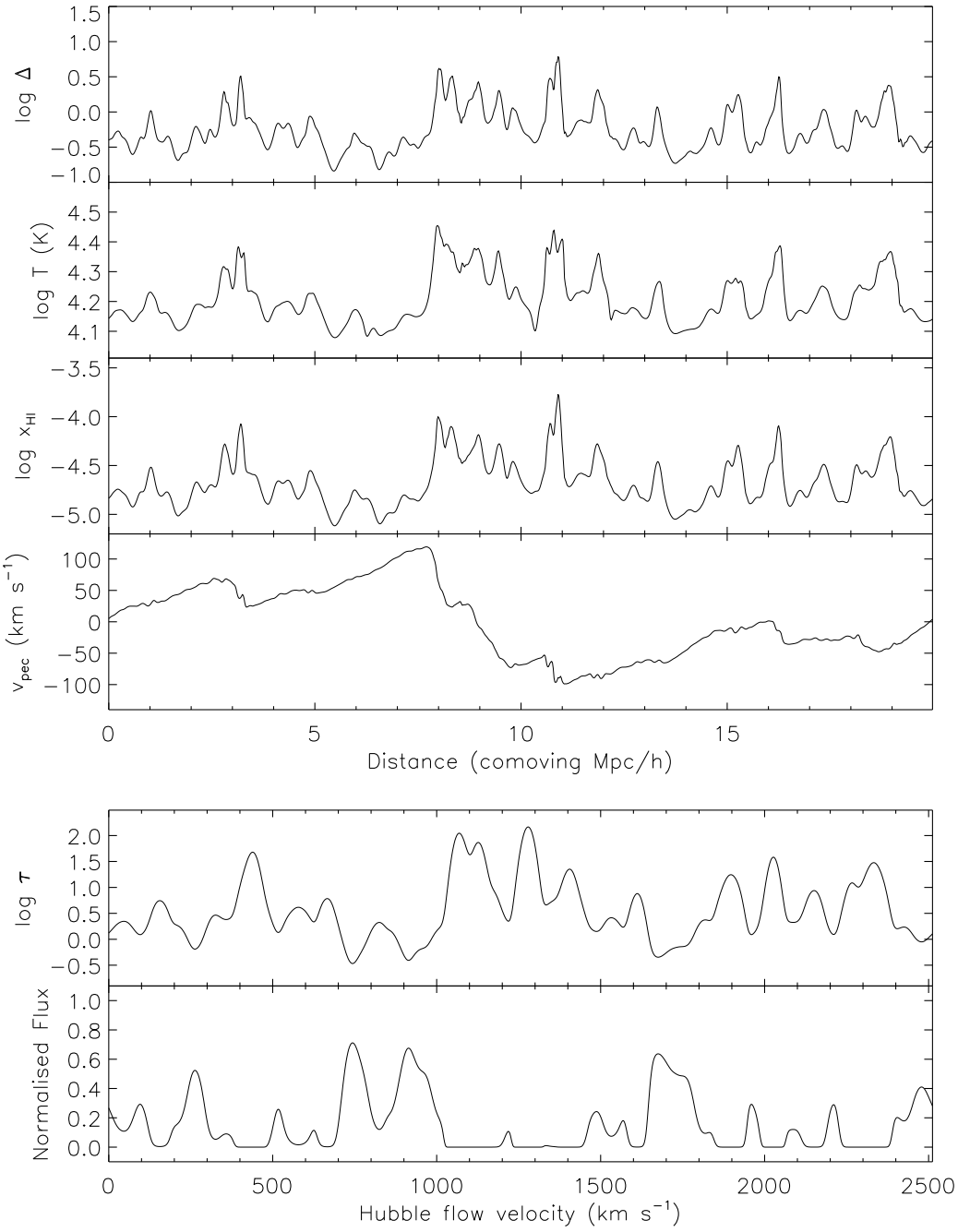


Figure 2.9: Simulation outputs used to construct synthetic Lyα spectra. The top four panels from top to bottom are the line-of-sight density, temperature, neutral hydrogen fraction and peculiar velocity fields, plotted against their position within the box, for a random sightline at $z = 5$ in Model C. The optical depth, τ, is then calculated using Equation 2.2. In the bottom panel is the resulting noiseless spectrum. The cyclic nature of the simulation box is clear in each of these fields, and allows τ to be calculated for the pixels at either end of the sightline.

all pixels along the sightline that, due to redshifting and blueshifting by a non-zero $v_{\rm pec}$, correspond to a particular pixel in velocity space. Assuming a Gaussian line profile, this gives

$$\tau(i) = \frac{c\sigma_\alpha \delta R}{\pi^{1/2}} \sum_{j=1}^{N} \frac{n_{\rm HI}(j)}{b_{\rm HI}(j)} \exp\left[-\left(\frac{v_{\rm H}(i) - u(j)}{b_{\rm HI}(j)}\right)^2\right], \tag{2.2}$$

where $b_{\rm HI} = (2k_{\rm B}T/m_{\rm H})^{1/2}$ is the Doppler parameter, $\sigma_\alpha = 4.48\times10^{-18}$ cm^2 is the Lyα scattering cross-section, δR is the physical width of each pixel, and $u(j) = v_{\rm H}(j) + v_{\rm pec}(j)$. Finally, this optical depth is converted into a normalised flux, $F = e^{-\tau}$. Fig. 2.9 shows the four output fields of Δ, T, $x_{\rm HI}$, and $v_{\rm pec}$ (in physical space) and the two derived fields of τ and F (in velocity space) for a random sightline at $z = 5$ in Model C. In order to compare with the real data, the simulated spectra were convolved with a Gaussian kernel with FWHM equal to the velocity resolution of the instrument being modelled (6.7 km s^{-1} for HIRES and 13.6 km s^{-1} for MIKE) before being resampled at the instrument pixel resolution (2.1 and 5.0 km s^{-1}, respectively). Gaussian-distributed noise, as well as other imperfections, could then be added. All of the synthetic Lyα spectra used in this thesis were constructed in this way.

2.2.1 Simulating proximity regions

In Chapters 4 and 5, quasar proximity regions are inserted into the synthetic spectra. Assuming a simple isotropic ionizing flux, the photoionization rate of a quasar follows $\Gamma_{\rm q}(r) \propto 1/r^2$. In the proximity effect, the optical depths in the proximity region are modified as

$$\tau = \frac{\tau_{\rm forest}}{1+\omega(r)} \quad \text{where} \quad \omega(r) \equiv \frac{\Gamma_{\rm q}(r)}{\Gamma_{\rm bkg}} = \left(\frac{r}{R_{\rm eq}}\right)^{-2}. \tag{2.3}$$

Here, $\tau_{\rm forest}$ is the optical depth ignoring the presence of the quasar, $\Gamma_{\rm bkg}$ is the photoionization rate of the UVB and $R_{\rm eq}$ is the distance at which $\Gamma_{\rm q}(r) = \Gamma_{\rm bkg}$ (see Chapter 4 for more details). Rather than modify the τ of each sightline directly in velocity space, the neutral hydrogen number density $n_{\rm HI}$ is modified by the same parameter (to introduce the effect in physical space, which is much more realistic), such that

$$n_{\rm HI} = \frac{n_{HI,\,\rm forest}}{1+\omega(r)}. \tag{2.4}$$

This causes $n_{\rm HI}$ to be greatly decreased close to the quasar, which when propagated through Equation 2.2 leads to an area of enhanced transmission - a proximity region. Fig. 2.10 demonstrates this effect, with a quasar with $R_{\rm eq} = 5$ Mpc at $z = 5$ using random sightlines from Model C. In the top panel $n_{\rm HI}$ drops off rapidly upon approaching the quasar, following a roughly $1/r^2$ dependence (black line), creating a proximity region in the spectrum in the bottom panel. All of the synthetic quasar proximity regions used in this thesis were created in this way. Note that this neglects any heating of the gas by the quasar, which would affect $b_{\rm HI}$ and hence also τ in the proximity region.

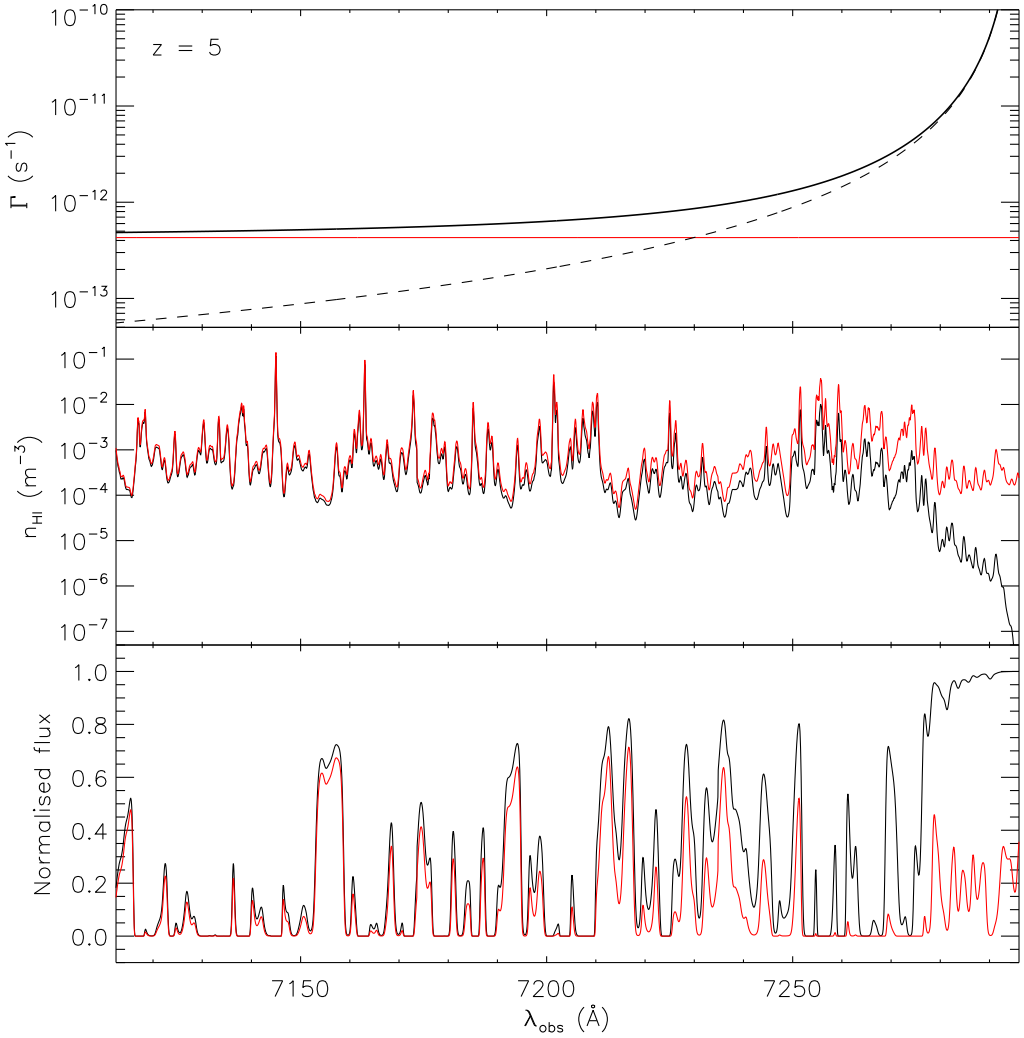


Figure 2.10: A synthetic Lyα sightline at $z = 5$ in Model C, modified to have a proximity region. In the top panel the solid black line is the total HI photoionization rate, Γ, along the spectrum. The solid red line is the background ionization rate $\Gamma_{\rm bkg}$ at $z = 5$ from Haardt and Madau (2001) and is assumed to be spatially uniform. The dashed black line is the photoionization rate due to the ionizing output of the quasar, $\Gamma_{\rm q}$, and it equals the background at $r = R_{\rm eq}$. In the middle panel is $n_{\rm HI}$ for both no quasar (red line) and with a quasar at the far right-hand side (black line). The quasar has an ionizing field that drops off as $1/r^2$, and so close to the quasar the IGM is much more highly ionized, causing $n_{\rm HI}$ to decrease. The bottom panel shows the resulting spectrum for the two cases against the observed wavelength. A small $n_{\rm HI}$ corresponds to a small τ, and thus a large normalised flux. Consequently, close to the quasar is a region of enhanced transmission, known as the proximity region. In this figure, the sightline is $60h^{-1}$ comoving Mpc long, created by combining three random sightlines, and the quasar has a luminosity chosen such that for a Haardt and Madau (2001) $\Gamma_{\rm bkg}$, $R_{\rm eq} = 5$ Mpc.

All models are wrong, but some are useful.

GEORGE E. P. BOX

3

A new method for measuring the UV background using the transmitted flux in the Lyα forest

THE redshift evolution of the mean transmitted flux in the Lyα forest can be used to estimate the evolution of the intensity of the ultraviolet background (UVB). By matching the mean optical depth in simulations to the observed value, the UVB hydrogen photoionization rate, $\Gamma_{\rm bkg}$, can be inferred. Using synthetic spectra extracted from a hydrodynamical simulation of the intergalactic medium (IGM), we investigate whether it is possible to improve upon estimates that use only the mean flux by considering the entire normalised flux probability distribution function (PDF). We develop a maximum-likelihood approach where the observed flux PDF is compared to a set of synthetic PDFs generated from simulations with a range of UVB intensities. We find that there is minimal decrease in the scatter in the distribution of $\Gamma_{\rm bkg}$ returned from spectra using our new flux-decrement technique compared to simply using the mean flux of the spectra. This holds true even for spectra with low velocity resolution or low signal-to-noise ratio, and implies that pixels with fluxes near the mean dominate our likelihood function. Our result helps validate use of the mean flux decrement, and suggests that the precision of estimates of $\Gamma_{\rm bkg}$ cannot be substantially improved with knowledge of the Lyα forest optical depth distribution. We note, however, that the use of the entire PDF can help to identify other discrepancies between the data and simulations, as used by several authors to constrain other IGM parameters.

3.1 Introduction

Much of our understanding of the intergalactic medium (IGM) comes from comparison between numerical hydrodynamical simulations of cosmic structure formation and the observed Lyα forest in the spectra of high-redshift quasars. Perhaps the most fundamental advance has been the understanding that the Lyα forest itself arises from absorption associated with density fluctuations within a warm, photoionized, and diffuse IGM, which is a prediction of hierarchical structure formation within a ΛCDM cosmology (e.g. Cen et al., 1994; Zhang et al., 1995; Hernquist et al., 1996; Rauch, 1998, and references therein).

Simulations are now able to reproduce with high accuracy many of the statistical properties of the Lyα forest through careful tuning of various cosmological and astrophysical parameters (e.g. Meiksin and White, 2003; Tytler et al., 2004; Jena et al., 2005). One of the simplest statistics to measure is the mean flux, which provides insight into the physical state of the IGM through its relationship to the optical depth of the absorption. Dividing the spectra through by a suitable continuum, the mean normalised flux, $\langle F \rangle$, is related to the so-called effective optical depth, $\tau_{\rm eff} \equiv -\ln \langle F \rangle$. The value of $\tau_{\rm eff}$ is dependent on both the assumed cosmology and the ionization and thermal state of the IGM. If all but one of these parameters are fixed, based on other independent observations, then measurements of $\tau_{\rm eff}$ can be used to constrain the value of the remaining parameter.

After reionization, the ionization state of the IGM is determined by the intensity of ultraviolet background (UVB); the net radiation field of ionizing photons emitted by stars and active galactic nuclei (AGN). Classically, this has been measured by observing the "proximity effect" in quasar spectra (e.g. Bajtlik et al., 1988). Within this model, the region of enhanced transmission blueward of the Lyα emission line is attributed to the extra ionizing photons emitted by the quasar. Consequently, given a quasar luminosity and the size of the region over which the quasar is dominant, the UVB intensity can be inferred. In addition, the mean level of absorption in the Lyα forest has also been used to estimate the UVB intensity, since it will be highly sensitive to the ionization state of the gas (e.g. Rauch et al., 1997; Songaila et al., 1999; McDonald and Miralda-Escudé, 2001; Cen and McDonald, 2002; Meiksin and White, 2004; Tytler et al., 2004; Bolton et al., 2005; Jena et al., 2005; Kirkman et al., 2005; Bolton and Haehnelt, 2007c; Faucher-Giguère et al., 2008a; Wyithe and Bolton, 2011). The majority of these papers measured the distribution function of the "flux decrement", $D_A = \langle 1 - F \rangle$, an index first introduced by Oke and Korycansky (1982), and compared it with those from simulations. Good agreement between the two distribution functions was found when the UVB in the simulations was tuned such that the mean flux decrement (and thus $\tau_{\rm eff}$) was matched to the observations. By applying this technique to many quasar spectra, it can be used to measure the redshift evolution of the UVB intensity, which in turn can be used to constrain the redshift evolution of the source population.

The observed normalised flux probability distribution function (PDF), however, has a rather complicated shape (e.g. Becker et al., 2007). It is not unreasonable, therefore, to question whether the mean flux alone is a good descriptor of the entire distribution. In this Chapter we compare estimates of the UVB hydrogen photoionization rate, $\Gamma_{\rm bkg}$, based on a more sophisticated maximum-likelihood estimator that uses the whole flux PDF to those that use the mean flux alone.

The rest of this Chapter is organised as follows. In Section 3.2 we very briefly describe the simulation we used. In Section 3.3 we describe how we convert from a mean flux to a photoionization rate, and similarly we describe our maximum-likelihood technique in Section 3.4. The two methods are then compared in Section 3.5, before we present our conclusions in Section 3.6.

3.2 Simulated spectra

We use the simulations described in Bolton and Becker (2009). These were performed using a customised version of the parallel Tree-SPH code GADGET-3, an updated version of the publicly available code GADGET-2 (Springel, 2005). The simulations have both dark matter and gas components, and assume the cosmological parameters $(h, \Omega_\mathrm{m}, \Omega_\Lambda, \Omega_\mathrm{b}h^2, \sigma_8) = (0.72, 0.26, 0.74, 0.024, 0.85)$, in agreement with the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 5-yr data set (Komatsu et al., 2009). They were started at $z = 99$ with initial conditions generated with the transfer function of Eisenstein and Hu (1999), and use the UVB model of Haardt and Madau (2001), hereafter HM01, with contributions from both galaxies and quasars. This UVB is switched on at $z = 9$ and applied in the optically thin limit. Our fiducial run is a $20h^{-1}$ comoving Mpc box with 2×400^3 particles and has a mass resolution of $1.61 \times 10^6\ h^{-1}\mathrm{M}_\odot$ (Model C in Table 4.2). Artificial Lyα spectra were constructed from line-of-sight density, peculiar velocity, H I fraction and temperature fields for 1024 random sightlines drawn parallel to the box boundaries (e.g. Theuns et al., 1998), with outputs at $z = (2, 3, 4, 5, 6)$. Each sightline is 1024 pixels long.

The simulated spectra were analysed at three different velocity resolutions; that implicit to a sightline 1024 pixels long, and that of the High Resolution Echelle Spectrometer (HIRES; Vogt et al., 1994) on Keck, and the Magellan Inamori Kyocera Echelle (MIKE) spectrograph (Bernstein et al., 2003) on Magellan. The spectra were convolved with a Gaussian with FWHM equal to the velocity resolution of the instrument that was being modelled (6.7 $\mathrm{km\,s^{-1}}$ for HIRES and 13.6 $\mathrm{km\,s^{-1}}$ for MIKE), before being resampled at the instrument pixel resolution (2.1 and 5.0 $\mathrm{km\,s^{-1}}$, respectively). Gaussian-distributed noise, as well as other imperfections, could then be added. All simulated spectra used in this Chapter are $80h^{-1}$ comoving Mpc long, created by splicing together four random sightlines.

3.3 The flux-decrement method

The optical depth, τ, of the Lyα forest is determined by both cosmological and astrophysical parameters, the most important of the latter being the thermal and ionization state of the gas. After reionization, the thermal state of the IGM is determined by competition between photoheating and adiabatic cooling. For the low density gas responsible for the absorption in the Lyα forest, the temperature and density follow a tight power-law relation of the form $T(\Delta) = T_0\Delta^{\gamma-1}$, where T_0 is the temperature at mean density, and $\Delta \equiv \rho/\langle\rho\rangle$ is the normalised gas density (Hui and Gnedin, 1997). This is known as the temperature-density ($T-\Delta$) relation. In photoionization equilibrium, $\tau \propto n_\mathrm{HI}$, where n_HI is the number density of neutral hydrogen. Since we are only interested in low density ($\Delta \lesssim 10$) gas, collisional ionization can be ignored, and thus $n_\mathrm{HI} \propto \Delta^2 T^{-0.7}/\Gamma_\mathrm{bkg}$, for temperature T,

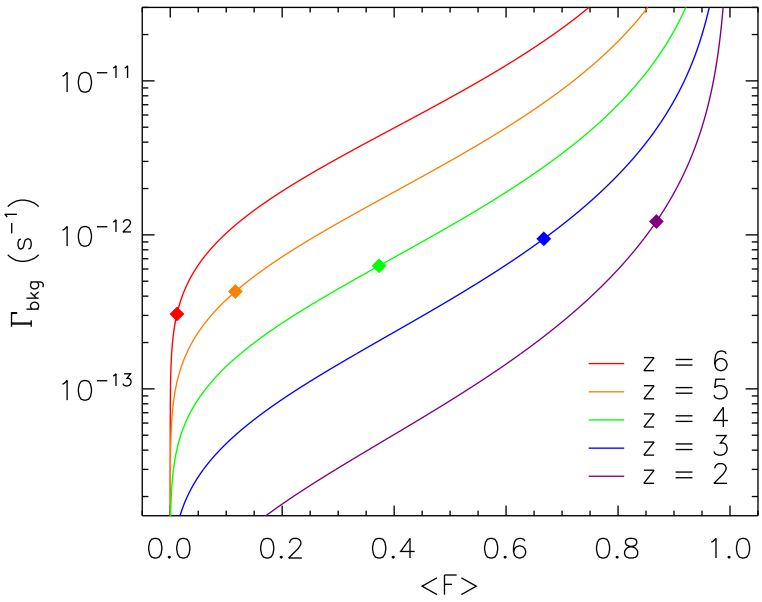


Figure 3.1: Our one-to-one functions between the mean flux of the Lyα forest and $\Gamma_{\rm bkg}$ for each redshift. This was derived by scaling the optical depths in the simulations around the Haardt and Madau (2001) $\Gamma_{\rm bkg}$ and measuring the concomitant mean flux. We use this relationship to directly convert from the mean flux of a spectrum to a UVB intensity. The filled diamonds mark the fiducial Haardt and Madau (2001) values.

and UVB hydrogen photoionization rate per atom $\Gamma_{\rm bkg}$. Combining these relations for both astrophysical parameters gives a simple relation between τ and the physical state of the gas, and is known as the fluctuating Gunn-Peterson approximation (FGPA). It describes the Lyα absorption due to an inhomogeneous IGM in photoionization equilibrium, and is related to the simple model for a homogeneous IGM first described by Gunn and Peterson (1965).

If we neglect thermal broadening and assume peculiar velocities for the typical sizes and densities of Lyα forest absorbers (Weinberg et al., 1997), then the optical depth to Lyα scattering is given by (e.g. McDonald and Miralda-Escudé, 2001)

$$\tau = \tau_0 \frac{(1+z)^6(\Omega_{\rm b}h^2)^2}{T_0^{0.7}H(z)\Gamma_{\rm bkg}(z)}\Delta^{2-0.7(\gamma-1)} \tag{3.1}$$

which holds for gas where shock heating is not important (Weinberg et al., 1999). In the above formula, $\Omega_{\rm b}$ is the fractional baryon density, h is the present day value of the Hubble constant in units of 100 km s^{-1} Mpc^{-1}, T_0 and γ are defined in the $T-\Delta$ relation, and τ_0 is a constant. The Hubble parameter $H(z) = H_0\sqrt{\Omega_{\rm m}(1+z)^3+\Omega_\Lambda}$ for a flat Universe, and so for $z \geq 2$ then $H(z) \approx H_0\Omega_{\rm m}^{1/2}(1+z)^{3/2}$. Therefore, with knowledge of the underlying density distribution and $T-\Delta$ relation at a given redshift $z \geq 2$, $\tau_{\rm eff}$ should scale with the parameter combination (e.g. Rauch et al., 1997)

$$\mu = \frac{\Omega_{\rm b}^2 h^3}{T_0^{0.7}\Omega_{\rm m}^{1/2}\Gamma_{\rm bkg}} \, . \tag{3.2}$$

Keeping all other parameters fixed, $\langle\tau\rangle \propto \Gamma_{\rm bkg}^{-1}$, and so scaling $\langle\tau\rangle$, and thus $\tau_{\rm eff}$, is equivalent to scaling $\Gamma_{\rm bkg}$. This is the critical behaviour that governs the flux-decrement method. The optical depths of the artificial spectra from a simulation are rescaled until the mean flux matches the observed mean flux measured from quasar spectra. If the simulation was run using cosmological and other astrophysical parameters determined from independent observations and assuming the FGPA approximation holds, the same scaling can be applied to the UVB assumed in the simulation, and thus $\Gamma_{\rm bkg}$ can be derived.

Our simulations use a HM01 UVB assuming contributions from stars and galaxies. We scale the optical depths of each pixel in all of the simulated sightlines by a scale factor ζ, such that

$$\tau_{\rm scaled} = \zeta \tau_{\rm HM01} \, . \tag{3.3}$$

The flux of each pixel is then calculated using $F = e^{-\tau}$, and the mean flux $\langle F\rangle$ is derived for each value of ζ. We then convert this into a photoionization rate as

$$\Gamma_{\rm bkg}(z) = \frac{\Gamma_{\rm bkg}^{\rm HM01}(z)}{\zeta} \, . \tag{3.4}$$

Consequently, for each redshift we are able to create a one-to-one function between $\langle F\rangle$ and $\Gamma_{\rm bkg}$. These are shown in Fig. 3.1. We use these functions to test the flux-decrement method on our simulated spectra. The scatter in $\Gamma_{\rm bkg}$ from this method should be completely determined by the variation in the mean flux from sightline to sightline, which is primarily due to large-scale density fluctuations.

3.4 The maximum-likelihood technique

Several papers in the literature have used maximum-likelihood estimators with measurements of the strength of the proximity effect to calculate values of $\Gamma_{\rm bkg}$ (Kulkarni and Fall, 1993; Scott et al., 2000, 2002; Faucher-Giguère et al., 2008b; Dall'Aglio and Gnedin, 2010). This work, however, is the first to apply this technique to flux decrements. A likelihood function for each redshift is constructed of the form

$$L_z = \prod_i P[F(i); z, \Gamma_{\rm bkg}] \tag{3.5}$$

where $P[F(i); z, \Gamma_{\rm bkg}]$ is the probability distribution of fluxes for pixel i for a given z and $\Gamma_{\rm bkg}$. The preferred value of $\Gamma_{\rm bkg}$ is the one that maximises L_z. We derive the optical depth distribution at a given redshift directly from the simulation. This PDF is then rescaled for a given UVB with Equation 3.3. Varying $\Gamma_{\rm bkg}$ (and thus ζ) corresponds to a translation of the PDF in $\log(\tau)$, as demonstrated in the left-hand panels of Fig. 3.2. With this underlying optical depth distribution for a given redshift and $\Gamma_{\rm bkg}$, $P(F)$ is calculated using $F = e^{-\tau}$ (middle panels in Fig. 3.2). Each of these histograms have a bin width of 0.01 in flux. Finally, the flux PDFs are convolved with a Gaussian kernel with standard deviation equal to that of the Gaussian-distributed noise in the spectra (right-hand panels in Fig. 3.2). From this final flux PDF, each pixel in a spectrum can be given a probability based upon its observed

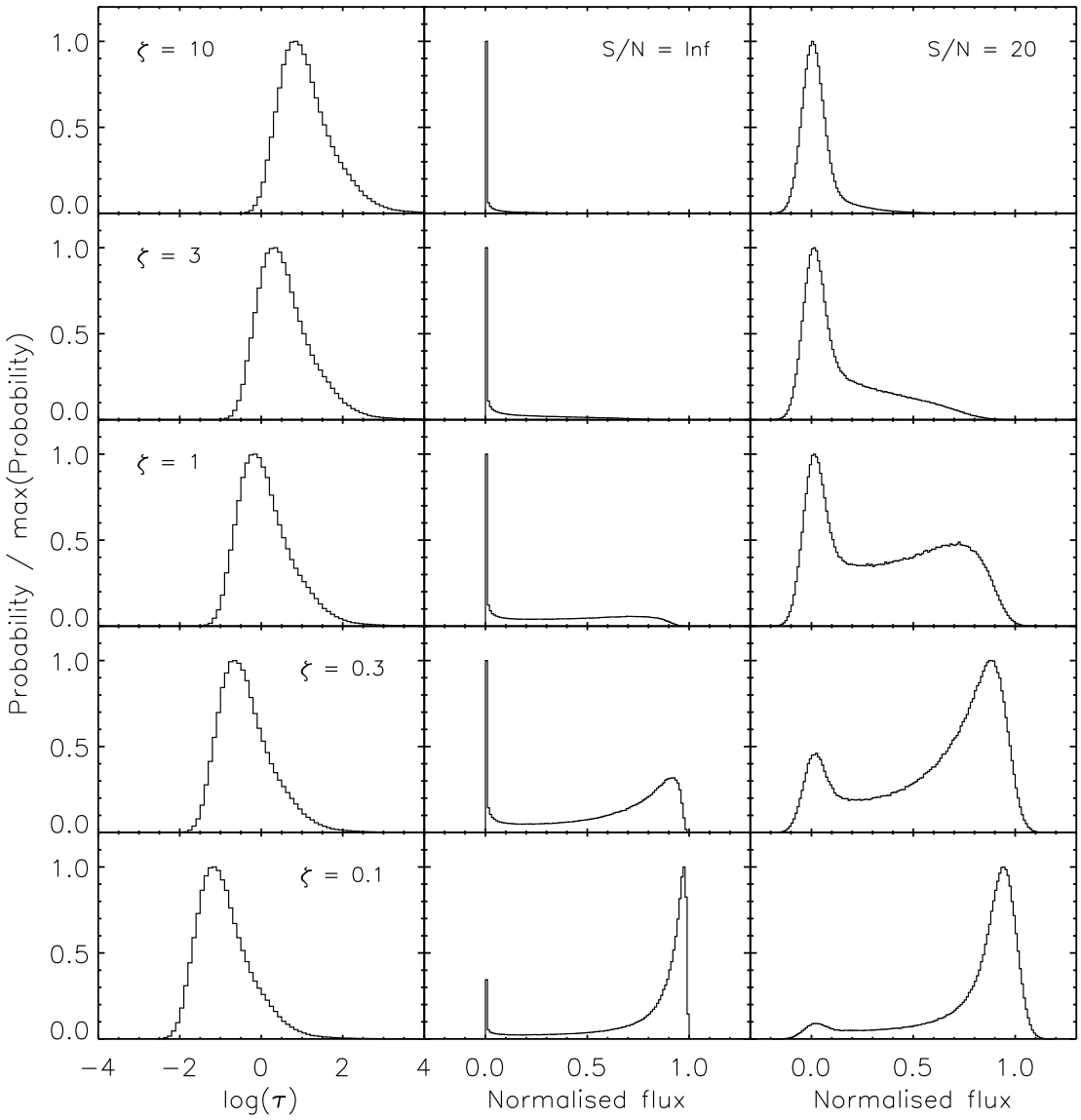


Figure 3.2: *Left-hand panels:* Histograms showing the $\log(\tau)$ distribution from the $z = 4$ simulation. These are scaled using Equation 3.3, where $\zeta = 1$ represents the UVB of Haardt and Madau (2001), whilst larger values represent weaker UVB intensities and smaller values represent stronger UVB intensities. In $\log(\tau)$ this is simply a translation. *Middle panels:* The flux probability histograms derived from each optical depth distribution, equivalent to noiseless spectra. These have a bin width of 0.01 in flux. *Right-hand panels:* Same as the middle panels, but after convolving with a Gaussian kernel with standard deviation 0.05 (i.e. equivalent to a signal-to-noise ratio of 20). For each spectrum, a flux PDF was constructed (for a given redshift and trial $\Gamma_{\rm bkg}$) and then convolved with a Gaussian with standard deviation equal to the standard deviation of the Gaussian-distributed noise. From this histogram each pixel was assigned a probability based upon the bin that contained its observed flux.

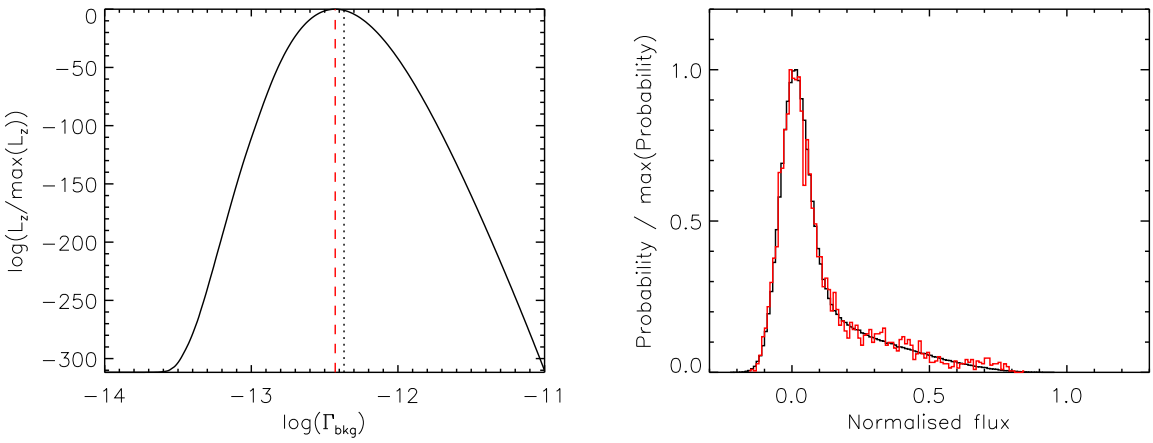


Figure 3.3: *Left-hand panel:* The output likelihood curve for a simulated quasar spectrum at $z = 5$ with HIRES velocity resolution and S/N = 20. For this particular input spectrum the maximum likelihood was for $\log(\Gamma_{\rm bkg}) = -12.428$, denoted by the red dashed line, whilst the black dotted line represents the input UVB. *Right-hand panel:* Comparison of the two flux PDFs. In black is the synthetic PDF generated for the favoured UVB, whilst in red is the flux PDF of the quasar spectrum.

flux. The preferred value of $\Gamma_{\rm bkg}$ maximises the combined probability for all pixels along a sightline. Trial values ranged from $-14 < \Gamma_{\rm bkg} < -11$, and the iteration was done to 0.001 precision in $\Gamma_{\rm bkg}$. An example of the output likelihood curve of a simulated HIRES spectrum at $z = 5$ is shown in Fig. 3.3.

3.5 Comparison of the two methods

The deviation of the estimated value of $\log(\Gamma_{\rm bkg})$ from the input value as determined by both methods over $2 < z < 6$ is shown for the fiducial HM01 UVB in Fig. 3.4, and for a flat UVB with $\Gamma_{\rm bkg} = 10^{-12}$ s^{-1} in Fig. 3.5. In both cases, the returned distributions are very similar. They are approximately Gaussian in $\log\Gamma$ with standard deviations that are roughly the same, particularly when noise is added to the spectra. The estimates from the maximum-likelihood technique are systematically lower in all cases by approximately 0.05 dex. This behaviour of small systematic bias but very similar scatter to estimates from the mean flux was apparent even in spectra with a velocity resolution as low as MIKE, and signal-to-noise ratio (S/N) as low as 10.

There are two issues that must be addressed from this result. First, why does the maximum-likelihood systematically underestimate $\Gamma_{\rm bkg}$? The answer to this lies in the construction of the flux PDF. We derive our optical depth distributions straight from the simulations, and hence they are numerical rather than analytical in nature. Consequently, we choose to bin the optical depths in order to recover a relatively smooth PDF. This approach is carried through in the creation of the flux PDF, where again we bin, using a bin width of 0.01 in flux. We choose this width as a compromise between recovering the subtleties of the shape of the PDF, yet still recovering a relatively smooth function. Unfortunately, the discrete nature of our flux PDF is equivalent to smoothing the input spectra in flux, subtly changing the underlying optical depth distribution (since averaging over fluxes is not equivalent to averaging over optical depths due to their non-linear relation). The flux PDF being used is

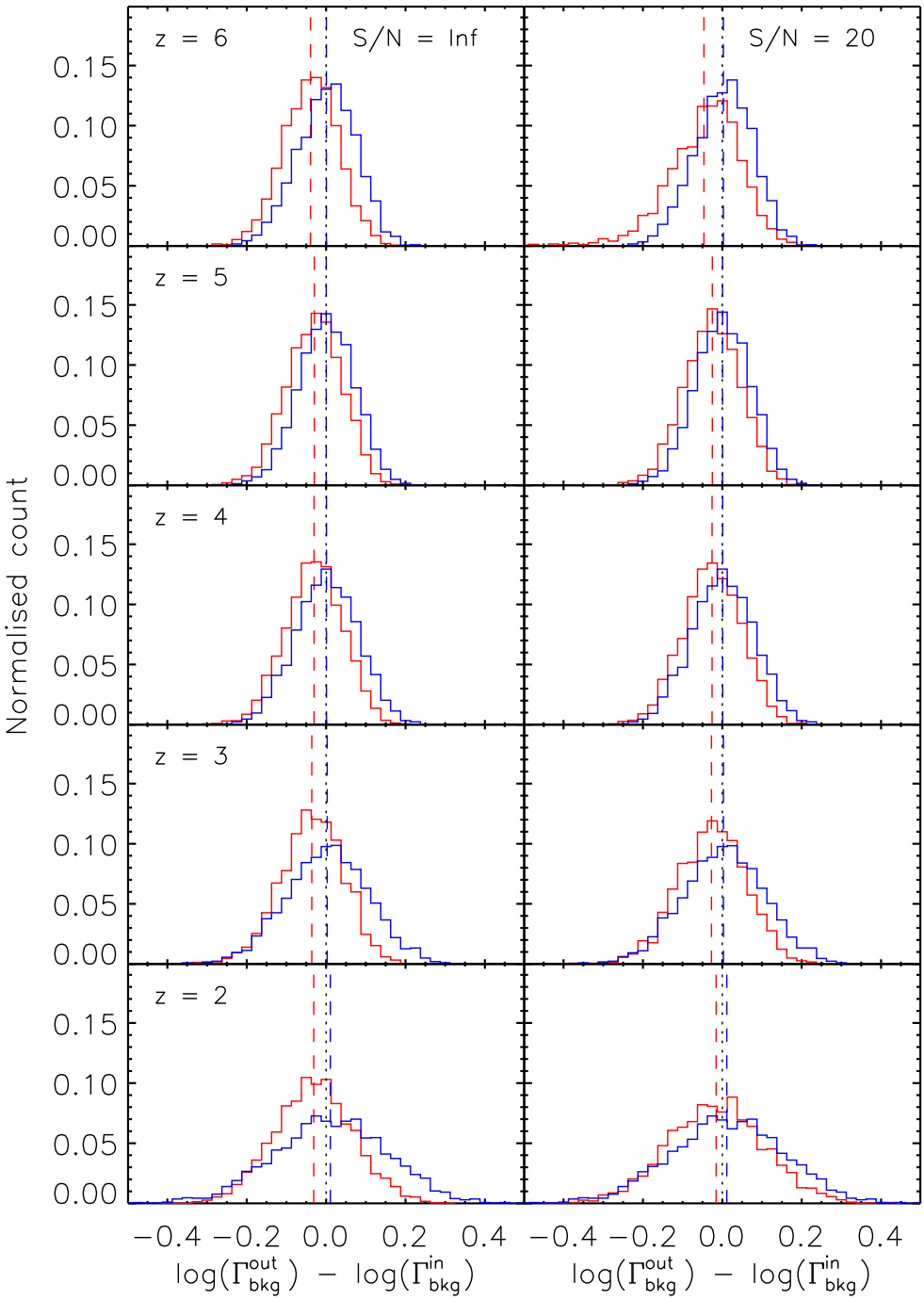


Figure 3.4: Comparison of estimates of $\Gamma_{\rm bkg}$ for 3000 simulated spectra from maximum-likelihood analysis (red histogram), and from mean flux (blue histogram) for a HM01 UVB (i.e. $\zeta = 1$). Both distributions are approximately Gaussian in $\log\Gamma$. The dashed lines are the mean of the distribution, whilst the dotted line marks the input value. *Left-hand panels:* Noiseless spectra at the output velocity resolution of the simulation. The maximum-likelihood technique systematically underestimates $\log(\Gamma_{\rm bkg})$ by ~ 0.05 dex. The scatter in both distributions is very similar, although for $z \lesssim 3$ the scatter is slightly less for maximum-likelihood estimates. *Right-hand panels:* Same as in the left-hand panels, but with Gaussian distributed noise added into the spectra to simulate spectra with a signal-to-noise (S/N) of 20. Again, both methods perform very similarly across all redshifts.

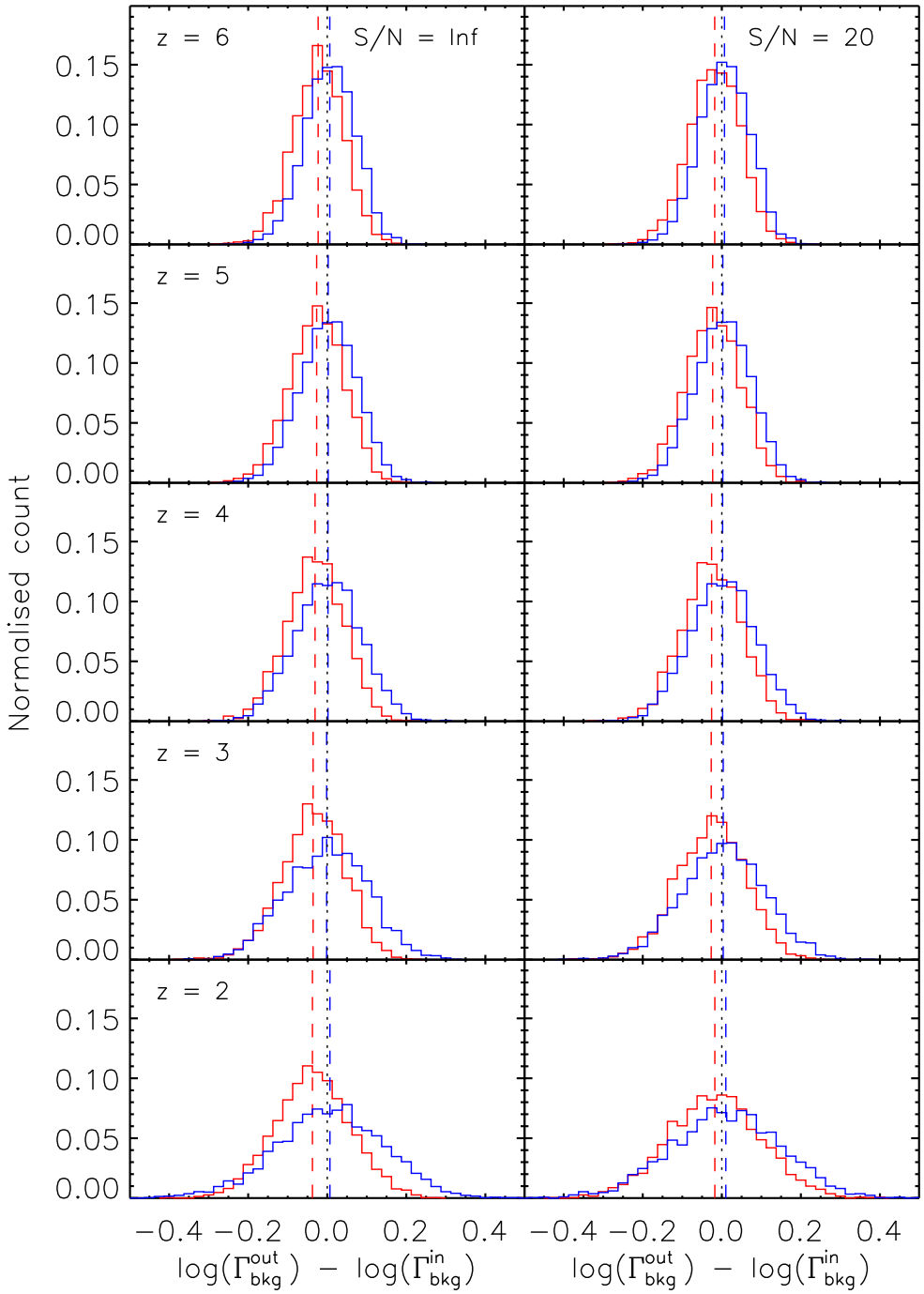


Figure 3.5: The same as Fig. 3.4 but for a constant UVB with $\Gamma_{\rm bkg} = 10^{-12}$ s^{-1} (i.e. ζ is varied with redshift). Despite this rather different input UVB to the previous figure, the distributions are still very similar in scatter.

therefore a slightly different shape to the PDF that it is sensitive to, and this results in a small systematic offset. We tested the sensitivity of this offset to choice of bin width. It was found to decrease with decreasing bin width, and would be zero in the limiting case if the flux PDF being used was described by an analytical function.

The second issue is somewhat less straightforward: why does our maximum-likelihood method, using information from the whole of the flux PDF, return estimates in $\Gamma_{\rm bkg}$ with a similar scatter to those from a method that uses only the mean flux? The scatter in $\Gamma_{\rm bkg}$ from the flux-decrement method comes from both the scatter in the mean flux of a sightline, and the shape of the relationship between the two in Fig. 3.1 within the range of interest. These conspire to produce a roughly Gaussian scatter in the estimated $\log(\Gamma_{\rm bkg})$ across all redshifts tested. Similarly, the rough symmetry within ± 0.4 dex of the peak likelihood in Fig. 3.3 would also suggest a Gaussian scatter in $\log(\Gamma_{\rm bkg})$ using our maximum-likelihood technique. Both these methods lead to distributions that have very similar scatter. Ultimately, this is because the likelihood function is driven by pixels near the mean flux, so it is only the shape of the PDF around the mean flux that matters. When this is roughly flat (as for example at $z = 4$) then this is equivalent to the implicit uniform PDF of the flux-decrement method, and so the scatter is nearly identical despite the rather complicated overall shape of the PDF. In contrast, if the flux PDF varies rapidly around the mean flux, such as when it sharply peaks at $F \sim 0$ ($z = 6$) or $F \sim 1$ ($z = 2$) then the maximum-likelihood technique has a mild advantage and the estimates of $\Gamma_{\rm bkg}$ are slightly more precise. The addition of noise, however, once again flattens the PDF, and so this advantage is suppressed to the point where it is statistically insignificant for spectra with $\mathrm{S/N} \leq 20$.

We highlight that we have only used artificial Lyα forest spectra to test both the flux-decrement and maximum-likelihood methods. Observed spectra will suffer from a variety of complicating issues, such as metal line absorption, continuum fitting difficulties, and assumptions about the underlying density distribution (particularly difficult to model at the low densities relevant at $z \gtrsim 5$; Miralda-Escudé et al., 2000) and $T - \Delta$ relation. All of these would bias the estimates of $\Gamma_{\rm bkg}$ using either method, although the relative sensitivities we leave to future work. We also note that the shape of the flux PDF has been used by various authors to constrain uncertainties in the latter two issues (e.g. Becker et al., 2007; Bolton et al., 2008). What we wish to outline here, though, is that in the limiting case of perfect knowledge about the continuum, density distribution and thermal state of the gas, these two methods will be statistically equivalent for all but the highest S/N spectra. Given the ease with which the mean flux can be computed from an observed spectrum, the classical flux-decrement method is therefore more suitable for estimating $\Gamma_{\rm bkg}$ than the maximum-likelihood method under the restricted conditions tested here.

3.6 Conclusions

The transmitted flux of the Lyα forest is dependent upon several cosmological parameters and the physical state of the gas in the IGM. Many studies in the literature have used the mean transmission to infer the UVB hydrogen photoionization rate, $\Gamma_{\rm bkg}$, by making it the only free parameter in their

simulations and varying it to match the observed transmission (e.g. Rauch et al., 1997). This is known classically as the flux-decrement method.

We have investigated a new variant on this method utilising a maximum-likelihood approach, where each pixel is assigned a probability based upon its observed flux. The distribution of optical depths at a given redshift is derived directly from numerical hydrodynamical simulations. This distribution is then scaled to model different values of $\Gamma_{\rm bkg}$, and the flux PDF is calculated.

Both methods were compared using synthetic spectra drawn from the same simulations, so that any difference in the estimates of $\Gamma_{\rm bkg}$ would be purely due to the methodology. They were found to perform very similarly across a wide range of input UVB intensities, spectral velocity resolutions, and noise properties. The maximum-likelihood approach systematically underestimates $\Gamma_{\rm bkg}$ due to the finite width of the bins in the flux PDF histogram. This width is necessary as the PDF has been derived numerically directly from the simulations, rather than analytically. For a sensible choice of width, however, this bias is small. The scatter in the estimates of $\Gamma_{\rm bkg}$ from both techniques were very similar. This was somewhat surprising since the maximum-likelihood technique uses information about the shape of the entire flux PDF, rather than just the mean flux. We find that this is because our likelihood function is driven by pixels near the mean flux, and so it is the shape of the flux PDF near the mean that is most important. The classical flux-decrement method implicitly assumes a flat flux PDF near the mean, and so our maximum-likelihood technique can only improve upon it when this is a poor approximation for the true shape. This is true when the flux PDF is sharply peaked close to the mean, such as when the majority of pixels are near zero flux or the continuum. However, noise smoothes out these steep gradients such that for spectra with S/N ≤ 20 this advantage is essentially eliminated.

We note that the mean flux of a section of Lyα forest is computationally much easier to calculate than the iteration of our likelihood function. Also, with regard to the observed Lyα forest, both methods will be susceptible to biases from assumptions about modelling the IGM density field, the thermal state of the gas, and placement of the continuum. We leave to future work discussion about the relative sensitivities of each method to errors in these assumptions, however in the case of perfect knowledge about these parameters, both methods give estimates of $\Gamma_{\rm bkg}$ that are statistically equivalent. Therefore, in the situation explored here, we advocate the use of the classical flux-decrement method and conclude that knowledge of the shape of the flux PDF does not directly improve the precision of estimates of $\Gamma_{\rm bkg}$. We emphasise, however, that it may do so indirectly, as it can be used to help constrain the underlying density distribution (e.g. Becker et al., 2007) and thermal state of the gas (e.g. Bolton et al., 2008), both of which are large sources of uncertainty in observational estimates of $\Gamma_{\rm bkg}$ from flux decrements.

Eventually, we reach the utmost limits of our telescopes. There, we measure shadows and search among ghostly errors of measurement for landmarks that are scarcely more substantial.

The Realm of the Nebulae

EDWIN HUBBLE

4

Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect

We present measurements of the ionizing ultraviolet background (UVB) at $z \sim 5-6$ using the quasar proximity effect. The 15 quasars in our sample cover the range $4.6 < z_q < 6.4$, enabling the first proximity-effect measurements of the UVB at $z > 5$. The metagalactic hydrogen ionization rate, Γ_{bkg}, was determined by modelling the combined ionization field from the quasar and the UVB in the proximity zone on a pixel-by-pixel basis. The optical depths in the spectra were corrected for the expected effect of the quasar until the mean flux in the proximity region equalled that in the average Lyα forest, and from this we make a measurement of Γ_{bkg}. A number of systematic effects were tested using synthetic spectra. Noise in the flux was found to be the largest source of bias at $z \sim 5$, while uncertainties in the mean transmitted Lyα flux were responsible for the largest bias at $z \sim 6$. We find a decline in Γ_{bkg} with redshift, from $\log(\Gamma_{bkg}) = -12.15 \pm 0.16$ at $z \sim 5$ to $\log(\Gamma_{bkg}) = -12.84 \pm 0.18$ at $z \sim 6$ (1σ errors). Compared to UVB measurements at lower redshifts, our measurements suggest a drop of a factor of 5 in the H I photoionization rate between $z \sim 4$ and 6. The decline of Γ_{bkg} appears to be gradual, and we find no evidence for a sudden change in the UVB at any redshift that would indicate a rapid change in the attenuation length of ionizing photons. Combined with recent measurements of the evolution of the mean free path of ionizing photons, our results imply a decline in the emissivity of ionizing photons by roughly a factor of 2 from $z \sim 5$ to 6, albeit with significant uncertainty due to the measurement errors in both Γ_{bkg} and the mean free path.

4.1 Introduction

The metagalactic ultraviolet background (UVB) is the net radiation field responsible for keeping the Universe ionized from the end of reionization to the present day. The relative contributions from galaxies and quasars, as well as filtering by the intergalactic medium (IGM) itself, determine the intensity and spectrum of the UVB (e.g. Bechtold et al., 1987; Haardt and Madau, 1996; Fardal et al., 1998; Haardt and Madau, 2001). Thus by measuring the UVB one can hope to place constraints on the evolution of the source population with redshift.

Of particular interest is the evolution of the UVB at $z \sim 6$. The appearance of Gunn-Peterson (GP) troughs in the spectra of the highest redshift known quasars has been interpreted as evidence for a sharp downturn in the UVB at $z > 6$ signalling the end of reionization (e.g. Fan et al., 2006c). However, the diminishing transmitted flux is also consistent with a more slowly evolving UVB and IGM density field (Becker et al., 2007; Bolton and Haehnelt, 2007c).

The bulk of the ionizing photons at $z > 6$ that make up the UVB are believed to come from very low-luminosity galaxies (Richard et al., 2006; Stark et al., 2007; Richard et al., 2008; Oesch et al., 2010; Srbinovsky and Wyithe, 2010). Direct searches for these sources at $z \gtrsim 6-10$ have taken advantage of recent very deep optical and near-infrared imaging from both the ground and space (Bunker et al., 2004; Bouwens et al., 2006; Yoshida et al., 2006; Bouwens et al., 2008; Ouchi et al., 2009). The majority of these faint galaxies, however, still remain below current detection thresholds (e.g. Bouwens et al., 2011b).

The quasar proximity effect has been a classic tool for directly measuring the intensity of the UVB at high redshifts. Since quasars are highly luminous, their output of ionizing photons will dominate over that of the UVB out to large (up to several proper Mpc) distances. This produces a region of enhanced transmission near the redshift of the quasar, first noted by Carswell et al. (1982), known as the 'proximity region'. The size of this region depends both on the quasar luminosity and the intensity of the UVB. For a known quasar luminosity, therefore, the UVB can be estimated by measuring the extent of the proximity zone. Classically, the proximity effect has been measured by comparing column densities of the H I Lyα absorption lines in the forest with those close to the quasar (Murdoch et al., 1986; Tytler, 1987; Carswell et al., 1987; Scott et al., 2000; Bajtlik et al., 1988, hereafter BDO), although a variety of flux statistics have also been used (Liske and Williger, 2001; Dall'Aglio et al., 2008; Dall'Aglio and Gnedin, 2010). Altogether, proximity-effect studies have delivered measurements of the UVB from $z \sim 0.5$ (Kulkarni and Fall, 1993) to $z \sim 4.5$ (Dall'Aglio et al., 2009).

More recently, an alternative method of estimating the UVB has been developed which uses the mean flux in the Lyα forest in combination with numerical simulations. The UVB in the simulation is adjusted until the mean flux in artificial Lyα forest spectra is equal to that in the real data (see e.g., Rauch et al., 1997; Songaila et al., 1999; Tytler et al., 2004; Bolton et al., 2005; Jena et al., 2005; Kirkman et al., 2005). Above $z \sim 4.5$, the UVB has so far only been determined using this type of method (McDonald and Miralda-Escudé, 2001; Meiksin and White, 2004; Bolton and Haehnelt, 2007c; Wyithe and Bolton, 2011). At $z \gtrsim 5$, however, converting the mean flux into an ionization rate depends sensitively on modelling the gas density distribution at very low densities (e.g. Miralda-Escudé et al.,

2000). Not only is this a challenge numerically (Bolton and Becker, 2009), but the optical depth distribution will depend on the properties of the simulation, including the gas temperature.

In this Chapter, we further develop a variant on the 'flux-transmission' method (Liske and Williger, 2001) to measure the UVB intensity at $z > 4.5$ via the proximity effect. This method provide an alternative to identifying individual lines, which becomes increasingly difficult at $z > 4$. At lower redshifts flux statistics have been used to compare the mean flux averaged over extended sections of the spectrum near the quasar redshift to the mean flux in the forest. Rather than computing the mean flux in sections, however, we consider individual pixel optical depths across the proximity region. The optical depths are modified to remove the presumed effect of the quasar, until the proximity region has the same mean flux as the forest at that redshift. The characteristic scalelength of the quasar model is then combined with the quasar luminosity to estimate the UVB. This simple approach avoids a direct dependence on simulations as it does not require the optical depth distribution of the forest to be known a priori. We do, however, use simulations extensively to estimate the bias and uncertainties of our method.

The remainder of this Chapter is organised as follows. In Section 4.2 we describe the observational data as well as the hydrodynamical simulations used. In Section 4.3 we detail the proximity-effect analysis, and the sources of systematic bias. The results are presented and their implications discussed in Section 4.4. Finally, we present our conclusions in Section 4.5. Throughout this Chapter we use the cosmological parameters taken from the mean of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 5-yr data set (Komatsu et al., 2009), with Hubble constant $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and density parameters $(\Omega_\mathrm{m}, \Omega_\Lambda) = (0.26, 0.74)$.

4.2 Data and models

4.2.1 Observed spectra

The quasar spectra used in this Chapter were taken with either the Keck or Magellan telescopes. The majority of the quasars at $z > 5$ were observed with the High Resolution Echelle Spectrometer (HIRES; Vogt et al., 1994) on the 10-m Keck I telescope, and reduced using a custom set of IDL routines and optimal sky-subtraction techniques as detailed in Becker et al. (2006, 2007). All of these observations were made with the 0.86-arcsec slit with $R = 40\,000$, and so the velocity resolution is $6.7\,\mathrm{km\,s^{-1}}$.

The majority of the quasars at $z < 5$ were observed using the 6.5-m Magellan-II Clay and the Magellan Inamori Kyocera Echelle (MIKE) spectrograph (Bernstein et al., 2003) and reduced with a similar custom pipeline. The velocity resolution is roughly half that of the HIRES spectra at $13.6\,\mathrm{km\,s^{-1}}$. A list of the targets is given in Table 4.1.

Quasar redshifts were taken either from the CO and Mg II redshifts presented in Carilli et al. (2010), or from the spectra themselves by identifying the redshift at which the Lyα forest appears to start. Errors on the redshifts measured from the apparent start of the Lyα forest were estimated by comparing to more precise redshifts from Mg II and CO where available, or those in the Sloan Digital Sky Survey (SDSS). All the objects have photometry in the SDSS, and those at $z < 5.5$ also have flux-

Name	z_q	Inst.	Dates	t_{exp} (hrs)	Ref.	S/N
SDSS J1148+5251	6.42	HIRES	Jan 2005 - Feb 2005	14.2[a]	1	16
SDSS J1030+0524	6.31	HIRES	Feb 2005	10.0	1	12
SDSS J1623+3112	6.25	HIRES	Jun 2005	12.5	1	11
SDSS J0818+1722	6.02	HIRES	Feb 2006	8.3	2	12
SDSS J1306+0356	6.02	MIKE	Feb 2007	6.7	4	14
SDSS J0002+2550	5.82	HIRES	Jan 2005 - Jul 2008	14.2	1,4	21
SDSS J0836+0054	5.81	HIRES	Jan 2005	12.5[a]	1	19
SDSS J0231-0728	5.41	HIRES	Jan 2005 - Feb 2005	10.0	1	14
SDSS J1659+2709	5.33	HIRES	Sep 2007 - Jul 2008	11.7	3	32
SDSS J0915+4924	5.20	HIRES	Feb 2005	10.0	1	23
SDSS J1204-0021	5.09	HIRES	Jan 2005 - Feb 2005	6.7	1	17
SDSS J0011+1440	4.97	HIRES	Sep 2007	6.7	3	47
SDSS J2225-0014	4.89	MIKE	Oct 2007	5.0	4	23
SDSS J1616+0501	4.88	MIKE	Mar 2008	3.3	4	21
SDSS J2147-0838	4.59	MIKE	Oct 2007	8.3	3	51

Table 4.1: The list of quasars included in this Chapter. Columns give the quasar name and redshift, details of the observations, and the average signal-to-noise per pixel in the $40h^{-1}$ comoving Mpc closest to the quasar redshift, after masking out skyline residuals (see Section 4.3.3).
[a] The present reductions include only data taken with the upgraded detector.
References: (1) Becker et al. (2006); (2) Becker et al. (2007); (3) Becker et al. (2011a); (4) this work.

calibrated spectra in the SDSS archive. Continuum magnitudes for those at $z > 5.5$ were taken from the discovery papers (Fan et al., 2001, 2003, 2004, 2006c), whilst fluxes for those at $z < 5.5$ were measured from the SDSS spectra. In both cases the continuum flux was measured at a rest wavelength of 1280 Å. The error in the flux measured from the SDSS spectra is conservatively taken to be 0.5×10^{17} erg cm^{-2} s^{-1} Å^{-1}.

All the spectra were normalised following the method described in Bolton et al. (2010). The spectrum is first divided through by a power law $F_\nu \propto \nu^{-0.5}$, normalised at $1280\,(1+z)$ Å, and the Lyα emission line is then fitted with a slowly varying spline. It is difficult to fit the continuum over the forest at these redshifts due to the low flux levels; however, in the proximity region the transmitted flux maxima will be nearer to the continuum, and so is estimated to be within ~ 20 per cent of the correct value over the region of interest.

4.2.2 Simulated spectra

Simulations of the IGM at high redshift were used to test the method and to explore potential sources of systematic error. The simulations are the same as those in Bolton and Becker (2009) and were performed using a customised version of the parallel Tree-SPH code `GADGET-3`, an updated version of the publicly available code `GADGET-2` (Springel, 2005). The simulations assume the cosmological parameters $(h, \Omega_m, \Omega_\Lambda, \Omega_b h^2, \sigma_8) = (0.72, 0.26, 0.74, 0.024, 0.85)$, have both dark matter and gas components and were started at $z = 99$ with initial conditions generated with the transfer function of Eisenstein and Hu (1999). Each simulation uses the UVB model of Haardt and Madau (2001) with contributions from both galaxies and quasars, and is switched on at $z = 9$ and applied in the optically

Model	L $[h^{-1}\,\mathrm{Mpc}]$	Total particle number	M_{gas} $[h^{-1}\,\mathrm{M}_{\odot}]$
A	20	2×100^3	1.03×10^8
B	20	2×200^3	1.29×10^7
C	20	2×400^3	1.61×10^6
D	40	2×200^3	1.03×10^8
E	40	2×400^3	1.29×10^7
F	80	2×400^3	1.03×10^8

Table 4.2: Mass resolution and box size (comoving) of the hydrodynamical simulations used in this work. Model C was primarily used to simulate quasar proximity zones.

thin limit. Our fiducial run is a $20h^{-1}$ (comoving) Mpc box with 2×400^3 particles (Model C) and is the only one used in this Chapter. Simulations with other box sizes and particle numbers were used in Chapter 5 to test the dependence of the environmental bias due to an enhanced average density on the mass of the host halo, and are summarised in Table 4.2.

Simulated Lyα spectra were constructed from line-of-sight density, peculiar velocity, H I fraction and temperature fields. For the main analysis this was done for 1024 random sightlines drawn parallel to the box boundaries (e.g. Theuns et al., 1998), with outputs at $z=(2,3,4,5,6)$. Each sightline is 1024 pixels long. Proximity zones were introduced into the spectra by modifying the neutral H I fraction (in real space) with the ionizing intensity falling off as $1/r^2$, before convolving with the other fields to derive the optical depth (in velocity space). The spectra were convolved with a Gaussian with FWHM equal to the velocity resolution of the instrument that was being modelled (6.7 $\mathrm{km\,s^{-1}}$ for HIRES and 13.6 $\mathrm{km\,s^{-1}}$ for MIKE), before being resampled at the instrument pixel resolution (2.1 and 5.0 $\mathrm{km\,s^{-1}}$, respectively). Gaussian-distributed noise, as well as other imperfections, could then be added. Additionally, sightlines were drawn through the most massive haloes for the analysis described in Section 5.3. The haloes were identified using a friends-of-friends algorithm with a linking length of 0.2. For the analysis in both this Chapter and the next, sightlines of $40h^{-1}$ comoving Mpc were used. For Models A-C several random sightlines were combined for this.

4.3 Analysis

4.3.1 Proximity-effect formalism

As mentioned in Section 4.1, the ionizing flux from a quasar will locally dominate over the UVB in setting the ionization state of the IGM. This leads to increased transmission near the quasar compared to the transmission in the forest (the 'proximity effect'). BDO first translated this increased transmission (through a reduction in the number of strong absorption troughs) into a measure of the photoionization rate of hydrogen caused by the UVB, Γ_{bkg}. More recently, measurements have been made of the proximity effect using flux transmission statistics rather than line-counting as done by BDO. The average transmission measured close to the quasar is thereby compared to that of the average Lyα forest (e.g. Liske and Williger, 2001).

The optical depth, τ, is related to the normalised flux, $F=e^{-\tau}$. In the simplest model where the

quasar lies in a typical region of the IGM, and neglecting all motion of the gas and any temperature gradients that may exist as one approaches the quasar, the optical depth at any point can be described as

$$\tau = \tau_{\rm forest}[1+\omega(r)]^{-1}\,, \tag{4.1}$$

where $\tau_{\rm forest}$ is the optical depth that would be measured in the absence of the quasar (i.e. the typical value for the forest at that redshift), and

$$\omega(r) = \frac{\Gamma_{\rm q}(r)}{\Gamma_{\rm bkg}}\,. \tag{4.2}$$

Here, $\Gamma_{\rm q}(r)$ is the H I photoionization rate of the quasar at proper distance r and $\Gamma_{\rm bkg}$ is that of the UVB, assumed to be spatially uniform for a given redshift. The distance from the quasar, r, is approximately

$$r \simeq \frac{c}{H(z)}\frac{\Delta z}{1+z}\,. \tag{4.3}$$

We further define a characteristic length $R_{\rm eq}$ to be the distance from the quasar where the photoionization rate from the UVB equals that from the quasar, i.e. $\Gamma_{\rm q}(R_{\rm eq}) = \Gamma_{\rm bkg}$. With this definition, Equation 4.1 can then be rewritten as

$$\tau = \tau_{\rm forest}\left[1+\left(\frac{r}{R_{\rm eq}}\right)^{-2}\right]^{-1}\,. \tag{4.4}$$

It must be emphasised that $R_{\rm eq}$ is different to proximity region sizes as defined in the literature. At high redshift the proximity region 'size' is typically defined to be the maximum extent of the enhanced transmitted flux (e.g. the first point at which the transmission drops to 0.1 in the spectrum when smoothed with a 20 Å filter, Fan et al., 2006b; Carilli et al., 2010), but this is an observational, rather than a physical quantity. In this Chapter we choose to define the proximity region size as the scalelength out to which the ionizing flux from the quasar dominates over that from the background, i.e. where $\omega = 1$. As such any comparison between our proximity region sizes (values of $R_{\rm eq}$) and those found by other methods for the same quasars should be made keeping this difference in mind. For further discussion on the distribution of proximity region sizes at high redshift see Maselli et al. (2009) and Bolton and Haehnelt (2007a,b).

A measurement of the UVB intensity can be expressed as a value of $R_{\rm eq}$ if the flux of ionizing photons from the quasar and its fall-off with distance are known. We assume a $1/r^2$ fall-off. For this the flux at the Lyman limit, f_{ν_0}, needs to be determined. At the redshifts of the quasars discussed in this Chapter insufficient flux is transmitted through the forest for direct measurement of f_{ν_0}, and so we extrapolate the continuum flux (measured at a rest wavelength of 1280 Å as described in Section 4.2.1) by assuming a power-law relation of the form $f_\nu \sim \nu^{-\alpha}$. The value of α used in this Chapter is 1.61 ± 0.86 and is based on the $z < 2$ radio-quiet quasar sample of Telfer et al. (2002), who measured this index in the range 500 to 1200 Å for 39 individual AGN. The error quoted is the rms scatter of that sample. By contrast, Scott et al. (2004) found a harder index ($\alpha = 0.74$) in their $z < 0.67$ sample. The Scott et al. (2004) sample, however, covers a luminosity range that is an order of magnitude lower than either the Telfer et al. (2002) sample, or the quasars analysed here. The Telfer et al. (2002) mean value

and dispersion should therefore be the most appropriate for this study. Using f_{ν_0}, the luminosity of the quasar at the Lyman limit, L_{ν_0}, is calculated as

$$L_{\nu_0} = 4\pi d_{\rm L}^2 \frac{f_{\nu_0}}{(1+z_{\rm q})} \,, \tag{4.5}$$

where $d_{\rm L}$ is the luminosity distance to the quasar. The quoted errors in L_{ν_0} take into account the error on α, the error in $d_{\rm L}$ (from the error in the redshift), and the error in the apparent magnitude of the continuum, m_{1280}. For quasars with $z_{\rm q} < 5.5$ the error in m_{1280} is calculated from the error in the measured continuum flux for SDSS spectra as described in Section 4.2.1, whilst for those with $z > 5.5$ it is taken to be the same as the error on the SDSS z-band photometry. For a given distance from the quasar, r, the Lyman limit flux density is

$$F^{\rm Q}_{\nu_0}(r) = \frac{L_{\nu_0}}{4\pi r^2} \,. \tag{4.6}$$

The photoionization rate of H I (in units of $\rm s^{-1}$) by a source of UV flux is given by

$$\Gamma = \int_{\nu_0}^{\infty} \frac{4\pi J(\nu)\sigma_{\rm HI}(\nu)}{{\rm h}\nu} d\nu \,, \tag{4.7}$$

where $J(\nu)$ is the intensity of the source, $\sigma_{\rm HI}(\nu)$ is the ionization cross-section of neutral hydrogen, and h is in this case Planck's constant. By definition $\Gamma_{\rm bkg} = \Gamma_{\rm q}(R_{\rm eq})$, and so using $\sigma_{\rm HI}(\nu) = 6.3\times10^{-18}(\nu_0/\nu)^{2.75}$ $\rm cm^2$ (Kirkman and Tytler, 2008, note that an exponent of 3 is often used in similar analyses) and integrating Equation 4.7 for the photoionization rate by the background gives

$$\Gamma_{\rm bkg} = \frac{9.5\times10^8 F^{\rm Q}_{\nu_0}(R_{\rm eq})}{(\alpha+2.75)} \tag{4.8}$$

in units $\rm s^{-1}$, where $F^{\rm Q}_{\nu_0}(R_{\rm eq})$ is the Lyman limit flux density in erg $\rm cm^{-2}$ evaluated at a distance $R_{\rm eq}$ away from the quasar. Therefore, using Equation 4.6, $\Gamma_{\rm bkg}$ can be expressed as

$$\Gamma_{\rm bkg} = \frac{9.5\times10^8 L_{\nu_0}}{(\alpha+2.75)4\pi R^2_{\rm eq}} \,, \tag{4.9}$$

where L_{ν_0} is in erg $\rm s^{-1}$ $\rm Hz^{-1}$ and $R_{\rm eq}$ is in cm.

4.3.2 Measurement method

The value of $\Gamma_{\rm bkg}$ was inferred for each quasar by increasing the optical depths of each pixel by the expected effect of the quasar, using Equation 4.4, until the mean flux across the spectra was the same as that expected from the Lyα forest at that redshift. Details of this procedure are given below.

The analysis was carried out on the section of the spectrum immediately bluewards of the Lyα line and $40h^{-1}$ comoving Mpc ($\sim 8-9$ proper Mpc, or $\Delta z \sim 0.10-0.15$) in length. Each pixel is converted from a normalised flux into an optical depth using $\tau = -\ln(F)$. For a given quasar luminosity and trial

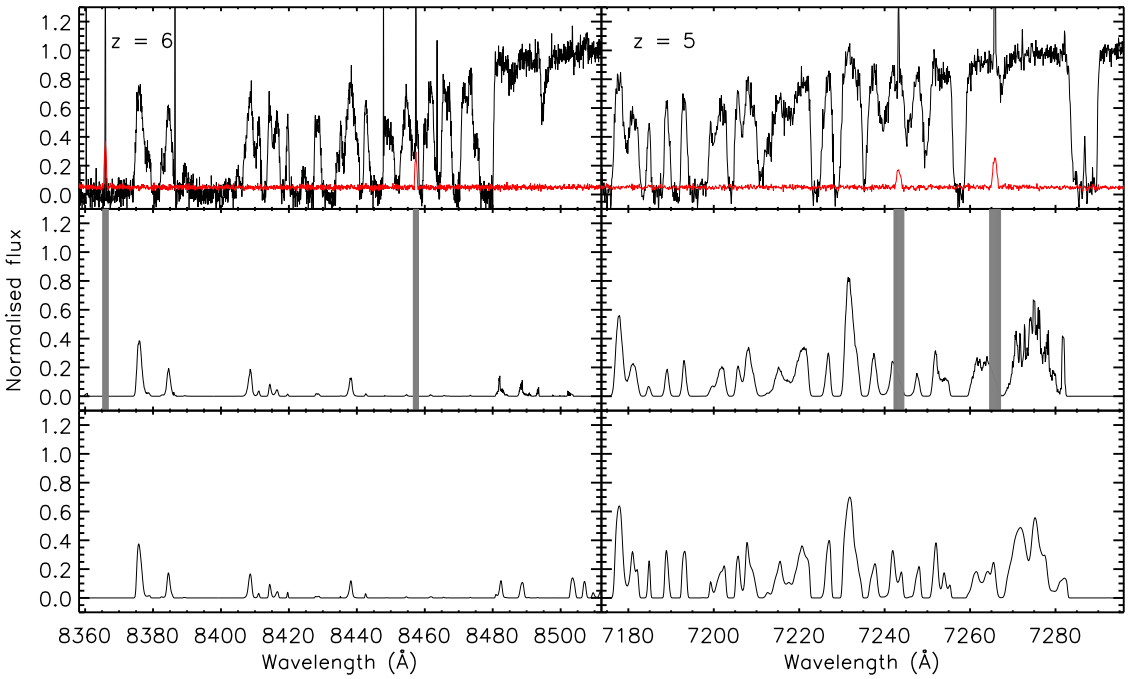


Figure 4.1: *Top panels:* Simulated sightlines at $z = 6$ (left) and $z = 5$ (right). Each sightline is $40h^{-1}$ comoving Mpc long, with $R_{\rm eq}$ = 10 proper Mpc and S/N = 20. The resolution for the $z = 6$ spectrum is equal to that of HIRES, whilst the spectrum at $z = 5$ has resolution equal to that of MIKE. The red line is the simulated error array. Artefacts such as bad pixels and skyline residuals have also been introduced. Lyα at the redshift of the quasar is on the far right-hand side of each panel. *Middle panels:* The spectra smoothed according to their noise properties (see Section 4.3.3), and with optical depths altered such that $\Delta F = 0$ (see Equation 4.10). The grey regions mark parts of the spectra that were masked out automatically (primarily skyline residuals). *Bottom panels:* The original Lyα forest for these sightlines. The similarity between these spectra and those in the middle panels provides a good check that the method is working correctly.

value of $\Gamma_{\rm bkg}$, Equation 4.9 can be inverted to give a trial value of $R_{\rm eq}$. Using this trial value of $R_{\rm eq}$, the factor $[1+\omega(r)]^{-1}$ can be calculated for every pixel in the spectrum. The pixel optical depths are then adjusted, before being converted back into fluxes. This process is continued until the section of the spectrum resembles the Lyα forest at the same redshift. The most robust statistic for determining this was found to be the difference between the mean flux in the spectrum and the mean flux of the Lyα forest, $F_{\rm forest}$. The difference between the two, ΔF, was defined as

$$\Delta F = \langle F_{\rm new} \rangle - F_{\rm forest} \, , \tag{4.10}$$

where $F_{\rm new} = e^{-\tau_{\rm new}}$ and $\tau_{\rm new} = \tau[1+\omega(r)]$. The value for which $\Delta F = 0$ we took as our estimate of $\Gamma_{\rm bkg}$. The mean flux, $\langle F_{\rm new} \rangle$, was calculated as a weighted mean, with each pixel weighted by the inverse variance of the flux. For a given original error in the transmission, σ_F, then $\sigma_{F_{\rm new}} = \sigma_F[1+\omega(r)]^{-1} F_{\rm new}/F$. The weighted mean flux was then calculated as $\langle F_{\rm new} \rangle = \sum (F_{\rm new}/\sigma^2_{F_{\rm new}}) / \sum (1/\sigma^2_{F_{\rm new}})$. In order to stop $\sigma_{F_{\rm new}}$ from reaching zero and strongly biasing the weighting, a noise 'floor' was used such that $\sigma_{F_{\rm new}} \geq 0.01$.

Trial values of $\log(\Gamma_{\rm bkg})$ ranged within $-14 < \log(\Gamma_{\rm bkg}) < -11$ and the iteration was done until a

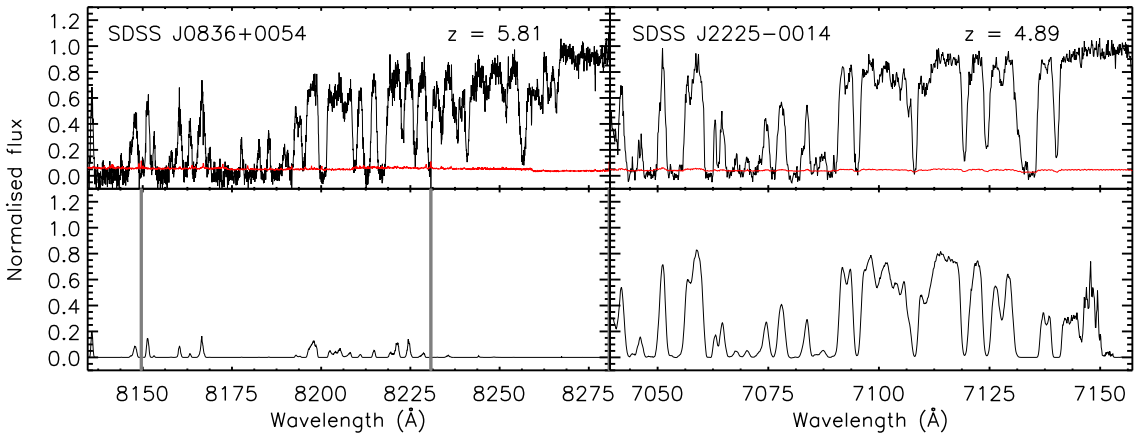


Figure 4.2: Same as Fig. 4.1, but for a small sample of the observed spectra. *Top panels:* The normalised spectrum of SDSS J0836+0054 at $z = 5.810$ taken with the HIRES instrument (left), and the normalised spectrum of SDSS J2225-0014 at $z = 4.886$ taken with MIKE (right). The length of each spectrum corresponds to $40h^{-1}$ comoving Mpc, and the red line is the observed error array. Lyα at the redshift of the quasar is on the far right-hand side of each panel. *Bottom panels:* The smoothed spectra with optical depths modified assuming an $R_{\rm eq}$ of the size quoted in Table 4.4. The grey regions mark parts of the spectra that were masked out automatically.

precision of 0.001 in $\log(\Gamma_{\rm bkg})$ was reached. Fig. 4.1 shows the effect of the result of this iteration on simulated sightlines at $z = 5$ and $z = 6$. The 'recovered' Lyα forest (middle panels) is strikingly similar to the actual Lyα forest in those sightlines (bottom panels). Two observed spectra are presented in a similar fashion in Fig. 4.2. The bottom panels show the spectra after smoothing to mitigate the effect of pixel noise (see Section 4.3.3) and after the optical depths have been modified using the estimated value of $R_{\rm eq}$ presented in Section 4.4.1.

The assumed value of $F_{\rm forest}$ is calculated at each redshift from the fit to measured optical depths of the Lyα forest in Becker et al. (2007). The fit is reproduced along with some observed values of $\tau_{\rm eff}$ from the literature in Fig. 4.3. It assumes a simple evolution of a lognormal distribution of optical depths and matches the observed values excellently from $2 < z < 6$. Above $z = 6$ there are few measurements of the flux in the forest and so it is unknown if this relationship still holds. For the $z > 6$ quasars in this Chapter it was assumed that it does, although for large values of effective optical depth ($\tau > 6$), a substantial change in τ corresponds to only a very small (absolute) change in $F_{\rm forest}$. Only a small change in $R_{\rm eq}$, therefore, is required for ΔF to again be zero. This is apparent in Fig. 4.4, which shows an example of the results for one of the simulated spectra at $z = 5$ and 6. If the mean flux is overestimated by a factor of 5, for example (corresponding to the dotted line being down at -0.10 and -0.010 in the top and bottom panels, respectively), then the shift in the estimated $\log(\Gamma_{\rm bkg})$ is much greater at $z = 5$ (~ 0.6 dex) than at $z = 6$ (~ 0.3 dex). The results from the $z > 6$ quasars should therefore be rather robust to significant uncertainties in the mean flux.

Probable skylines are identified as regions with $\sigma_F > 2\,\langle\sigma_F\rangle$, where the mean is defined over the whole spectrum. Those pixels, plus 5 pixels either side, are masked out. The data is also smoothed with a boxcar with a smoothing window that is proportional to the amount of noise in the data (window width $= 10[\langle\sigma_F\rangle/0.05]$ pixels), in order to balance between a smoothly varying function whilst

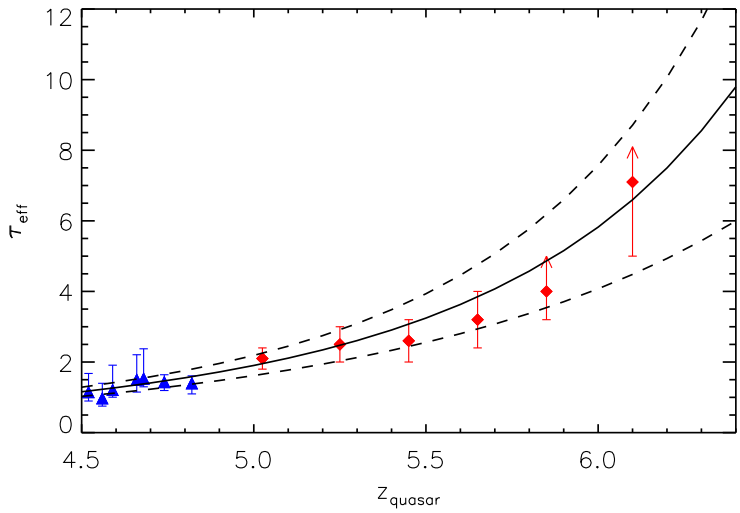


Figure 4.3: Our adopted fit to the evolution of $\tau_{\rm eff}$ with redshift. The blue triangles are from measures of the flux in the forest from Songaila (2004), whilst the red diamonds are the binned values from Fan et al. (2006b). The solid line gives the Becker et al. (2007) relation for the evolution of $\tau_{\rm eff}$ with redshift, which we adopt for this study. The 1σ error in $\tau_{\rm eff}$ is marked with the dashed lines and was taken to be 10% at $z = 4$, 15% at $z = 5$ and 30% at $z = 6$, with the error at other redshifts calculated by quadratic extrapolation.

maintaining some of the contrast and resolution in the proximity region. Regions with a smoothed flux below zero were treated as though they were positive, but remembering the sign, such that $F_{\rm new} = -e^{-(-\ln|F|)[1+\omega(r)]}$. Smoothed fluxes that were greater than or equal to 1 (i.e. at the continuum) were set to be equal to 0.99. These effects, designed to limit the effects of observational artefacts, introduce a bias such that as the noise increases the method will systematically overestimate the UVB. These biases were therefore extensively modelled using the simulated spectra (see Appendix A).

There are several advantages of our method over other recently presented methods to measure the intensity of the UVB from the proximity effect. One of the key attributes of our method is that the intrinsic distribution of Lyα optical depths need not be known a priori. This means that we are not directly dependent on numerical simulations. Another key advantage is that a slope, β, of the column density distribution of neutral hydrogen is not assumed, as was required in e.g. Dall'Aglio et al. (2008). We believe that with our method the largest remaining uncertainty is the evolution of $\tau_{\rm eff}$ at $z > 6$.

4.3.3 Statistical accuracy of the method

In order to test the accuracy of the method, simulated spectra were created using random sightlines through the simulations. The test spectra were all created using Model C at $z = 5$ and 6, and were $40h^{-1}$ (comoving) Mpc long, constructed by joining together multiple sightlines. In our standard mock spectra the proximity region was assumed to have a size of 10 proper Mpc. This is in general agreement with the derived values from the data (see Section 4.4.1) and that estimated by Wyithe and Loeb (2004). Assuming a Haardt and Madau (2001) UVB this corresponds to a quasar with $M_{1450} = -27.10$ (AB) at $z = 6$ and $M_{1450} = -27.47$ (AB) at $z = 5$. The noise was assumed to be Gaussian with a signal-to-noise (S/N) of 20, close to the average value of the S/N in the data.

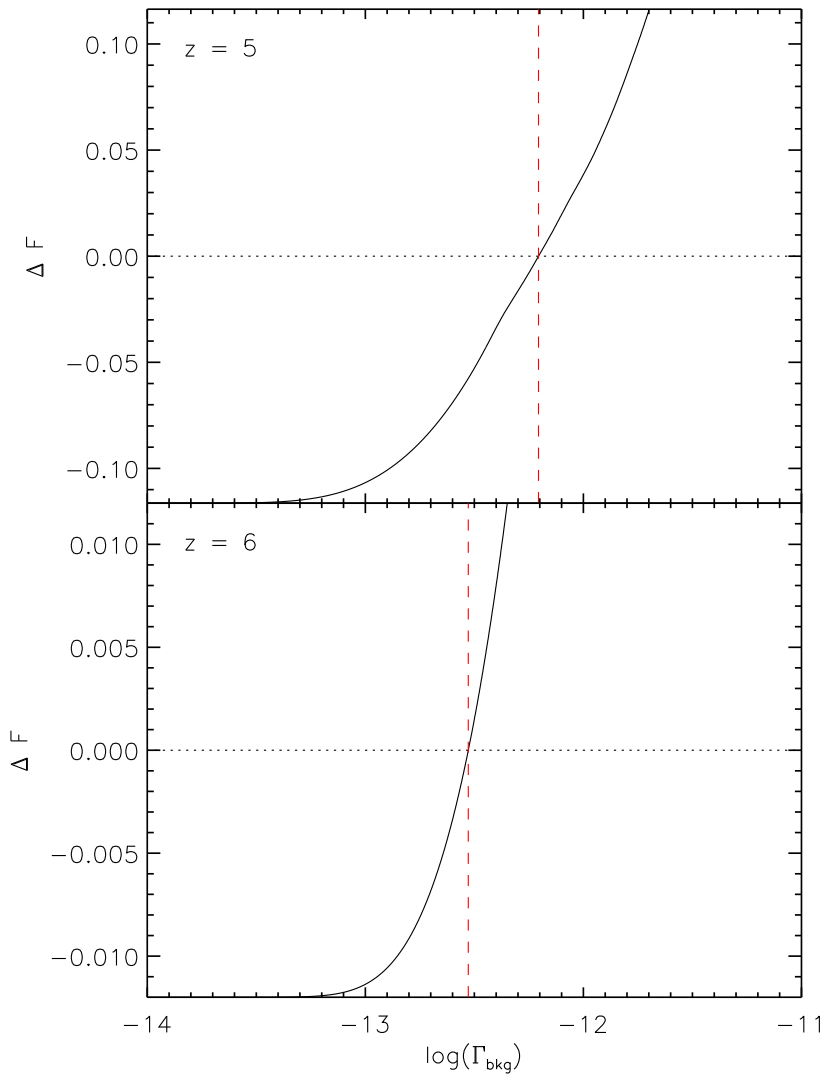


Figure 4.4: *Top panel:* The variation of ΔF with $\log(\Gamma_{\rm bkg})$ of a simulated quasar spectrum at $z = 5$. The dotted line marks $\Delta F = 0$ and the dashed red line shows the estimated value of $\log(\Gamma_{\rm bkg})$. For this particular input spectrum $\Delta F = 0$ at $\log(\Gamma_{\rm bkg}) = -12.206$. The curve takes an asymptotic shape as the mean flux becomes zero. *Bottom panel:* Same as the top but for a simulated quasar at $z = 6$. The curve is steeper, and so if the mean flux has been overestimated by some factor, this will result in a smaller change in the estimated $\log(\Gamma_{\rm bkg})$ at $z = 6$ than at $z = 5$.

In order to better mimic the real data, a variety of artefacts were introduced into the simulated spectra. Sky line residuals were added with an associated peak in the error spectrum, as well as bad pixels for which there was no associated effect on the error spectrum. Overall the simulated spectra are qualitatively very similar to the real data, as demonstrated in Fig. 4.5.

We applied our analysis to artificial spectra with similar S/N to the data. On average the method estimated the correct value of $\Gamma_{\rm bkg}$, with the distribution of values approximately Gaussian in $\log(\Gamma_{\rm bkg})$. Several factors were thoroughly investigated for their effect on the size of the errors, and any bias on the estimated UVB intensity. We investigated the effects of peculiar velocities in the gas, $v_{\rm pec}$, which will introduce distortions in redshift space to the expected radial profile of the transmitted flux. This

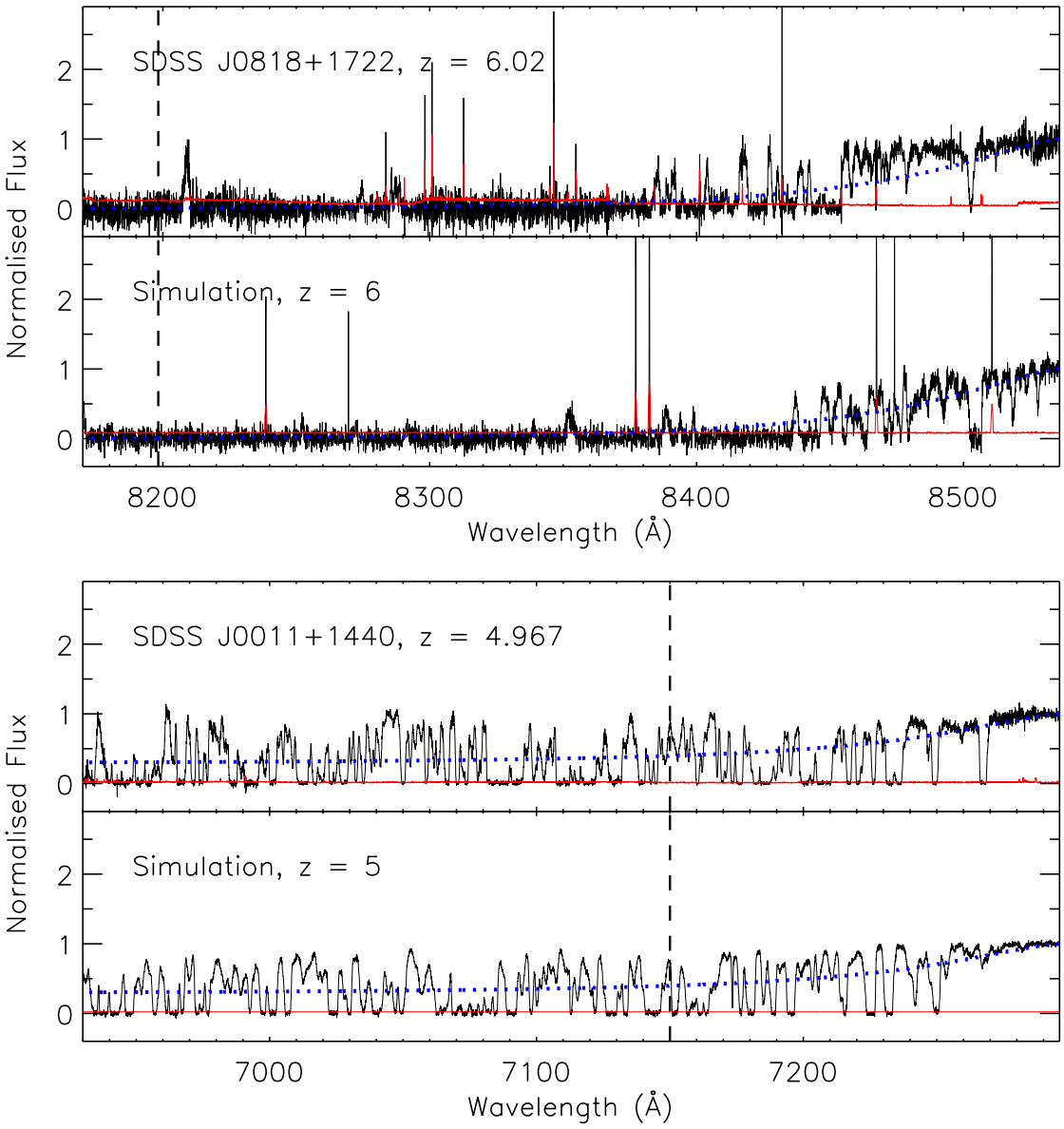


Figure 4.5: *Top:* The upper panel shows the normalised spectrum of SDSS J0818+1722 at $z = 6.02$, whilst the lower panel is a simulated spectrum at $z = 6$ with the same $R_{\rm eq}$ and $\Gamma_{\rm bkg}$ as presented in Table 4.4, as well as similar noise properties. The red solid line is the normalised error spectrum, the black dashed line is the derived value of $R_{\rm eq}$, and the dotted blue line is the expected drop-off in the mean normalised flux. Lyα at the redshift of the quasar is on the far right-hand side of each plot, and the simulated spectrum has been shifted so that it is at the same redshift as the observed spectrum. *Bottom:* The same as above but for SDSS J0011+1440 at $z = 4.967$ and the $z = 5$ simulation. Again there is a strong similarity between the two spectra. The value of $R_{\rm eq}$ is smaller at $z = 5$ than at $z = 6$ because the intensity of the UVB is higher, and so the ionizing flux from the quasar dominates over the background out to a shorter distance.

	$z = 5$						$z = 6$					
	$R_{\rm eq}$ = 10 Mpc			$R_{\rm eq}$ = 5 Mpc			$R_{\rm eq}$ = 10 Mpc			$R_{\rm eq}$ = 5 Mpc		
Property	Value	σ_{x_i}	ϵ_{x_i}	Value	σ_{x_i}	ϵ_{x_i}	Value	σ_{x_i}	ϵ_{x_i}	Value	σ_{x_i}	ϵ_{x_i}
$\log L_{\nu_0}$	31.37	0.07	0.01	30.77	0.09	0.01	31.22	0.07	0.01	30.62	0.09	0.01
$v_{\rm pec}$	-	0.09	0.05		0.12	0.05	-	0.04	-0.02		0.07	0.00
Sightline δF	21%	0.13	0.02		0.14	0.01	43%	0.10	0.02		0.13	0.02
$\Delta\tau_{\rm eff}$	15%	0.22	0.03		0.23	0.01	30%	0.45	0.10		0.59	0.11
Δz	0.005	0.18	-		0.31	-	0.01	0.20	-		0.39	-
ΔL_{ν_0}	30%	0.14	-		0.14	-	30%	0.15	-		0.14	-
S/N	20	0.13	0.16		0.11	0.29	20	0.03	0.04		0.09	0.09
Inst. res.	13.6 km s^{-1}	0.03	0.05		0.04	0.03	6.7 km s^{-1}	0.02	0.01		0.04	0.00
Halo host mass	10^{13}M$_\odot$	0.10	0.12		0.16	0.33	10^{13}M$_\odot$	0.05	0.03		0.09	0.09
Thermal prox. effect	10^4 K	0.07	-0.10		0.01	-0.17	10^4 K	0.02	-0.11		0.09	-0.19
Total		0.41	0.33		0.51	0.57		0.54	0.08		0.76	0.13
Model		0.43	0.31		0.56	0.57		0.54	0.09		0.86	0.18

Table 4.3: The contribution to the error and systematic shift in $\log(\Gamma_{\rm bkg})$ from a variety of causes estimated from 1000 simulated spectra at $z = 5$ and 6, each with either $R_{\rm eq}$ = 10 or 5 proper Mpc. The errors from specified properties were found to have a nearly Gaussian distribution in $\log(\Gamma_{\rm bkg})$. Each property, x_i, contributes an error σ_{x_i} and a systematic shift, ϵ_{x_i} (both measured in dex). The two instrumental resolutions represent those of MIKE and HIRES at $z = 5$ and 6, respectively. The bottom two sources of error are discussed in Chapter 5, and are included here for comparison. To simulate a thermal proximity effect, the temperature of the gas in the closest 5 proper Mpc has been raised by 10^4 K, similar to if the quasar had ionized HeII (see Section 5.5). The total error is calculated by adding the individual errors in quadrature, whilst the total shift is the sum of the individual shifts. For comparison, we also simulated spectra with all possible sources of error included simultaneously ('model'). The total error, as well as the total systematic shift, are similar to those expected from combining the individual effects.

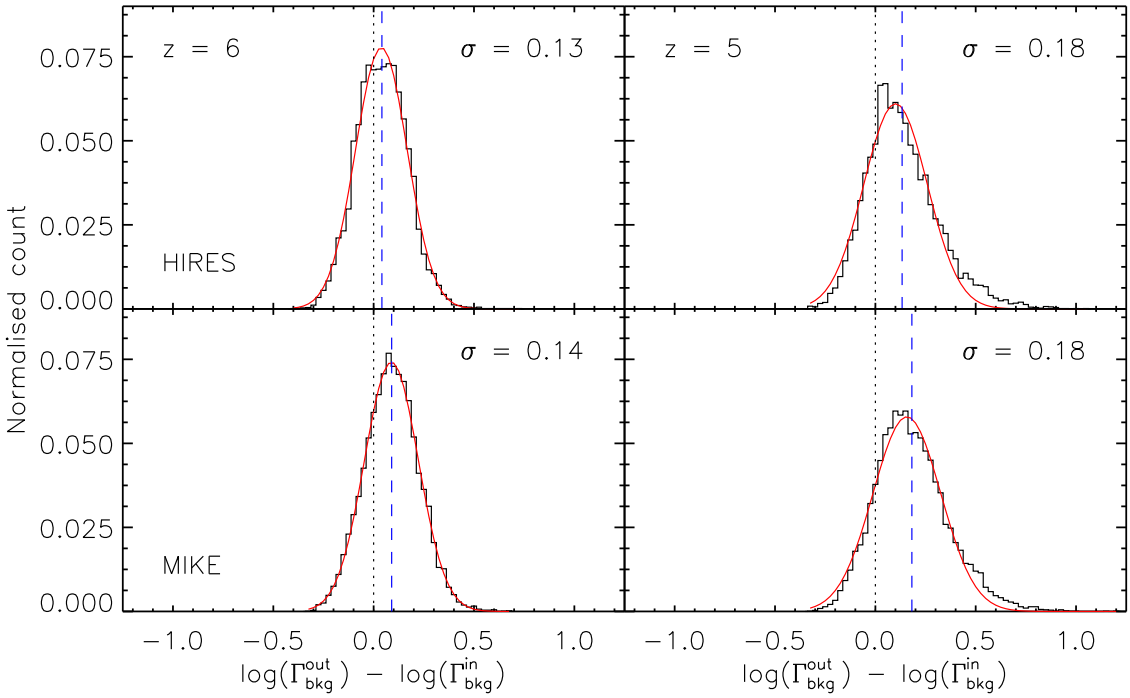


Figure 4.6: The deviation of the estimated value of $\log(\Gamma_{\rm bkg})$ from the input value for 10 000 simulated spectra at $z = 6$ (left) and $z = 5$ (right) with input proximity regions of 10 proper Mpc in size. The spectra had a fixed S/N = 20, and the resolution of either the HIRES (top) or MIKE (bottom) spectrographs. The dotted black line marks the input value, whilst the blue dashed line marks the mean of the simulated data set. The spread is approximately Gaussian in $\log(\Gamma)$ with a small systematic offset caused by noise in the spectra. Each offset in the estimated $\Gamma_{\rm bkg}$ corresponds to an overestimate of the input value. The width of the Gaussian and size of the systematic offset varied with redshift, S/N, luminosity and the resolution. Detailed modelling of the effect of each of these parameters was carried out to calculate the statistical error and bias in the estimated $\Gamma_{\rm bkg}$ for both individual quasars and a grouped sample. A detailed error budget is presented in Table 4.3, and related figures are presented in Appendix A.

was done by rerunning our analysis on spectra generated with peculiar velocities set to zero. Similarly we examined the impact of sightline-to-sightline variations in the mean flux away from the global mean flux due to local fluctuations in the density field. For this we ran the analysis on spectra from which the mean flux of that sightline was known, and compared it to when simply assuming a global mean flux. The spread was also dependent on the noise levels in the simulated spectra, the redshift, the instrumental resolution, the luminosity, and the error in $\tau_{\rm forest}$ (which varied with redshift). This analysis is shown in more detail in Appendix A. Each contributor to the error (see Table 4.3) was varied between reasonable limits for 1000 simulated spectra at both $z = 5$ and 6, and linearly extrapolated to other redshifts. This meant that the contribution to the error budget and associated bias from all the variables could be calculated for the parameter space and redshift range present in the data. Using this analysis an overall statistical error and bias were determined for each individual quasar. For demonstration, the dependence on S/N and instrumental resolution is shown in Fig. 4.6. As expected, if the S/N decreases then the scatter in $\log(\Gamma_{\rm bkg})$ increases. There is also a systematic bias, due to the smoothing of the spectra. Averaging over fluxes is not equivalent to averaging over optical depths,

due to the non-linear relation between the two, and so the more the spectrum is smoothed the more the optical depths are underestimated. This effect is strongest for low fluxes (high optical depths). Consequently, smoothing (particularly at high redshift) increases the transmission in the forest so the fall-off in mean flux is not as steep. This corresponds to a smaller $R_{\rm eq}$, and so ultimately $\log(\Gamma_{\rm bkg})$ is overestimated. Since MIKE has a lower instrumental resolution than HIRES there is an additional bias brought in due to a similar effect. Indeed any effect that smoothes the spectra by averaging over pixel fluxes will bring in a bias of this nature. Table 4.3 breaks down the sources of scatter and systematic biases based on 1000 sightlines that have similar values to those in the data at $z = 5$ and 6. At both redshifts the errors are dominated by the error in $\tau_{\rm eff}$, and the errors in redshift are far more important than the errors in the luminosity. Even though the error in $\tau_{\rm eff}$ is symmetric (it is approximated as a Gaussian), the resulting bias is asymmetric, and becomes more so at high $\tau_{\rm eff}$ (i.e. high redshifts), as the equivalent distribution of $F_{\rm forest}$ also becomes highly asymmetric.

As we noted in Section 4.3.1 we assume a spatially uniform UVB. However, shortly after reionization has completed the mean free path of ionizing photons is short enough that significant spatial variations in the UVB may exist (e.g Lidz et al., 2007). Mesinger and Furlanetto (2009) show, however, that variations in the density field at $z \sim 5-6$ dominate over spatial variation in the UVB, and that even just after reionization assuming a uniform UVB underestimates $\Gamma_{\rm bkg}$ by at most a few percent. We thus do not try to correct for this uncertain but small effect.

We also performed a joint analysis of several sightlines simultaneously. These were modelled in a similar way, with the luminosities, noise properties and instrumental resolution equal to that of the component spectra in each bin being represented. From this a statistical error and bias could be determined for each bin. Finally, the environment of the quasar could be another source of systematic bias, and is investigated thoroughly in Chapter 5.

4.4 Results and discussion

4.4.1 Results

Table 4.4 summarises the main results of this Chapter for the individual quasar sightlines. The quoted errors in $\Gamma_{\rm bkg}$ consist of the measured error from 10 000 Monte Carlo realisations of the spectra, varying the luminosity, systemic redshift and forest $\tau_{\rm eff}$, added in quadrature to the expected sightline-to-sightline scatter as described in Section 4.3.3. The spectra with the largest errors in $L^{\rm Q}_{\nu_0}$ are those with continuum magnitudes measured directly from the SDSS spectra. Even though the errors in the luminosity could be rather large ($\sim 30\%$), the measured error was strongly dominated by errors in the redshift and $\tau_{\rm eff}$. The statistical power of a measurement from a single spectrum is rather limited and the error on $\Gamma_{\rm bkg}$ is substantial. Fig. 4.7 shows the estimated $\log(\Gamma_{\rm bkg})$ and $R_{\rm eq}$ for our sample as a function of redshift. The data points in Fig. 4.7 are consistent with a linear fit within the errors. There is therefore little we can say about spatial fluctuations in UVB intensity other than that they appear to be smaller than our measurement errors. We note that if we had used the α presented in Scott et al. (2004) then the $\log(\Gamma_{\rm bkg})$ values would have been on average 0.22 dex higher, and the error bars 10 per

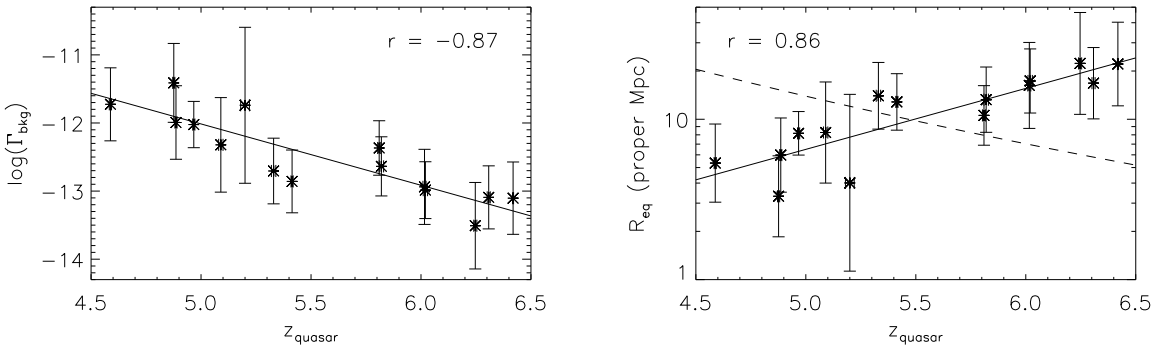


Figure 4.7: *Left-hand panel:* The bias corrected estimates of $\log(\Gamma_{\rm bkg})$ for the 15 individual quasars studied in this Chapter. The errors are calculated as the standard deviation in $\log(\Gamma_{\rm bkg})$ from 10 000 Monte Carlo realisations (varying the systemic redshift, luminosity and forest $\tau_{\rm eff}$) added in quadrature to the statistical error, which was individually calculated for each quasar sightline. Biases similar to those listed in Table 4.3 have also been removed. A smooth decline in $\Gamma_{\rm bkg}$ with redshift appears over this redshift range, with a formal correlation coefficient of −0.87. *Right-hand panel:* the inferred values of $R_{\rm eq}$ for each quasar. The error bars were calculated in the same way as those for $\Gamma_{\rm bkg}$. There is a strong increase towards higher redshifts, driven by the declining intensity of the UVB. A fit to this increase is shown by the solid line, and the formal correlation coefficient is 0.86. The dashed line marks the evolution in the mean free path of ionizing photons from the formula in Songaila and Cowie (2010). The size of the proximity region is larger than the mean free path at $z \gtrsim 5.5$. The implications of this are discussed in Section 4.4.2.

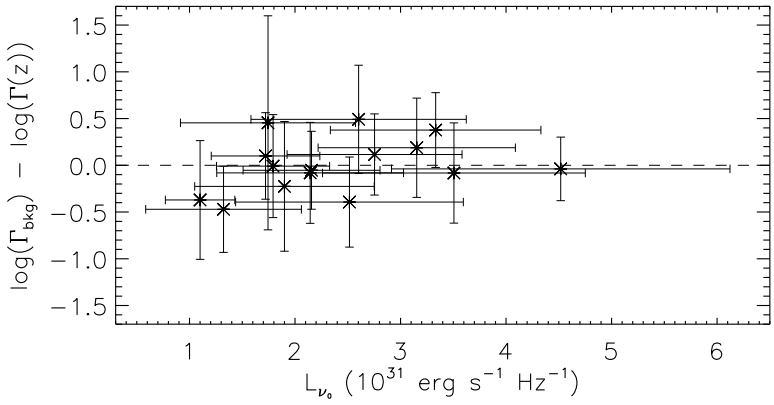


Figure 4.8: Plot of the deviance from the linear fit to $\log(\Gamma_{\rm bkg})$ in the left panel of Fig. 4.7 against the quasar luminosity. The dashed line marks zero deviation. All of the points are consistent with this line, which suggests that the measured values of $\log(\Gamma_{\rm bkg})$ have been adequately corrected for systematic biases from quasar luminosities.

cent larger. We also note that our redshift evolution of $R_{\rm eq}$ is opposite to the redshift evolution of the proximity region sizes as presented by Fan et al. (2006b). Note that this is entirely due to the different definition of the two sizes considered. For the proximity region size definition from Fan et al. (2006b) we find similar results.

In Fig. 4.8 we subtract out the linear evolution of the average $\log(\Gamma_{\rm bkg})$. There is no systematic trend with the luminosity of the quasar. This confirms that the systematic shifts induced by the luminosity of the quasar (see Table 4.3 and Fig. 4.15) have been suitably corrected for. Fig. 4.9 shows a sample of the observed spectra with their derived $R_{\rm eq}$ and expected average flux fall-off, in their rest

Name	z_q	σ_z	z source	$L^Q_{\nu_0}$ (10^{31} erg s^{-1} Hz^{-1})	M_{1450} (AB)	R_{eq} (Mpc)	Γ_{bkg} (10^{-12} s^{-1})	$\log(\Gamma_{bkg})$
SDSS J1148+5251	6.4189	0.002	CO	3.154 ± 0.935	–27.81	$22.1^{+18.2}_{-10.0}$	$0.08^{+0.19}_{-0.06}$	-13.10 ± 0.53
SDSS J1030+0524	6.308	0.007	Mg II	1.721 ± 0.516	–27.15	$16.8^{+11.2}_{-6.7}$	$0.08^{+0.15}_{-0.05}$	-13.09 ± 0.46
SDSS J1623+3112	6.247	0.007	Mg II	1.100 ± 0.329	–26.67	$22.3^{+24.0}_{-11.6}$	$0.03^{+0.10}_{-0.02}$	-13.51 ± 0.64
SDSS J0818+1722	6.02	0.01	Lyα forest	2.156 ± 0.649	–27.40	$17.3^{+10.1}_{-6.4}$	$0.10^{+0.17}_{-0.06}$	-12.99 ± 0.42
SDSS J1306+0356	6.016	0.007	Mg II	1.792 ± 0.536	–27.20	$16.3^{+13.9}_{-7.5}$	$0.12^{+0.30}_{-0.08}$	-12.94 ± 0.55
SDSS J0002+2550	5.82	0.01	Lyα forest	2.754 ± 0.829	–27.67	$13.3^{+7.9}_{-5.0}$	$0.23^{+0.40}_{-0.15}$	-12.64 ± 0.43
SDSS J0836+0054	5.810	0.007	Mg II	3.333 ± 0.998	–27.87	$10.6^{+5.6}_{-3.7}$	$0.43^{+0.65}_{-0.26}$	-12.37 ± 0.40
SDSS J0231-0728	5.41	0.01	Lyα forest	1.323 ± 0.738	–26.87	$12.8^{+6.4}_{-4.3}$	$0.14^{+0.26}_{-0.09}$	-12.86 ± 0.46
SDSS J1659+2709	5.33	0.01	Lyα forest	2.516 ± 1.079	–27.57	$14.0^{+8.6}_{-5.3}$	$0.20^{+0.40}_{-0.13}$	-12.71 ± 0.48
SDSS J0915+4924	5.20	0.02	Lyα + N V	1.744 ± 0.830	–27.17	$4.0^{+10.3}_{-2.9}$	$1.82^{+23.60}_{-1.69}$	-11.74 ± 1.15
SDSS J1204-0021	5.09	0.01	Lyα forest	1.900 ± 0.851	–27.26	$8.3^{+8.8}_{-4.3}$	$0.48^{+1.88}_{-0.38}$	-12.32 ± 0.69
SDSS J0011+1440	4.967	0.005	Lyα forest	4.518 ± 1.604	–28.20	$8.2^{+3.0}_{-2.2}$	$0.95^{+1.13}_{-0.52}$	-12.02 ± 0.34
SDSS J2225-0014	4.886	0.005	Lyα forest	2.144 ± 0.885	–27.39	$6.0^{+4.2}_{-2.5}$	$1.02^{+2.51}_{-0.72}$	-11.99 ± 0.54
SDSS J1616+0501	4.876	0.005	Lyα forest	2.603 ± 1.020	–27.60	$3.3^{+2.6}_{-1.5}$	$3.89^{+10.80}_{-2.86}$	-11.41 ± 0.58
SDSS J2147-0838	4.588	0.005	Lyα forest	3.506 ± 1.245	–27.93	$5.3^{+4.0}_{-2.3}$	$1.88^{+4.58}_{-1.33}$	-11.73 ± 0.54

Table 4.4: Tabulated results for each of the investigated quasar sightlines. Those with σ_z of 0.002 or 0.007 have systemic redshifts taken from the CO or Mg II emission lines (Carilli et al., 2010) while the others use either the Lyα + N V emission-line redshift, or (in most cases) use the redshift at which the Lyα forest appears to begin. Due to the high resolution of the spectra, errors for these are ≤ 0.01. The absolute magnitudes at rest-frame 1450 Å were calculated from either the published continuum magnitudes (Fan et al., 2001, 2003, 2004, 2006c) or the measured fluxes at rest-frame 1280 Å, extrapolated to 1450 Å by assuming a power-law continuum of the form $f_\nu \propto \nu^{-0.5}$.

frame, as well as denoting the section of each spectrum used for the measurement.

Our most robust results come from simultaneously fitting multiple sightlines. The data were grouped into low redshift ($z < 5.5$) and high redshift ($z > 5.5$) samples, and $\Gamma_{\rm bkg}$ was measured by finding where $\sum_i \Delta F_i = 0$, with ΔF calculated for each spectrum using the same trial $\Gamma_{\rm bkg}$. For the lower redshift bin, containing eight spectra, the average redshift was $z = 5.04$ and gave $\log(\Gamma_{\rm bkg}) = -12.15 \pm 0.16$ (0.32) at 68 (95) per cent confidence, whilst for the higher redshift bin with seven spectra and an average redshift of $z = 6.09$, $\log(\Gamma_{\rm bkg}) = -12.84 \pm 0.18$ (0.36). These will subsequently be referred to as the $z \sim 5$ and 6 samples, respectively, for ease of comparison to the simulations and previous work. The implications of these results will now be discussed.

4.4.2 Comparison to previous work

The results of fitting multiple lines of sight with a constant UVB are plotted in Fig. 4.10. We also plot a selection of UVB estimates from the literature, where the literature results have been scaled to the cosmology used in this Chapter, as well as to the same temperature-density relation parameters used by Bolton and Haehnelt (2007c), such that $T = T_0 \Delta^{\gamma-1}$, with T_0 and γ held constant at 10^4 K and 1.3, respectively. The UVB model of Haardt and Madau (2001) is also shown. As previously noted in Fig. 4.7, whilst the error bars are large for the individual sightlines, there is a clear trend of a decreasing UVB intensity in the redshift range $z \sim 5-6$. This decrease is more pronounced in the binned results.

The results presented here are the first proximity-effect measurements of $\Gamma_{\rm bkg}$ at these redshifts. Dall'Aglio et al. (2009) detected the proximity effect in 1733 spectra from the SDSS over the range $2 \lesssim z \lesssim 4.5$ and found that the UVB seemed to be remarkably flat over this redshift range, with $\log(\Gamma_{\rm bkg}) = -11.78 \pm 0.07$. This value is consistent with our results from individual sightlines over $4.5 < z < 5$. A more direct comparison can be made with the results of Dall'Aglio et al. (2008), who measured the proximity effect in 40 spectra with the UV-Visual Echelle Spectrograph (UVES) instrument on the Very Large Telescope. Comparing our results to the Dall'Aglio et al. (2008) results may be more appropriate, as both studies used high-resolution spectra. Combined, the proximity effect results suggest a smooth decrease in UVB from $z \sim 2$ to $z \sim 6$ by an order of magnitude (Fig. 4.10).

The UVB in the range $z \sim 5-6$ has previously only been probed using measurements of flux decrements (McDonald and Miralda-Escudé, 2001; Meiksin and White, 2004; Bolton and Haehnelt, 2007c; Wyithe and Bolton, 2011). There is excellent agreement between our binned results and those from flux decrements (see Fig. 4.10), which also suggest a significant decline in the UVB from $z \sim 4$ to $z \sim 6$.

Our measured evolution of the UVB intensity has important implications for reionization. The H I photoionization rate should scale as $\Gamma(z) \propto l(\nu_0, z)\epsilon_{\nu_0}$, where $l(\nu_0, z)$ is the mean free path of ionizing photons and ϵ_{ν_0} is the ionizing emissivity. The redshift evolution of $\Gamma_{\rm bkg}$ therefore gives an insight into the evolution of these two key variables. During the 'percolation' stage at the end of reionization, $l(\nu_0, z)$ is expected to evolve rapidly, in marked contrast to its gradual evolution in the post-reionization Universe (Gnedin and Fan, 2006). Consequently, the end of reionization should be indicated by a break in the evolution of $l(\nu_0, z)$. The smooth redshift evolution of our measurements of the UVB intensity (see left-hand panel of Fig. 4.7) implies that both $l(\nu_0, z)$ and ϵ_{ν_0} are evolving

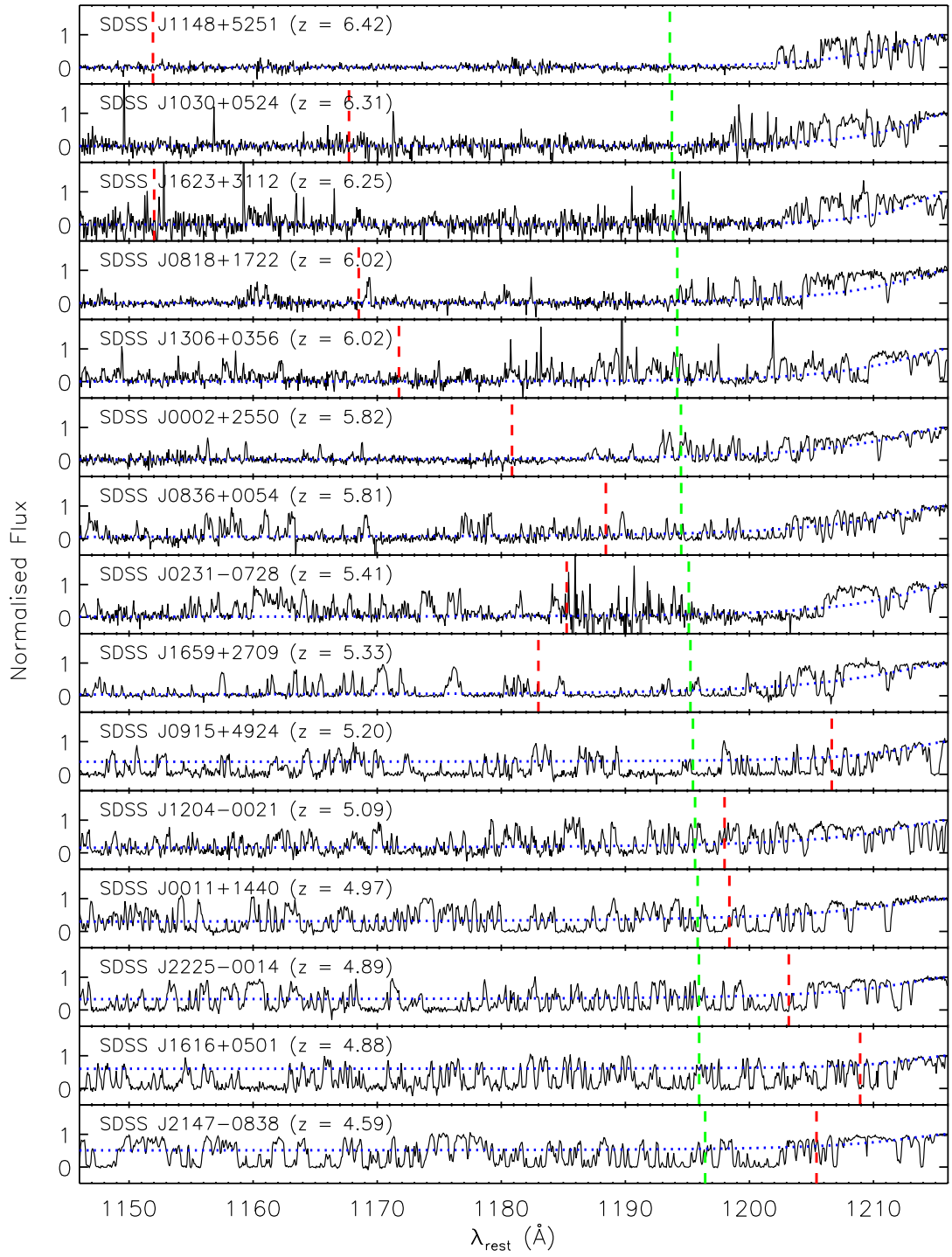


Figure 4.9: The 15 spectra used in this study, ordered by redshift. The spectra are presented in their rest wavelength, $\lambda_{\rm rest}$, in order to emphasise the proximity-effect region bluewards of the Lyα emission line, on the far right-hand side of the plot at 1216 Å. They have been normalised with the continuum fitting process described in Section 4.2.1, and have also been smoothed to a common pixel size of 16.7 km s^{-1} (observed frame) for clarity. The red dashed line indicates the derived value of $R_{\rm eq}$, whilst the blue dotted line is the expected fall-off in average flux. The area of spectrum to the right of the green dashed line was the section used for the proximity-effect measurement, and is $40h^{-1}$ comoving Mpc long.

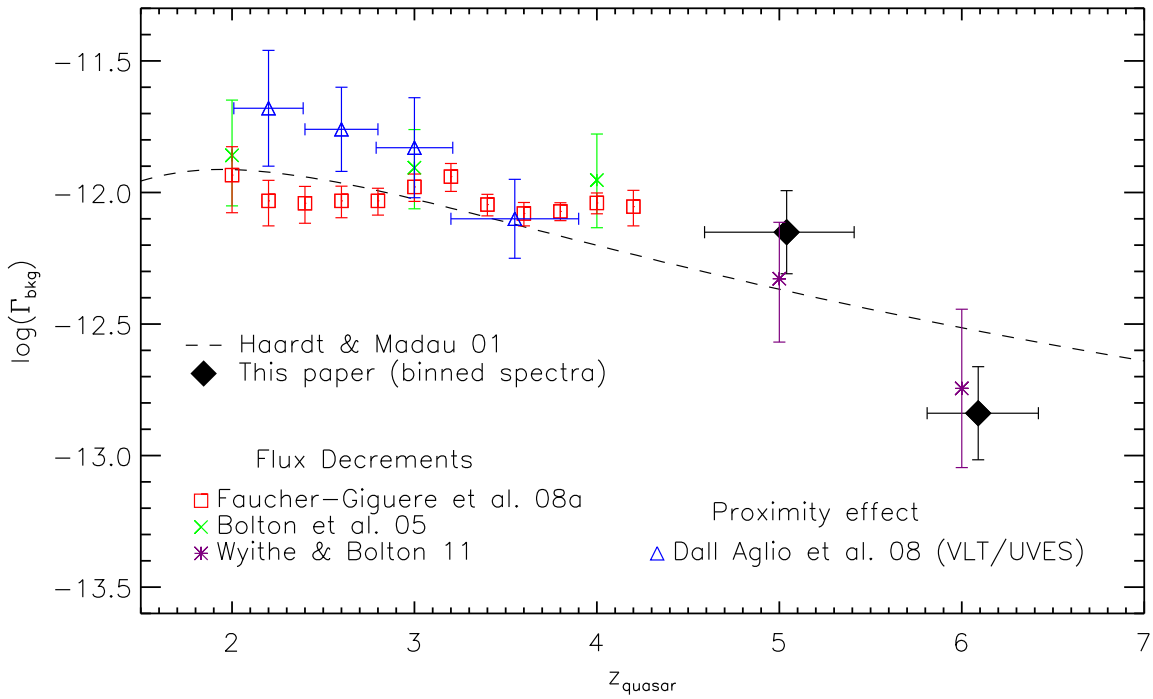


Figure 4.10: The evolution of the UVB in the redshift range $z = 2-6$. The solid points are the results of this Chapter, and represent the aggregate analysis over multiple sightlines. The spectra were coarsely binned into two subsets, namely $z > 5.5$ and $z < 5.5$. The lower redshift subset contained eight spectra and had an average redshift of $z_{\rm av} = 5.04$ whilst the upper had seven spectra and $z_{\rm av} = 6.09$. The errors displayed on our points are the 68 per cent confidence intervals, and exclude any correction due to environmental biases. Recent estimates from the literature are also plotted with their 1σ errors (Bolton et al., 2005; Dall'Aglio et al., 2008; Faucher-Giguère et al., 2008a; Wyithe and Bolton, 2011). The flux-decrement results have all been scaled to our adopted cosmology $(h, \Omega_{\rm m}, \Omega_{\rm b}h^2, \sigma_8) = (0.72, 0.26, 0.024, 0.85)$ and to the same temperature-density relation, $T = T_0\Delta^{\gamma-1}$, with T_0 and γ held constant at 10^4 K and 1.3, respectively, using the scaling relations from Bolton et al. (2005) and Bolton and Haehnelt (2007c). Our results at both $z \sim 5$ and 6 agree very well with those from flux-decrement measurements. The theoretical curve of Haardt and Madau (2001), assuming contributions from quasars and star-forming galaxies, is also plotted.

smoothly in the redshift range $4.6 < z < 6.4$, as otherwise they would have to rapidly evolve simultaneously in opposite directions, which appears very unlikely. This suggests that percolation has occurred at higher redshifts than are probed by our sample (although see Furlanetto and Mesinger, 2009).

Our measurements of the UVB can be combined with measurements of the mean free path to place constraints on the evolution of the ionizing emissivity. Songaila and Cowie (2010) recently measured the incidence of Lyman limit systems (LLSs; see Section 5.4) over $0 < z < 6$. They find an evolution in the mean free path which can be approximated as $l(\nu_0, z) = 50[(1+z)/4.5]^{-4.44^{+0.36}_{-0.32}}$. Using this fit, we infer that $l(\nu_0, z)$ decreases by a factor $\sim 1.5-2.5$ from $z = 5$ to 6. At face value the decrease of our measurements of $\Gamma_{\rm bkg}$ with redshift (a factor of $\sim 2.5-8$) therefore imply an emissivity that is either roughly constant or drops by up to a factor of ~ 5. Note that this is consistent with the decrease in the (dust-corrected) UV luminosity density between $z = 5$ and 6 measured by Bouwens et al. (2009) for rest wavelength ~ 1600 Å and integrated to either $0.3L^*_{z=3}$ or $0.04L^*_{z=3}$ [where $L^*_{z=3}$ is the luminosity derived by Steidel et al. (1999) at $z \sim 3$ and corresponds to $M_{1700,AB} = -21.07$].

The fit to the mean free path from Songaila and Cowie (2010) is plotted with the measured proximity region sizes in the right-hand panel of Fig. 4.7. Our proximity region sizes become larger than the mean free path as measured by Songaila and Cowie (2010) at $z \gtrsim 5.5$. If the quasar is able to ionize LLSs within the proximity region, then the mean free path within the proximity region will increase, allowing an enhanced contribution to the ionization rate from local galaxies and AGN. As a result, the proximity region may appear larger, and so the UVB may be underestimated in a proximity-effect analysis. Careful modeling of LLSs will be required to determine whether this is an important effect. However, it may partly explain the decrease in ionizing emissivity we infer from $z \sim 5$ to 6 (see also Section 5.3).

The ionizing emissivity is already very low at $z = 6$ (~ 1.5 ionizing photons per hydrogen atom per Hubble time) and so this evolution cannot continue to much higher redshifts without reionization failing to complete by $z = 6$ (Bolton and Haehnelt, 2007c). For this reason, unless there is a substantial increase in ϵ_{ν_0} at very high redshifts, the end of reionization, while potentially before $z = 6.4$, appears unlikely to occur much earlier.

4.5 Conclusions

We have presented new measurements of the photoionization rate of hydrogen by the ultraviolet background (UVB) using the proximity effect in quasar spectra. The 15 spectra in the sample cover the range $4.6 < z < 6.4$, allowing us to conduct the first proximity-effect measurements of $\Gamma_{\rm bkg}$ at these high redshifts.

For each quasar, $\Gamma_{\rm bkg}$ was calculated by modelling the total photoionization rate as a function of distance from the quasar, taking into account contributions from both the quasar and the UVB. The optical depths in the proximity zone were then modified to those expected in the absence of the enhanced ionization from the quasar. The preferred value of $\Gamma_{\rm bkg}$ was the one for which the resulting mean flux in the proximity zone was equal to that of the average Lyα forest at the same redshift.

We investigated a wide range of potential errors and biases affecting the proximity-effect measurements using numerical simulations. We found the error in $\Gamma_{\rm bkg}$ for an individual line of sight to be dominated by the error in the quasar redshift and $\tau_{\rm eff}$. Our raw measurements of the UVB intensity from individual sightlines at $z = 5$ should be typically overestimated by a factor of two, with random errors that are also roughly a factor of two. At this redshift, the largest bias results from the finite S/N of the spectra (50%). At $z = 6$, the UVB is overestimated by 30%, with a factor of 2.5 random error. The largest source of bias at this redshift is the uncertainty in $\tau_{\rm eff}$.

Our measured values of $\Gamma_{\rm bkg}$, corrected for biases, decline significantly from $z \sim 5$ to 6. For $z \sim 5$ we find $\log(\Gamma_{\rm bkg}) = -12.15 \pm 0.16$ (0.32) at 68 (95) per cent confidence, whilst at $z \sim 6$ we find $\log(\Gamma_{\rm bkg}) = -12.84 \pm 0.18$ (0.36), a decline significant at roughly the 3σ level. Within our sample, the UVB intensity measured from individual sightlines is seen to decline smoothly with redshift over $4.6 < z < 6.4$, but shows no sign of the rapid decline which may be expected in the late stages of reionization, when there is a rapid change in the attenuation length of ionizing photons.

Our results are in good agreement with UVB estimates from measurements of the mean flux decre-

ment in the redshift range $z = 5-6$, assuming an IGM temperature $T_0 = 10^4$ K. Both the proximity effect and flux-decrement measurements imply a decline in the intensity of the UVB by nearly an order of magnitude from $z = 4$ to 6.

Finally, we have combined our estimates of $\Gamma_{\rm bkg}$ with the evolution of the mean free path measured by Songaila and Cowie (2010). At face value, the combined measurements imply a decline in the ionizing emissivity of a factor of roughly two from $z \sim 5$ to 6, but it is important to keep in mind that at these redshifts measurements of the mean free path of ionizing photons are extremely difficult, and that with such a low emissivity reionization could barely have been completed by $z = 6.4$.

The results presented here represent some of the highest-redshift measurements of the UV background made to date, enabling us to probe deeper into the late(st) stages of hydrogen reionization. The next generation of optical and NIR telescopes will enable access to high-S/N quasar spectra at even higher redshifts, leading to improved measurements of the ionizing background and helping to establish a more complete picture of the final stages of the hydrogen reionization epoch.

Appendix A: Detailed discussion of the error analysis

In this appendix we show how varying some of the parameters in Table 4.3 affects the scatter and bias in the measured value of $\log(\Gamma_{\rm bkg})$ and induce a systematic bias (Fig. 4.11−4.16). By modelling the effects of these parameters, an expected sightline-to-sightline scatter, $\sigma_{\rm exp}$, and bias, $\epsilon_{\rm exp}$, can be calculated for each individual quasar. The former contributes to our quoted errors in $\log(\Gamma_{\rm bkg})$, and the latter is used such that $\log(\Gamma^{\rm corr}_{\rm bkg}) = \log(\Gamma^{\rm raw}_{\rm bkg}) - \epsilon_{\rm exp}$. All values of $\log(\Gamma_{\rm bkg})$ quoted in Table 4.4 have been corrected for these biases. For most of the spectra, the combined correction for all the sources of error is $\epsilon_{\rm exp} \lesssim 0.2$ dex.

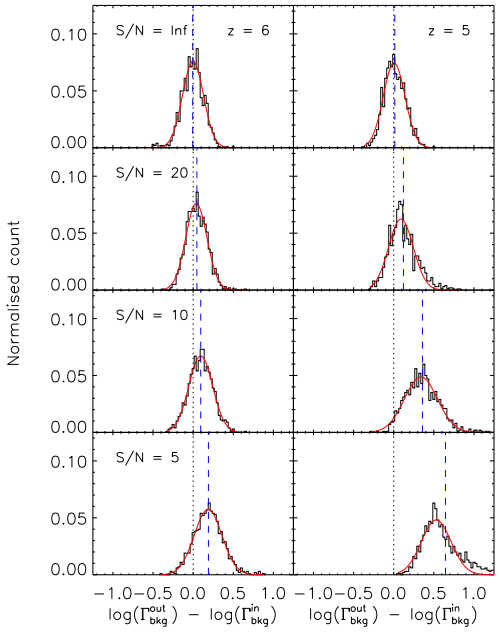


Figure 4.11: Distribution of the estimated values of $\log(\Gamma_{\rm bkg})$ as a function of S/N for simulated spectra with HIRES resolution. A proximity region size of 10 Mpc is assumed. Apart from the luminosity of the quasar and the peculiar velocities of the gas, there are no other sources of error. The dotted line marks the input value, while the blue dashed line marks the mean of the simulated data set. The estimated values of $\log(\Gamma_{\rm bkg})$ are well fit by a Gaussian distribution.

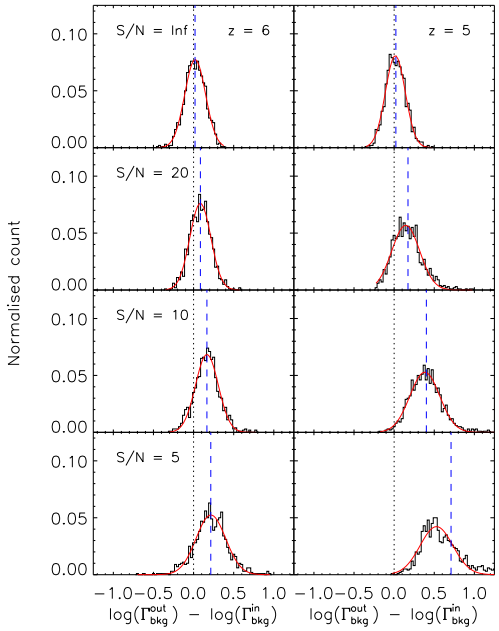


Figure 4.12: Same as Fig. 4.11, but for simulated data with MIKE velocity resolution.

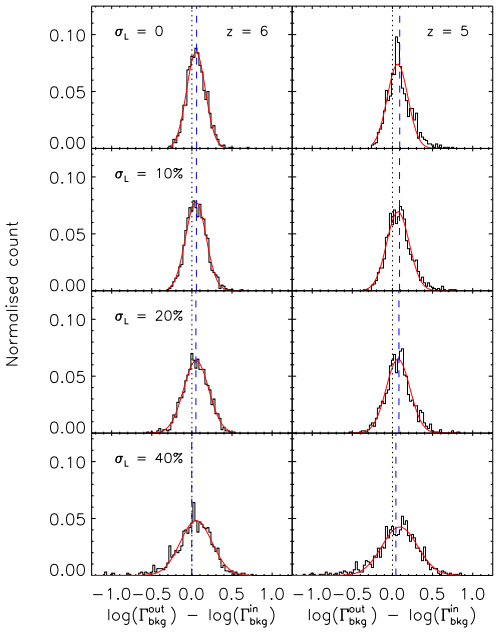


Figure 4.13: Expected distribution of errors in $\log(\Gamma_{\rm bkg})$ for various errors in the luminosity, assuming a fixed S/N of 20, and HIRES resolution.

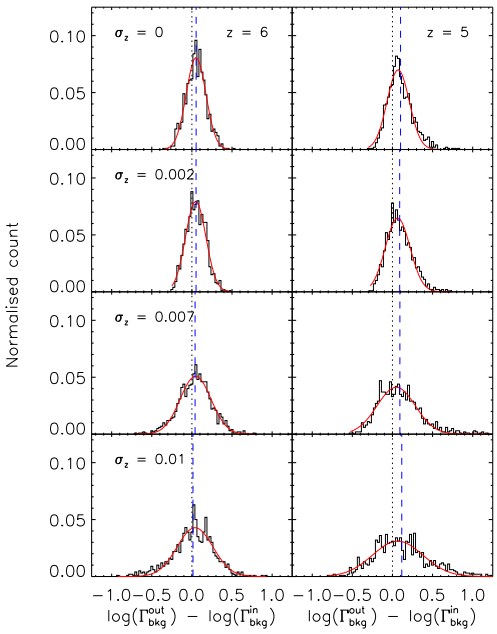


Figure 4.14: Expected distribution of errors in $\log(\Gamma_{\rm bkg})$ for various errors in the redshift, for an assumed S/N of 20, and HIRES resolution. The top panels assume no redshift error. The next two panels down represent CO and MgII redshifts, respectively, and the bottom panels represent the errors on redshifts determined from the onset of the Lyα forest.

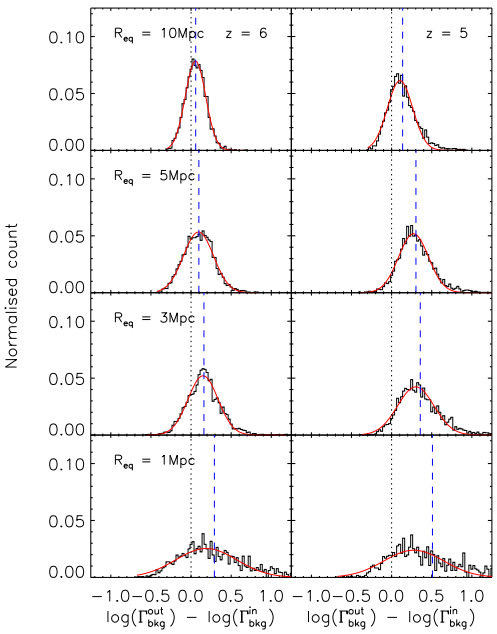


Figure 4.15: Expected distribution of errors in $\log(\Gamma_{\rm bkg})$ for various input proximity region sizes (i.e. luminosity of the quasar), for a fixed S/N of 20, and HIRES resolution.

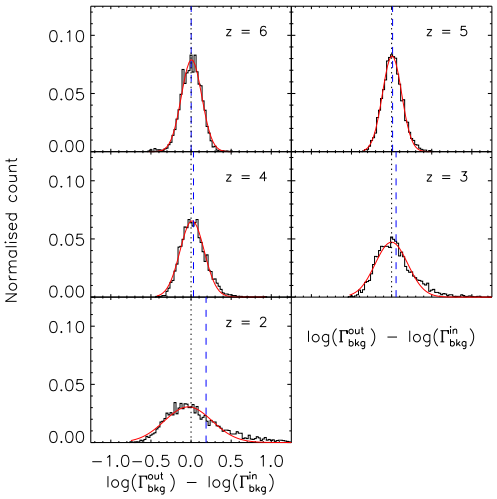


Figure 4.16: Expected distribution of errors in $\log(\Gamma_{\rm bkg})$ for various quasar redshifts, for noiseless spectra (i.e. S/N = ∞), and at HIRES resolution. Even for noiseless data, the method struggles at lower redshift. The distribution loses its Gaussian symmetry, although the mode of the distribution does maintain its position close to the input value (see also Dall'Aglio et al., 2008).

At night astronomers agree.

Phillis's Age

MATTHEW PRIOR

5

Investigating systematic biases in proximity-effect measurements of the ultraviolet background

ESTIMATES of the ionizing ultraviolet background (UVB) at $2 \lesssim z \lesssim 4$ from observations of the proximity effect in quasar spectra tend to be systematically higher than the estimates from other independent measures. This is often attributed to the biased environment in which the quasar resides. We investigate the significance of this and other sources of systematic bias at $z = 5$ and 6 in order to test the robustness of the results for the metagalactic hydrogen ionization rate, $\Gamma_{\rm bkg}$, presented in Chapter 4. We find that the impacts of large-scale overdensities and Lyman limit systems on $\Gamma_{\rm bkg}$ are small at $z > 5$, and that any overestimation in $\Gamma_{\rm bkg}$ they may cause is potentially partially counteracted by the underestimation in $\Gamma_{\rm bkg}$ from ignoring quasar heating effects. The proximity effect should therefore give relatively unbiased estimates of $\Gamma_{\rm bkg}$ at $z > 5$, in contrast to lower redshifts, and so differences with other independent measures are unlikely to be due to the quasar environment.

5.1 Introduction

Upon the completion of reionization, the Universe is kept ionized through the presence of an ionizing radiation field, known as the metagalactic ultraviolet background (UVB). Historically, the intensity of the UVB has been measured from the proximity effect seen in quasar spectra (e.g. Bajtlik et al., 1988, hereafter BDO). This effect attributes the region of enhanced transmission observed in quasar spectra blueward of the Lyα emission line to the dominance of the quasar's ionizing output over that of the UVB. Thus for a given ionizing luminosity and size of the region of enhanced transmission, the intensity of the UVB can be estimated.

An alternative method to estimate the intensity of the UVB is to measure the mean flux of the Lyα forest and compare the result with simulations (e.g. Rauch et al., 1997). The mean flux is determined by both the cosmology and the UVB, so by using independent measures of the required cosmological parameters the intensity of the UVB can be varied in the simulation until the flux decrement matches the observed value.

Before the work presented in Chapter 4, the UVB at $z > 4.5$ had only been determined from flux decrements. At $z < 4.5$, however, estimates of the UVB from the proximity effect and flux-decrement methods from the literature can be directly compared. These are generally discrepant at $2 < z < 4$, with the proximity-effect estimates systematically higher for most studies. Whilst the proximity effect can be used on both individual sightlines (e.g. Williger et al., 1994; Cristiani et al., 1995; Lu et al., 1996) and in a statistical sense across many spectra (e.g. BDO; Cooke et al., 1997; Scott et al., 2000) to help constrain the redshift evolution of the UVB, several issues must be addressed before the UVB estimates from the proximity effect can be considered robust. The proximity effect essentially measures the direct signature of an enhanced ionization field near the quasar by noting a reduced fraction of H I extending out across a region of a few physical Mpc. Therefore, the conversion from this physical size to a value for the UVB will be highly sensitive to any assumptions made about the quasar and its environment. There are three main systematic uncertainties in this conversion.

First, it is assumed that the quasar turned on a sufficiently long time ago that its proximity region is in photoionization equilibrium with the intergalactic medium (IGM), and that it has been at its current luminosity over a similar timescale. The unknown level of variability in quasar luminosity on the order of 10^4 years (the time-scales required for a region of the highly ionized IGM to reach ionization equilibrium) means that the measured proximity region size may not correspond to a region in ionization equilibrium, and so any measurement of the UVB may be biased (Pentericci et al., 2002; Schirber et al., 2004; Dall'Aglio et al., 2008).

The second issue is that in order to calculate an accurate size of the proximity region, an accurate systemic redshift for the quasar is needed. Redshifts determined from broad high-ionization lines typically underestimate the systemic redshift (e.g. Richards et al., 2002), which causes the UVB to be overestimated (Espey, 1993). It is worth noting, however, that an accurate systemic redshift can be determined from low-ionization lines, and so for some spectra this effect can be avoided. Also, it should be noted that at high redshift the extent of enhanced transmitted flux will be larger than even the largest expected error in systemic redshifts.

Finally, in the classical proximity-effect analysis by BDO, the IGM within the proximity region is presumed to have the same density distribution as the general IGM. Quasars are hosted by massive galaxies, however, which are expected to reside in environments with higher than average density out to rather large distances (Granato et al., 2004; Fontanot et al., 2006; da Ângela et al., 2008). In that case, the proximity region would be smaller than expected for a region of the same size with an average density close to the global mean density (D'Odorico et al., 2008), and thus the UVB may be overestimated (Loeb and Eisenstein, 1995). If quasar luminosity correlates with host halo mass, then this environmental bias may be worst for the most luminous quasars (e.g. Pascarelle et al., 2001; Adelberger et al., 2003; Kim and Croft, 2008). By using an independent measure of the UVB from

flux decrements, Rollinde et al. (2005) found tentative evidence that quasars at $z = 2-3$ may reside in haloes as massive as $\sim 10^{14}\mathrm{M}_{\odot}$, similar to the most massive halo in the Millennium Simulation. Guimarães et al. (2007) have claimed to detect a similar effect at $z \sim 4$. This is believed to be the main reason for the discrepancies between estimates of the UVB using the proximity effect and the flux-decrement method at redshift $z \sim 2-4$ (Faucher-Giguère et al., 2008b). Other recent papers, however, have claimed to have overcome the environmental bias of an enhanced average density in the proximity zone. Dall'Aglio et al. (2008) report that with their measurement method only 10% of their sample at $2 \lesssim z \lesssim 4.5$ showed significant excess absorption attributable to an overdense environment, and that for the majority of the quasars in their sample the proximity-effect measurement of the UVB did not appear to be affected by an overdense quasar environment (at least on scales > 3 Mpc).

Parcels of neutral gas in the vicinity of the quasar can also have an impact. Lyman limit systems are absorption features that are optically thick ($\tau > 1$) to Lyman limit photons (Tytler, 1982), and so prevent the quasar from ionizing as large a volume of the IGM as in their absence. Consequently, if in the line of sight to a quasar, they can truncate the extent of enhanced transmission in a spectrum and bias any estimate of the size of the proximity region.

Another environmental effect that is often neglected is the consequence of the quasar heating the surrounding IGM via ionization of He II . Bolton et al. (2010) found that the gas within ~ 5 proper Mpc of the $z \approx 6$ quasar SDSS J0818+1722 was $\sim 10^4$ K hotter than that presumed for the general IGM at that redshift. Higher temperatures within the proximity region will lead to more transmission, and so the UVB might then be underestimated.

In this Chapter we quantify the systematic biases associated with the environment of the quasar with regard to their effect on the proximity-effect measurements of the UVB presented in Chapter 4. All of these effects were investigated at $z = 5$ and 6 with the simulations described in Section 4.2.2, and using the analysis method described in Section 4.3.

The remainder of this Chapter is organised as follows. In Section 5.2 we describe the bias introduced by errors in the systemic redshift and ionizing luminosity. In Section 5.3 we explore the biases due to the quasar residing in an overdense region of the IGM, and the impact of the expected clustering of other ionizing sources around it. Deviations from the assumed $1/r^2$ fall-off in ionizing flux due to Lyman limit systems are discussed in Section 5.4, as is the impact of heating of the IGM close to the quasar in Section 5.5. Finally, we present our conclusions in Section 5.6. Throughout this Chapter we use the cosmological parameters taken from the mean of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 5-yr data set (Komatsu et al., 2009), with Hubble constant $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and density parameters $(\Omega_\mathrm{m}, \Omega_\Lambda) = (0.26, 0.74)$.

5.2 Luminosity and redshift errors

Errors in the Lyman limit luminosity of observed quasars are due to uncertainty in the power-law slope of the continuum blueward of the Lyα emission line, and uncertainty in the continuum flux. Within our simulated spectra the true quasar luminosity was constant, however we varied the luminosity associated with each sightline in accordance with the typical observed errors. Similarly, the

true systemic redshift of the simulated quasars was equal to the redshift of the simulation snapshot. To imitate errors in the redshift we either masked the pixels just blueward of the quasar, or appended pixels with $F = 1$ (before addition of noise) redward of the quasar.

Investigation with both the real and simulated spectra showed that the measured error on the UVB is dominated by errors in the redshift and the effective optical depth of the forest, and not by errors in the quasar luminosity. Even random errors in luminosity of up to 40% had only a comparable effect to small errors ($\sigma_z \sim 0.01$) in redshift (see Figs. 4.13 and 4.14). Of course, if the variability in quasar luminosity is not simply random scatter, but is instead systematic (for example, if the quasars are only observed in their brightest stages) then there will be a systematic bias as well as an additional random error. The method ultimately is sensitive to the gradient of the fall-off in the mean transmitted flux throughout the spectrum, and so will give the same estimate for $R_{\rm eq}$ regardless of the quasar luminosity. Therefore, if all the quasars are radiating at a systematically brighter luminosity than the one that established their proximity region, then the values of $\Gamma_{\rm bkg}$ will have been overestimated by the same amount, since they are linearly related (Equation 4.9).

Several quasars in Chapter 4 have very accurate systemic redshifts from either CO or Mg II emission lines (Carilli et al., 2010), and the high-resolution spectra themselves allow for a reasonably accurate determination of the onset of the Lyα forest. Consequently almost all the redshift errors are estimated to be ≤ 0.01. This corresponds to a line of sight size much smaller than the proximity region ($\sim$ 10 Mpc). We therefore believe that these two possible sources of error have been suitably dealt with.

5.3 Effect of quasars lying in overdensities

One of the more contentious assumptions used in proximity region measurements is that the environment of the quasar within the proximity region is similar to that of the general IGM. The quasar should lie in an overdense environment, and therefore could struggle to ionize as large a volume as it would in the general IGM, thus reducing the apparent size of the proximity region. As already discussed, there have recently been a number of studies suggesting that proximity-effect measurements may indeed suffer from a substantial bias due to this effect (Rollinde et al., 2005; Guimarães et al., 2007; Faucher-Giguère et al., 2008b).

We have attempted to estimate how such a bias may affect our Chapter 4 measurements, which are at significantly higher redshift than previous studies. As a first attempt we examined the effect of placing the quasars at the centres of massive dark matter haloes in our fiducial simulation, as well as in simulations with different box sizes and mass resolutions.

We chose sightlines through the largest (by total mass) 500 haloes in Models B-F in three perpendicular directions through the box. Due to the low resolution and small box size of Model A, only 100 haloes could be identified at $z = 6$ and 250 at $z = 5$. For each halo the sightline was adjusted so that it began at the centre of a massive halo. Since the simulation box is cyclic the sightlines needed to be split in half so the sightline only passed through the overdensity once, and consequently for each halo six sightlines could be drawn (i.e. leaving the halo from both sides in three perpendicular directions).

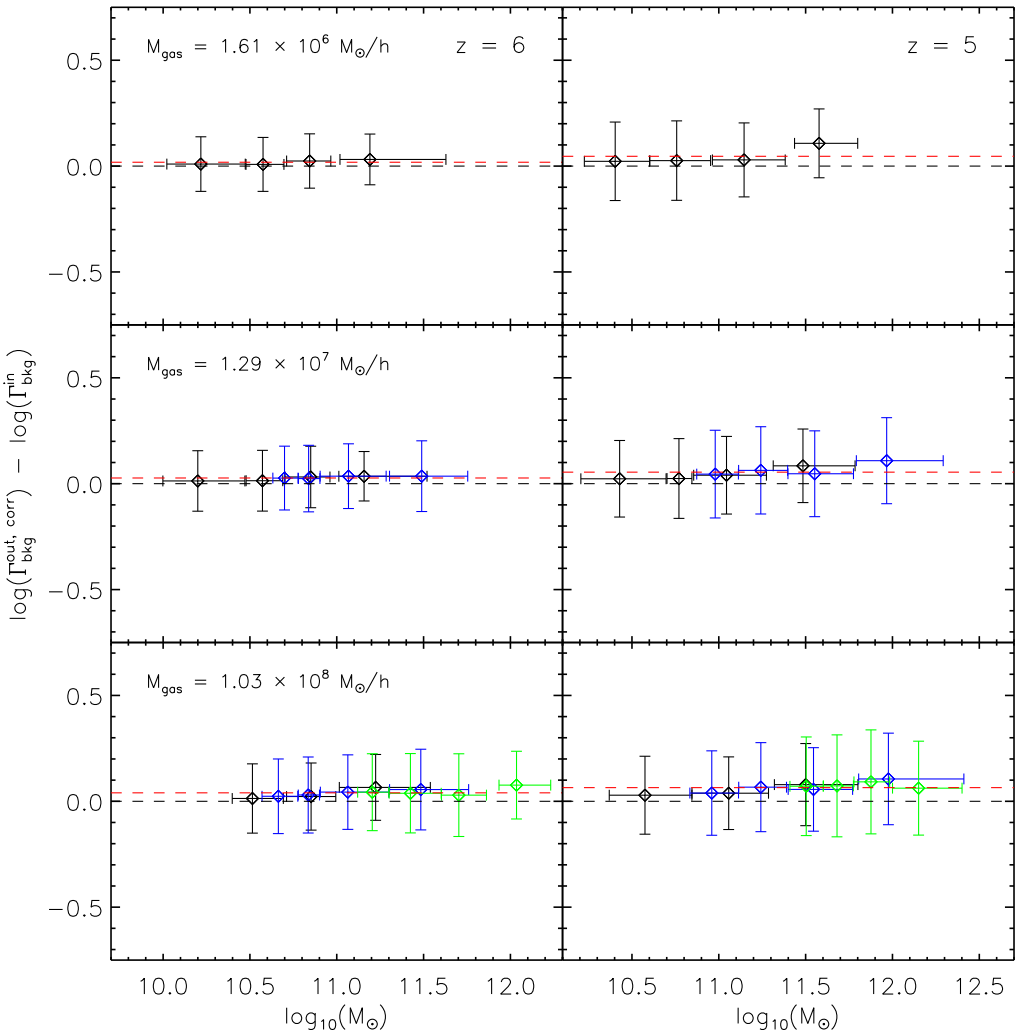


Figure 5.1: *Top row:* The ratio of the estimated $\Gamma_{\rm bkg}$ (corrected for systematic shifts due to noise) to the input $\Gamma_{\rm bkg}$ for sightlines starting in the 500 most massive haloes in the fiducial simulation, Model C. The proximity region in each spectra is 10 proper Mpc. On the left results are shown for $z = 6$ and on the right for $z = 5$. The black dashed line is for zero bias, whilst the red line marks the true average value estimated. For these haloes there is a systematic overestimation of $\lesssim 0.1$ dex, but apparently no dependence on halo mass. *Middle row:* Same as above, but for Models B (black) and E (blue). Model B has been resampled to the same pixel size as Model E. *Bottom row:* Same as above but for Models A (black), D (blue) and F (green). Due to the low resolution and small box size in Model A only 100 haloes could be identified at $z = 6$ and 250 haloes at $z = 5$. Models A and D have been resampled to the same pixel size as Model F. Lowering the mass resolution introduces a systematic overestimation in log(Γbkg). This is due to the fact that the voids are under-resolved and so the proximity regions appear systematically smaller. The larger pixel sizes used in the lower mass resolution simulations also smooth the transmitted flux, so the proximity region is again underestimated (see Section 4.3.3). This effect is the stronger of the two. With these offsets taken into account, there is still no significant increase in $\Gamma_{\rm bkg}$ with halo mass, even in the bottom panel, which includes an $80h^{-1}$ comoving Mpc box with haloes of mass $> 10^{12}$M$_\odot$.

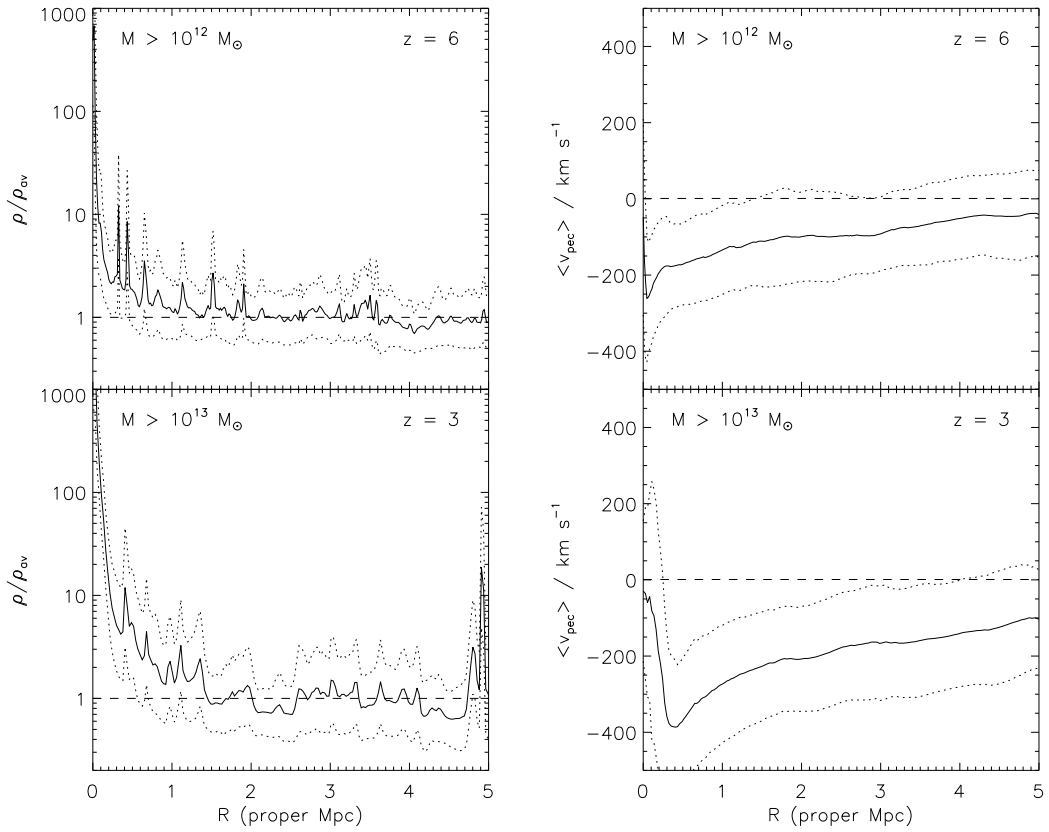


Figure 5.2: Mean density and peculiar velocity profiles around the most massive haloes in our simulation boxes. *Top left:* The solid line is the average density field in the sightlines that start in haloes of masses greater than $10^{12}\mathrm{M}_{\odot}$ in Model F at $z = 6$, plotted in proper Mpc. The dotted lines mark the $\pm 1\sigma$ deviations, and the mean density in the simulation is marked with the horizontal dashed line. Any significant overdensity is restricted to a region of ~ 1 Mpc in these haloes, and the overdense region will be even smaller for lower mass haloes. *Top right:* The solid line is the average peculiar velocity field in the same sightlines in Model F at $z = 6$, plotted in proper Mpc. The dotted lines mark the $\pm 1\sigma$ deviations, and $v_{\mathrm{pec}} = 0$ km s^{-1} is marked with the horizontal dashed line. The gravitational attraction of the overdensity means that the gas is infalling with a velocity of ~ 200 km s^{-1}, which will make the proximity region appear smaller. The infall velocity will be less for lower mass haloes. *Bottom left:* The same as in the top left panel but for haloes with mass greater than $10^{13}\mathrm{M}_{\odot}$ at $z = 3$. These are essentially the same haloes as those at $z = 6$, but have become more massive through hierarchical growth. The overdense region is larger, extending out to ~ 3 Mpc, with evidence for smaller haloes clustered nearby (the small peak at ~ 5 Mpc). If the UVB is higher at this redshift then proximity region sizes will decrease, and may be comparable to the extent of the overdensity. In that case measurements of the size of the proximity region could depend strongly on the mass of the quasar host halo. *Bottom right:* The same as in the top right panel but for haloes with mass greater than $10^{13}\mathrm{M}_{\odot}$ at $z = 3$. Due to the greater masses, the infall velocity is higher than at $z = 6$.

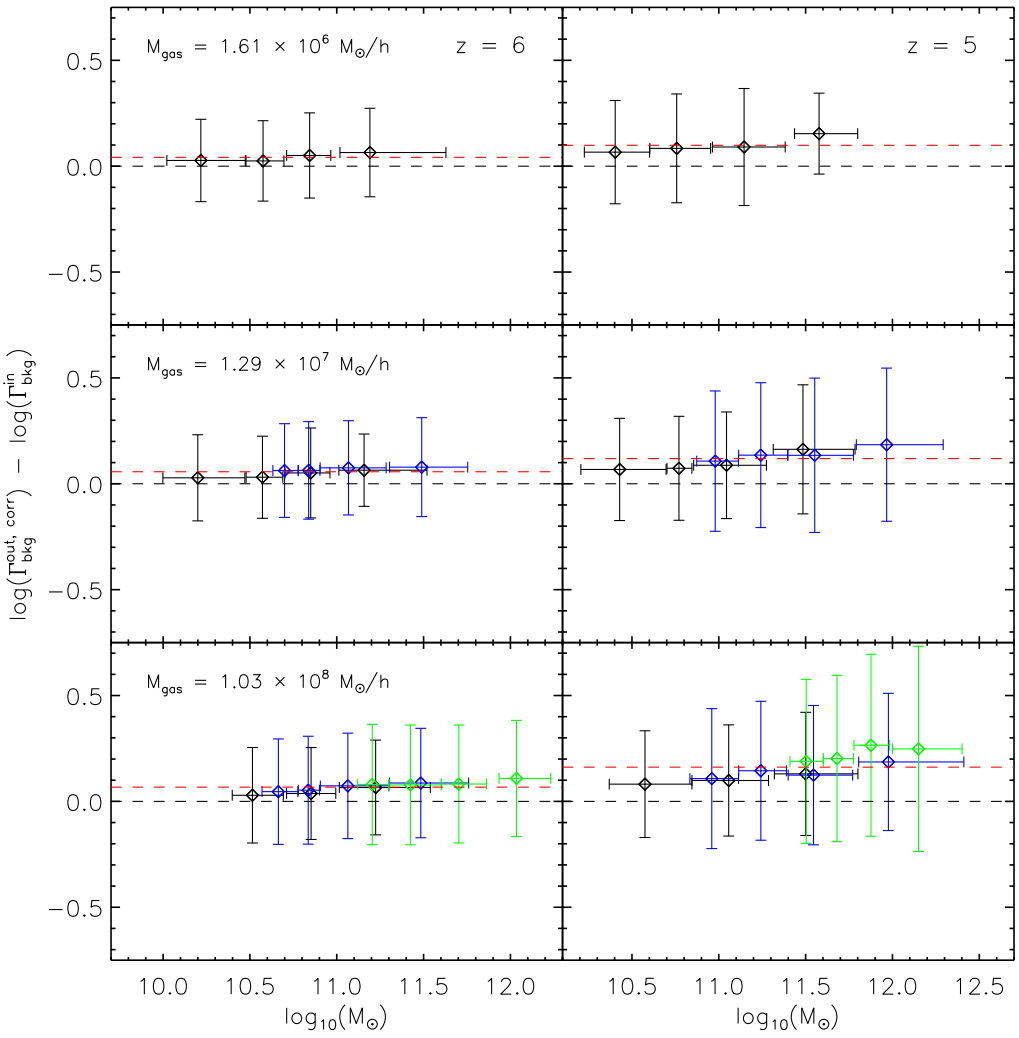


Figure 5.3: Same as Fig. 5.1, but the proximity region in each spectra is 5 proper Mpc. There is little change at $z = 6$ (left) where there is still a systematic overestimation of $\lesssim 0.1$ dex, but apparently no dependence on halo mass. However, at $z = 5$ (right), the systematic overestimation seems to show a weak dependence on halo mass, which is particularly pronounced in the bottom panel. This is because the extent of the overdensity is closer in size to that of the proximity region, and so the estimates of $\Gamma_{\rm bkg}$ become increasingly biased.

The halo sightlines were then spliced with random sightlines so that the total length of each spectrum was $40h^{-1}$ comoving Mpc. The proximity region size was chosen to be 10 proper Mpc, as in previous tests with the simulations. The results for the haloes in the fiducial simulation, Model C, are shown in the top panel of Fig. 5.1. There is no significant change as the mass increases, with a Gaussian spread in $\log(\Gamma_{\rm bkg})$ in each bin of similar width to that from the random sightlines (Fig. 4.6).

The box size for Model C is, however, only $20h^{-1}$ comoving Mpc, which limits the maximum halo mass to just over $4 \times 10^{11}{\rm M}_{\odot}$. Luminous quasars are believed to reside in more massive host hales. The best observational constraints are inferred from the clustering analysis of quasars combined with predictions from ΛCDM simulations and suggest host halo masses in the range 10^{12} – $10^{13}{\rm M}_{\odot}$ (e.g. Bonoli et al., 2009). It is important to note here that the very luminous quasars often

used for proximity-effect measurements are more luminous than those available for clustering analysis. It is thus perhaps not surprising that the proximity-effect studies by Rollinde et al. (2005) and Guimarães et al. (2007) suggest that, at $z \sim 2.5-4$, the host haloes of these very luminous quasars are even more massive. Fortunately, at the redshifts we consider here haloes more massive than $10^{13}\mathrm{M}_{\odot}$ are not yet expected to have formed in large enough numbers to be credible candidates to host even the most luminous observed quasars (Springel et al., 2005; Sijacki et al., 2009). In order to probe overdensities around haloes in this mass range, larger boxes are needed than we have considered so far. Thus, simulations with $40h^{-1}$ and $80h^{-1}$ comoving Mpc box size were also investigated, and compared with $20h^{-1}$ Mpc boxes with equal mass resolution. Model B was compared to Model E (see middle panel in Fig. 5.1), and Model A to Models D and F (bottom panel). Sightlines through Models A, B and D were resampled so that the pixel size (and mean forest flux) was the same as in the largest simulation to which they were being compared. The highest mass bin in Model F ($> 10^{12}\mathrm{M}_{\odot}$) still shows no more significant bias (in comparison to less massive haloes in Models A and D), seeming to confirm the results from Model C that, at least at $z = 5-6$, the mass dependence of the bias is rather weak.

The largest halo mass in the $z = 6$ simulation was $10^{12.2}\mathrm{M}_{\odot}$, which may still be nearly an order of magnitude less massive than the haloes in which luminous quasars reside. To get a rough estimate of the effect of more massive haloes we have taken the density profile of haloes with $M > 10^{13}\mathrm{M}_{\odot}$ at $z = 3$ in our simulation (the bottom panel in Fig. 5.2) and used this to modify the optical depths of spectra created from random sightlines in the $z = 6$ simulation. We modified the optical depths by applying the relation $\tau \propto \Delta^{2(\gamma-1)}$, where γ is from the temperature-density relation, $T(\Delta) \propto \Delta^{\gamma-1}$, and is taken to be 1.3. The extra systematic bias from such a massive halo is shown in Table 4.3. The effect is small at $z = 6$, and even though it is more important at $z = 5$, it is still not the dominant source of bias or main contributor to the error budget. Obviously this is just an approximation of how a halo this massive would affect the results, as the density profile was applied to the spectra artificially and will not be self-consistent with the other simulation outputs along the line of sight. However, since the effect is relatively minor we would not expect it to be a dominant source of error even if it were modelled self-consistently using a halo found in a much larger simulation box. Since the quasars at $z = 6$ will probably lie in haloes with $M < 10^{13}\mathrm{M}_{\odot}$ our estimate should be an upper limit on the effect of overdensities, and thus we conclude that at $z = 5-6$ the expected overdensity around our quasars is not critically important.

The most likely explanation for the lack of dependence of the bias of the UVB estimate on the mass of the host halo is that the proximity region is large in comparison to the size of the overdensity. This is demonstrated in the right-hand panels of Fig. 5.2 where the overdensity profile is shown at $z = 6$ and $z = 3$ for the largest haloes in Model F ($M > 10^{12}\mathrm{M}_{\odot}$ and $M > 10^{13}\mathrm{M}_{\odot}$ respectively). The gas will tend to be infalling (see left-hand panels of Fig. 5.2), and the corresponding peculiar velocities will further decrease the apparent size of the proximity region in velocity space (Faucher-Giguère et al., 2008b). In our simulated spectra the proximity regions were 10 proper Mpc, which is nearly an order of magnitude larger than the overdensity expected in the largest haloes at $z = 6$. The properties of the IGM within the proximity region should therefore not be strongly biased due to the overdensity. To

test this we reran the tests using less luminous quasars with a proximity region size of 5 proper Mpc, hosted in the same haloes. The results are shown in Fig. 5.3. As expected, as the proximity region size approaches the extent of the overdensity, the systematic bias becomes increasingly sensitive to the mass of the halo. This will be an important factor at lower redshifts, where not only has hierarchical growth increased the magnitude of the overdensity, but the UVB is also higher, making the proximity region smaller for a given quasar luminosity. Consequently, there may be strong systematic bias with halo mass at $z \lesssim 3$, as suggested by various studies (Rollinde et al., 2005; Faucher-Giguère et al., 2008b).

As a final point, we note that a very similar test was recently carried out by Partl et al. (2011). They used hydrodynamical simulations with outputs at $z = 3$, 4, and 4.8 and measured the strength of the proximity effect from a quasar with $\log L_{\nu_0} = 31$ (i.e. less luminous than those tested here) placed in the 20 largest haloes in their simulations to quantify any influence of the environmental density field. They find that the average density within the proximity region is more important than the halo mass. Given that higher mass haloes will have a higher average density out to a particular radius, but that the difference in average density between different halo masses will be small when measured out to large radii, then this is broadly in agreement with our findings. They also find that the scatter in $\Gamma_{\rm bkg}$ increases with decreasing redshift, which is again consistent with our results (see Fig. 5.1).

5.3.1 Clustering of LBGs around quasars

Some clustering of galaxies and faint active galactic nuclei (AGN) is expected around luminous quasars (e.g. Utsumi et al., 2010). The intensity of the UVB in the vicinity of luminous quasars could therefore be enhanced due to other nearby sources. Our knowledge about the clustering of faint AGN and galaxies around bright high-redshift quasars is rather sparse. We can nevertheless use the magnitudes and positions of the Lyman break galaxies (LBGs) found by Utsumi et al. (2010) for a rough estimate. In a field containing a quasar at $z = 6.43$ they found 7 LBGs at $z > 6.4$, based on their colours. Let us then calculate L_{ν_0} for each LBG from their z_R band magnitude, assuming they radiate isotropically, are all at the same redshift as the quasar, and have a ratio of far-UV to extreme-UV flux, $f_{1500}/f_{900} = 22$ (Shapley et al., 2006), as measured for $z \sim 3$ LBGs. With projected distances of $2-5$ Mpc the LBGs would all lie within the typical proximity zones of our sample if they were indeed at the same redshift. Even if the LBGs would have a Lyman continuum as blue as we assume for the quasars (i.e. $\alpha = 1.61$) they would increase the size of the proximity region by as little as ~ 0.1 Mpc, and the UVB would typically be underestimated by 0.01 dex (3%). There could obviously be a substantial contribution from more numerous fainter objects below the detection limit. The faint end slope of the luminosity function of LBGs at $z \sim 6.4$ is not well constrained, but if we assume a slope of -1.73 (Bouwens et al., 2011b) down to zero luminosity in the volume of the proximity region the UVB is still only underestimated by 0.04 dex (11%). This is, however, likely to be an upper limit, as in reality the 7 LBGs do not have precise redshifts and may not be within the proximity region of the quasar. This should outweigh the possibility that the galaxies may have a larger value of f_{1500}/f_{900} than we have assumed here. We therefore conclude that the effect of nearby LBGs could be noticeable but appears likely to be small.

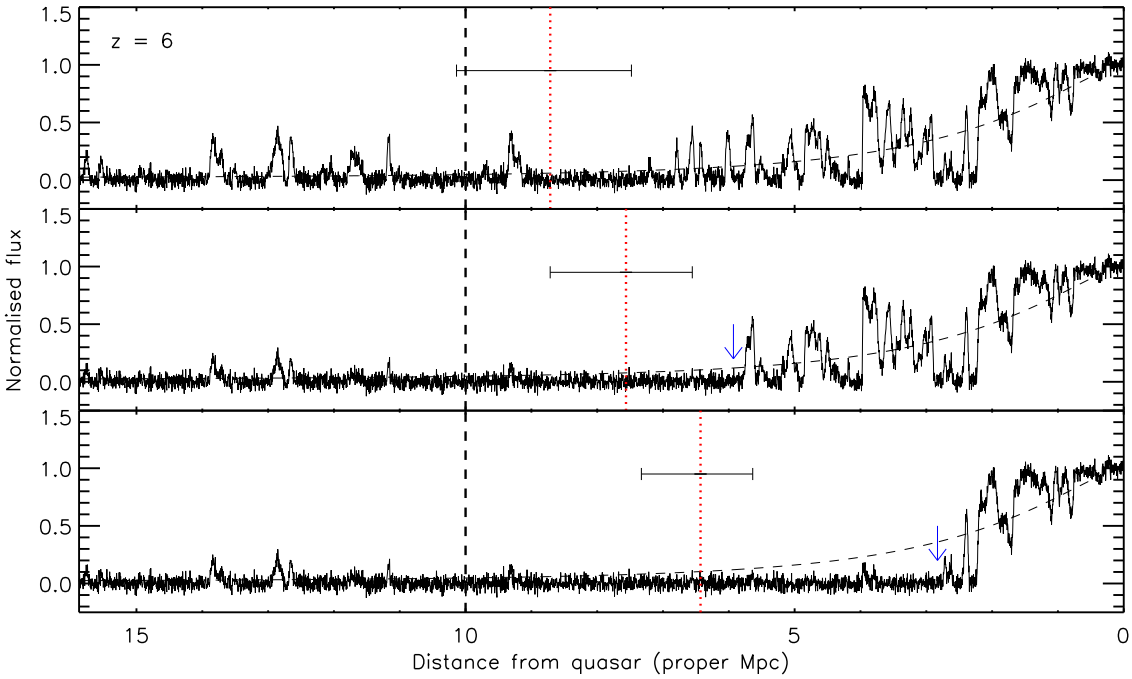


Figure 5.4: Demonstration of the impact of Lyman limit systems. *Top panel:* A standard simulated spectrum at $z = 6$ with $R_{\rm eq} = 10$ proper Mpc and S/N = 20. The thick black dashed line marks this input $R_{\rm eq}$ whilst the dotted red line shows the one estimated by our method. The error bars correspond to the statistical error in the estimated $R_{\rm eq}$, roughly 15%. The thin black dashed line is the expected drop in normalised flux for the input $R_{\rm eq}$. The method has, within error, recovered the input $R_{\rm eq}$. *Middle panel:* The same spectrum as above, but with a Lyman limit system with corresponding Lyα absorption at the location of the blue arrow. The region of enhanced transmitted flux is slightly truncated, and so the estimated $R_{\rm eq}$ is slightly smaller than in the case of no Lyman limit system, corresponding to a higher estimated value of $\Gamma_{\rm bkg}$. This is a well-known bias in proximity-effect measurements. *Bottom panel:* The same spectrum as in the top panel, but with the Lyman limit system much closer to the quasar, causing substantial shortening of the apparent size of the proximity region. There is again an associated decrease in the estimated $R_{\rm eq}$, and thus a higher estimated $\Gamma_{\rm bkg}$. However, the estimated values of $R_{\rm eq}$ are substantially more robust than would be expected from the extreme shortening of the apparent extent of the enhanced transmitted flux.

5.4 Effect of Lyman limit systems

Lyman limit systems (LLSs) are regions of neutral gas that are optically thick ($\tau > 1$) to Lyman limit photons ($\lambda_{\rm rest} = 912$ Å), having a neutral hydrogen column density of $N_{\rm HI} > 1.6 \times 10^{17}$ atom cm^{-2}. These are important for setting the intensity of the UVB, as their abundance limits the mean free path of ionizing photons (e.g. Storrie-Lombardi et al., 1994; Miralda-Escudé, 2003; Péroux et al., 2003; Furlanetto and Mesinger, 2009). If a LLS lies along the line of sight to a quasar it can truncate the extent of enhanced transmission in the spectrum. Fig. 5.4 demonstrates the case where a LLS has shortened the apparent proximity region. As the shortening becomes more severe, our method underestimates the true proximity region size [and thus overestimates the value of $\log(\Gamma_{\rm bkg})$] because the LLS modifies the assumed $1/r^2$ fall-off from the quasar.

Studies counting the number of LLSs in spectra struggle at high redshift as, due to the ever increasing blackness of the forest, features with $\tau > 1$ are difficult to detect. Consequently, there are very few

	Redshift range	N_0	$N_{3.5}$	γ	$z=5$			$z=6$		
					$n_{\rm lls}$	ϵ (ind)	ϵ (bin)	$n_{\rm lls}$	ϵ (ind)	ϵ (bin)
SL94	$0.40 < z < 4.69$	$0.27^{+0.20}_{-0.13}$	–	1.55 ± 0.45	0.2169	0.15	0.09	0.3459	0.17	0.08
P03	$2.40 < z < 4.93$	$0.07^{+0.13}_{-0.04}$	–	$2.45^{+0.75}_{-0.65}$	0.2810	0.20	0.12	0.5146	0.26	0.11
SC10	$0 < z < 6$	0.15^{a}	2.80 ± 0.33	$1.94^{+0.36}_{-0.32}$	0.2441	0.17	0.10	0.4134	0.22	0.09
SC10 1σ					0.3022	0.21	0.12	0.5409	0.27	0.12

Table 5.1: Summary of the parameters of the different power laws used to investigate the effect of LLSs on both individual and grouped spectra. The number of expected LLSs in a random sightline through Model C is calculated using Equation 5.1 with $z_{\rm max}$ equal to the simulation redshift and $z_{\rm max} - z_{\rm min}$ corresponds to $20h^{-1}$ comoving Mpc. SC10 parametrise their power law with N_0 evaluated at $z = 3.5$ such that instead of $N(z) = N_0(1+z)^{\gamma}$ they have $N(z) = N_{3.5}[(1+z)/4.5]^{\gamma}$. The shifts, ϵ, quoted are the average overestimation in $\log(\Gamma_{\rm bkg})$ due to the presence of LLSs. For the individual spectra, those at $z = 6$ have HIRES resolution, and those at $z = 5$ have MIKE resolution. For the grouped spectra the shifts correspond to the $z \sim 5$ and $z \sim 6$ redshift bins, as described in Section 4.4.1. The final row represents the results based upon a power law consistent with the 1 σ upper limit of the SC10 parameters.
[a] Derived from their $N_{3.5}$ and γ values.

studies that provide measured LLS frequency at the redshifts covered in Chapter 4. LLSs can lead to dramatic shortening of the region of enhanced transmission (see Fig. 5.4) and so, if abundant at high redshift, they could be a substantial source of systematic errors.

We inserted LLSs into our simulations following the method presented in Bolton and Haehnelt (2007b). A density threshold, $\rho_{\rm thresh}$, was chosen such that the average number of regions in a spectrum with $\rho \geq \rho_{\rm thresh}$ was the same as a chosen number of LLSs expected per sightline. The neutral fraction within those regions was then set to unity (i.e. they became self-shielded). As such, they absorb all the ionizing flux of the quasar; so further out in the spectrum the transmission in the forest is from the UVB alone. The expected number of LLSs in a particular sightline, $n_{\rm lls}$, was calculated from the power laws presented by Storrie-Lombardi et al. (1994), Péroux et al. (2003) and Songaila and Cowie (2010) (hereafter SL94, P03 and SC10, respectively), where $N(z) \equiv dN/dz = N_0(1+z)^\gamma$. SL94 used quasars covering the range $0.40 < z < 4.69$, and suggest that, for a sightline covering the redshift range $z_{\rm max}$ to $z_{\rm min}$,

$$n_{\rm lls} = \int_{z_{\rm min}}^{z_{\rm max}} N_0(1+z)^\gamma \, , \tag{5.1}$$

where $N_0 = 0.27$ and $\gamma = 1.55$. For a particular random sightline in Model C, $z_{\rm max}$ was taken to be the redshift of the simulation and $z_{\rm max} - z_{\rm min}$ corresponds to the simulation box size. This meant that $n_{\rm lls} = 0.2169$ and 0.3459 at redshifts $z = 5$ and 6, respectively. From this (using only the random sightlines) the threshold densities were derived as $\log(\Delta_{\rm thresh}) = 1.6310$ and 1.4080 (where $\Delta_{\rm thresh} \equiv \rho_{\rm thresh}/\langle\rho\rangle$). Regions above these densities were then presumed to be self-shielded. Similar calculations were also done for the power laws of P03 and SC10, and a worst-case scenario was investigated using the 1σ upper limits of SC10. The expected overestimations in $\log(\Gamma_{\rm bkg})$ for each of these power laws are presented in Table 5.1 for both individual spectra and grouped spectra. Consequently, assuming the predictions of SC10 to be the most reliable, as this is the only study probing the redshift range we are interested in, the effect would be to shift the results of the individual quasars down by 0.17 (0.22) dex at $z = 5$ (6), or 0.21 (0.27) dex in the 1σ limiting case.

SC10 is the only study into the actual spatial frequency of LLSs at $z = 6$. In both SL94 and P03 they found that towards higher redshifts, $N(z)$ starts to evolve rapidly, and so extrapolations become much more uncertain. Prochaska et al. (2010) note that both of those papers are subject to biases that may not have been adequately compensated for, and so they may have overestimated $N(z)$ by up to a factor of 3. Clearly if $n_{\rm lls}$ is a factor of 3 smaller then the effect on $\Gamma_{\rm bkg}$ would be greatly reduced. SC10 themselves also warn against extrapolating to $z > 6$ due to the potential rapid change in the mean free path of ionizing photons if reionization is being probed. Taking all these factors into account, and treating SC10 as the most reliable power law, we cautiously estimate using their 1σ upper limits on $n_{\rm lls}$, that the binned data points in Fig. 4.10 are possibly overestimated by ≤ 0.12 dex.

Judging by visual inspection, none of our observed spectra appears to be influenced by LLSs to the degree presented in the bottom panel of Fig. 5.4. It is, however, difficult to determine if any of the sightlines contained LLSs far from the quasar redshift, as in the middle panel of Fig. 5.4. SDSS J1148+5251 has a very short region of enhanced transmitted flux (Fig. 4.9) and is a possible candidate for being affected by a LLS.

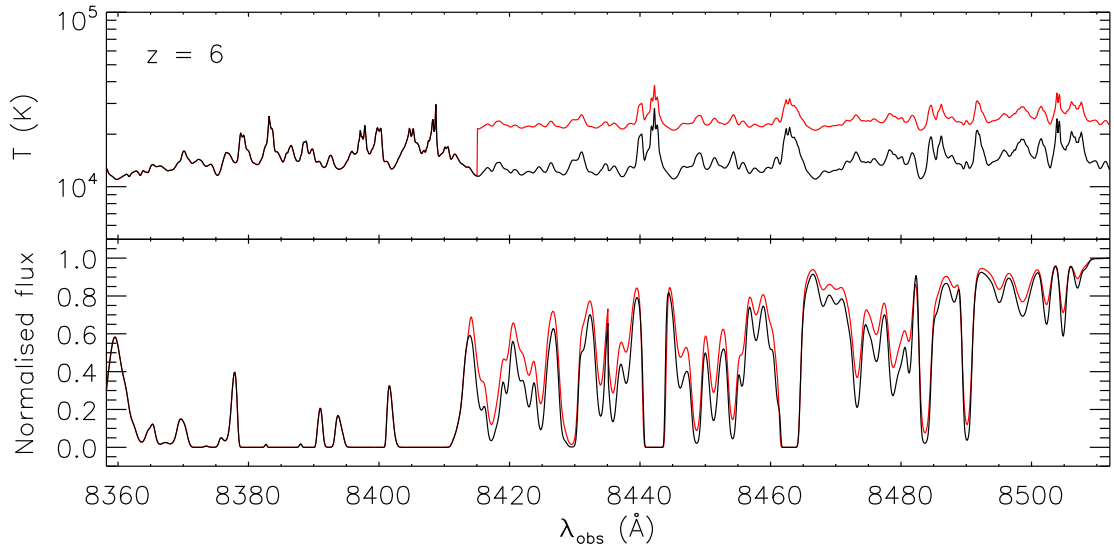


Figure 5.5: Demonstration of a thermal proximity effect within the classical proximity effect. *Top panel:* The IGM temperature for a $40h^{-1}$ sightline in Model C. The black line is the underlying temperature field, whilst the red line shows our toy model where the temperature within the 5 proper Mpc closest to the quasar has been raised by 10^4 K, similar to what is expected if He II in that region had been ionized. *Bottom panel:* The resulting spectra for both a classical proximity effect that ignores heating (black line), and one that includes the heating in our toy model (red line). The transmission within the closest 5 Mpc is higher for our toy model, so the derived proximity size will be greater, and thus $\Gamma_{\rm bkg}$ will be underestimated.

5.5 Thermal proximity effects

We have implicitly assumed that the temperature of the quasar proximity region is comparable to that of the general IGM. This may be not true, particularly prior to He II reionization when the ionization of He II by the quasar may heat the local IGM (e.g Bolton et al., 2010). This temperature gradient will cause higher mean flux in the proximity region, and so the proximity region size will be *overestimated* and the UVB *underestimated* (i.e. opposite to all the other environmental biases). We constructed a toy model in the simulations with the change in temperature as a step function, such that the gas within 5 proper Mpc of the quasar was 10^4 K hotter than that in the general IGM (see Fig. 5.5). This is similar to the result of Bolton et al. (2010), assuming general IGM temperatures at $z = 5-6$ of $\sim 10^4$ K (Becker et al., 2011a). In individual sightlines this caused $\log(\Gamma_{\rm bkg})$ to be underestimated by 0.10 (0.11) dex at $z = 5$ (6), and in the binned data by 0.14 (0.21). Consequently, we find that the hotter temperatures of gas close to the quasar could cause an underestimation in $\log(\Gamma_{\rm bkg})$ comparable to the overestimations from other environmental effects at the redshifts of our study.

5.6 Conclusions

In Chapter 4 we presented the intensity of the UVB at $z \sim 5-6$ as measured from the proximity effect in quasar spectra. Here, we have investigated the sensitivity of these results to sources of systematic bias from the environment of quasars.

Errors in the systemic redshift are of greater importance than errors in the ionizing luminosity caused by an unknown UV continuum power-law slope. An error in the assumed luminosity of the quasar of up to 40 per cent would still have a smaller effect than the typical redshift errors. Also, proximity regions are of order ~ 10 Mpc at $z \gtrsim 5$, which is larger than the shift from even substantial errors in the systemic redshift.

Systematic uncertainties caused by the overdensity of matter close to the quasar were found to be smaller at $z > 5$ than at lower redshifts. The effect of the quasars lying in large-scale overdensities was found to be small for host halo masses of $M \leq 10^{12.2}\mathrm{M}_{\odot}$, the most massive haloes in our largest ($80h^{-1}$ comoving Mpc) simulation. Even in a toy model emulating the density distribution surrounding a $10^{13}\mathrm{M}_{\odot}$ halo at $z = 6$ (and therefore as massive as the most massive haloes in the Millennium Simulation at this redshift), large-scale overdensities do not dominate the error budget. This is due to the proximity regions (as defined in Chapter 4) being substantially larger than the scale of the overdensities. It is however noted that at lower redshifts, where the UVB intensity and the mean flux level are higher, proximity regions may well be of similar size to the local overdensity. This effect is therefore likely to be important at $z \sim 2-3$, where discrepancies between measurements from flux decrements and the proximity effect have been noted (Rollinde et al., 2005; Guimarães et al., 2007).

Clustering of ionizing sources such as low-luminosity galaxies around quasars, as expected if the quasars reside in massive dark matter haloes, was also investigated. For realistic assumptions about the number and brightness of such sources, the quasar still dominates the ionizing photon budget within the proximity region. Consequently, the effect on our estimates of Γ_{bkg} from such galaxies is likely to be small.

We also tested the effect of LLSs. Using the recent measurements of the frequency of LLSs at high redshift from SC10, we found that LLSs also only have a small effect on our analysis, potentially causing us to overestimate Γ_{bkg} in our combined samples at $z \sim 5$ and 6 by roughly 0.1 dex.

Our measurements further assume that the gas temperature within the proximity region is similar to that in the general IGM. If present, a thermal proximity effect could have a sizeable impact on our estimates of the UVB intensity. For example, if the gas in the ~ 5 proper Mpc of the proximity region closest to the quasar were on average 10^4 K hotter than the general IGM, then the UVB will be underestimated by 0.14 (0.21) dex at $z = 5$ (6). This is comparable in magnitude to the biases from other environmental factors, although opposite in sign.

Overall we find that the environmental biases that plague proximity-effect measurements of Γ_{bkg} at lower redshift only produce small biases at $z \gtrsim 5$. Consequently, our high-redshift estimates of Γ_{bkg} from the proximity effect presented in Chapter 4 are likely to be robust. Differences between our estimates and those from flux decrements are therefore unlikely to be due to the quasar environment, and can instead be linked to assumptions used in the conversion of a flux decrement into a photoionization rate, such as the IGM temperature at the mean density.

Every day I go to work I'm making a bet that the universe is simple, symmetric, and aesthetically pleasing - a universe that we humans, with our limited perspective, will someday understand.

GEORGE SMOOT

6

Measurements of the IGM temperature at $3.8 < z < 5.8$

THE temperature of the intergalactic medium (IGM) provides a unique insight into the process behind the reionization of the Universe. Here we present the first measurements of the IGM temperature near the mean density, T_0, for $z > 5$ from high-resolution spectra of the Lyα forest. The small-scale structure of the forest is quantified using two independent statistics; the curvature, which has been successfully applied at $z < 5$, and the "peakiness", which we have developed for analysing the $z > 5$ forest. Both indicate a possible increase in T_0 between $z \sim 5$ and 6; however, the amplitude of this increase is sensitive to the smoothness of the underlying gas density distribution, which in turn depends on the thermal history spanning hydrogen reionization. Earlier or more extended reionization histories produce smoother gas distributions at $z \sim 5-6$, which can mimic the effects of higher instantaneous temperatures in the Lyα forest. If there is a sufficiently early period of reionization, then correcting for this effect can give temperatures that actually decrease from $z \sim 5$ to 6. The temperatures as derived by the curvature, after correction, are consistent with a wide range of possible thermal histories. This emphasises the need for independent measures of the smoothness of the gas density distribution at high redshift. On the other hand, the temperatures measured using the peakiness at $z \sim 4-5$, even with minimal corrections, are consistent with those expected if HI reionization was carried out by sources with soft spectra, such as Pop II stars. Large temperature corrections would lower these temperatures, requiring the spectra of the ionizing sources to be implausibly soft. As such, small corrections are preferred. Since reionization is likely to be extended, this suggests a late reionization scenario. When corrected for such a scenario, the temperatures at $z > 5$ drop by a factor of ~ 2 from $z = 5.6$ to $z = 4.8$, suggesting we may be witnessing the adiabatic cool-

ing expected after the end of hydrogen reionization. Assuming this temperature evolution continues to higher redshift, reasonable peak temperatures expected at the end of reionization are reached at $z \sim 6-7$. This implies the end of reionization may have occurred in this redshift range.

6.1 Introduction

Cosmic reionization was the last major phase change in the Universe's history. Transmission in the Gunn-Peterson troughs (Gunn and Peterson, 1965) of quasar spectra indicate that hydrogen was reionized at $z > 6$ (Fan et al., 2006c), and that helium is fully ionized at $z \lesssim 3$ (Kriss et al., 2001; Worseck et al., 2011). Photoionization heats the gas in the intergalactic medium (IGM), and so the thermal history of the IGM provides important constraints on the duration of reionization, as well as the ionizing sources (Miralda-Escudé and Rees, 1994; Gnedin, 2000; Hui and Haiman, 2003; Tittley and Meiksin, 2007).

Prior to hydrogen reionization, the temperature of the IGM was determined by adiabatic cooling ($T \propto (1+z)^2$) due to cosmic expansion, likely reaching ~ 30 K before the first structures formed. However, ionizing photons from the first stars and galaxies, responsible for the reionization of hydrogen, photoheated the IGM to a peak temperature of $\sim 1.5-3\times10^4$ K, as constrained by both theoretical motivations and simulations of the formation of cosmic structure (Miralda-Escudé and Rees, 1994; Hui and Gnedin, 1997; Hui and Haiman, 2003; Furlanetto and Oh, 2009; Benson and Bower, 2010). A similar maximum temperature is expected during He II reionization by quasars (Abel and Haehnelt, 1999; Gleser et al., 2005; Furlanetto and Oh, 2008b; Bolton et al., 2009; McQuinn et al., 2009). These peak temperatures occur as the ionized fraction approaches unity (e.g. McQuinn et al., 2009), and once reionization has finished the IGM cools adiabatically once more.

Since H I and He II reionization occur at distinct redshifts, each should be responsible for a recognisable peak in the IGM thermal history, corresponding roughly to its conclusion. The temperature of the IGM affects many of the observable properties of the Lyα forest, such as the width of individual absorption features, the power spectrum of the transmitted flux, the flux probability distribution function, and the effective optical depth (Hui and Gnedin, 1997; Haehnelt and Steinmetz, 1998; Dijkstra et al., 2004; Paschos et al., 2007; Bolton and Haehnelt, 2007b; Cen et al., 2009). The lower redshift of He II reionization means there are a large number of quasar sightlines that illuminate the IGM in the redshift range of interest. There have been several studies carried out on the $z \sim 2-4$ Lyα forest to detect this thermal evolution (Ricotti et al., 2000; Schaye et al., 2000; Theuns and Zaroubi, 2000; McDonald et al., 2001; Zaldarriaga et al., 2001; Theuns et al., 2002a,b; Zaldarriaga, 2002; Lidz et al., 2006a; Viel et al., 2009; Lidz et al., 2010; Becker et al., 2011a), although with somewhat mixed success.

Attempts to carry out similar measurements at $z > 6$ to detect the peak associated with hydrogen reionization are hampered by both a steep drop in the number density of bright quasars at high redshift (e.g. Hopkins et al., 2007), and the decreasing transmission of the Lyα forest (Songaila, 2004). Consequently, other probes have been used. After the temperature boost that accompanies reionization, the IGM initially cools adiabatically ($\propto (1+z)^2$), before reaching equilibrium with

sources of photoheating and cooling at a slower rate ($\propto (1+z)^{0.53}$). This well-defined late-time asymptote is essentially independent of the details of reionization and only dependent on the spectra of the ionizing sources (Miralda-Escudé and Rees, 1994; Hui and Gnedin, 1997; Theuns et al., 2002a; Hui and Haiman, 2003). Thus by measuring the temperature at redshifts approaching that of hydrogen reionization, one can learn about the sources responsible. If the measured temperature is higher than the asymptote for a reasonable ionizing spectrum, this can be used to place some upper limits on when it happened (Hui and Haiman, 2003). Indeed, in lieu of direct temperature measurements, the Lyα forest at $z \sim 4-5$ may provide the best restrictions on the ionizing sources that caused H I reionization, as there are a sufficient number of quasar sightlines for statistically significant constraints, as well as relatively high transmitted flux, and little contamination from He II reionization (Cen et al., 2009).

The main obstacle that hinders direct temperature measurements from the Lyα forest at $z \gtrsim 5$ is the decreasing transmission. The near-zones of quasars, however, have enhanced transmission compared to the forest. This allows the widths of individual absorbers to be measured, as can be done in the forest at lower redshift, and so in principle enables a temperature measurement, even at $z \gtrsim 6$ for which the forest is saturated (Bolton and Haehnelt, 2007b). Bolton et al. (2010) used this technique to measure the IGM temperature at mean density, T_0, within 33 comoving Mpc of a $z \sim 6$ quasar to be $T_0 = 23\,600 \pm^{5000}_{6900}$ K. This rather high temperature includes a boost due to local He II reionization and is suggestive of a relatively recent period of reionization. Since it is from just a single quasar, however, it is hard to draw conclusions about the global reionization history. Future measurements from the near-zones of other bright high-redshift quasars may be the best way of determining the thermal state of the IGM at $z \gtrsim 6$.

For the regime $z \leq 5$, Becker et al. (2011a) present some of the most robust IGM temperature measurements to date. They use a statistic called the curvature, which is well suited to quantifying the small-scale structure of the Lyα forest. Importantly, they only measure the temperature at a well-chosen overdensity, rather than trying to constrain the entire temperature-density ($T-\Delta$) relation. Assuming a $T-\Delta$ relation of the form $T(\Delta) = T_0 \Delta^{\gamma-1}$, where T_0 is the temperature at the mean density and $\Delta = \rho / \langle \rho \rangle$ (Hui and Gnedin, 1997), they can convert their temperatures into T_0 for an assumed γ. They find that T_0 increases from $\sim$ 8000 K at $z = 4.4$ (assuming $\gamma \sim 1.5$) to $\sim$ 15 000 K at $z = 2.8$ (assuming $\gamma = 1.3$, which is appropriate for the end of He II reionization; McQuinn et al., 2009), followed by cooling if γ returns to the asymptotic value of $\sim$ 1.5. This thermal evolution is consistent with theoretical predictions of the peak in T_0 expected due to He II reionization, and suggests that the process was rather extended. It should be noted that during reionization γ could be < 1.3, and the $T-\Delta$ relation may not even be monotonic if the contributions of high-energy blazars are significant (Broderick et al., 2011; Chang et al., 2011; Puchwein et al., 2011). There is some observational evidence for an inverted $T-\Delta$ relation at these redshifts (Becker et al., 2007; Bolton et al., 2008; Viel et al., 2009), so the peak T_0 during He II reionization may be substantially higher.

In this Chapter, we extend the use of the curvature statistic to higher redshift to look for similar thermal evolution due to H I reionization. We also introduce an independent statistic called the peakiness. This is less precise than the curvature as it can only take information from rare features in

the forest, but it does not require a continuum normalised spectrum, sidestepping one of the most challenging issues posed by the $z > 5$ Lyα forest.

The rest of this Chapter is organised as follows. In Section 6.2 we describe the observational data and the hydrodynamical simulations used. In Section 6.3 we outline the curvature and peakiness statistics, as well as their sensitivity to the underlying density field and Jeans smoothing. We present our results using both statistics and discuss the implications for hydrogen reionization in Section 6.4. Finally, we give our conclusions in Section 6.5. Throughout this paper we use the cosmological parameters taken from the mean of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 5-yr data set (Komatsu et al., 2009), with Hubble constant $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and density parameters $(\Omega_\mathrm{m}, \Omega_\Lambda) = (0.26, 0.74)$.

6.2 Data and models

6.2.1 Observed spectra

The data consists of 23 high-resolution quasar spectra, taken with either the Keck or Magellan telescopes. Almost all of those at $z > 5$ were observed with the High Resolution Echelle Spectrograph (HIRES; Vogt et al., 1994) on the 10-m Keck I telescope. The spectra were reduced using a custom set of IDL routines and optimal sky-subtraction techniques (Kelson, 2003), as detailed in Becker et al. (2006, 2007). The observations were made with the 0.86-arcsec slit with $R = 40\,000$, so the full width at half-maximum (FWHM) velocity resolution is $6.7\,\mathrm{km\,s^{-1}}$. The pixel resolution is $2.1\,\mathrm{km\,s^{-1}}$.

Most of the spectra at $z < 5$ were observed with the Magellan Inamori Kyocera Echelle spectrograph (MIKE; Bernstein et al., 2003) on the 6.5-m Magellan-II Clay telescope, and reduced with a similar custom pipeline. A 1-arcsec slit was used with $R = 22\,000$ in the red, so the FWHM velocity resolution is $13.6\,\mathrm{km\,s^{-1}}$. The pixel resolution for this instrument is $5.0\,\mathrm{km\,s^{-1}}$. The list of targets is given in Table 6.1.

The spectra were all normalised using the procedure described in Bolton et al. (2010). The spectrum is first divided through by a power law $f_\nu \propto \nu^{-0.5}$, normalised at $1280(1+z)$ Å, and the Lyα emission line is then fitted with a slowly varying spline. It is very difficult to fit the continuum over the forest at $z > 5$ due to the low flux levels, and so the sensitivity to continuum placement is discussed in Section 6.3.4. The $5000\,\mathrm{km\,s^{-1}}$ bluewards of the Lyα transition and the $2000\,\mathrm{km\,s^{-1}}$ redwards of the Lyβ transition were excluded from the analysis, with the intervening Lyα forest divided up into consecutive sections $10h^{-1}$ comoving Mpc long. Out of these short sections, only those with S/N > 5 were used.

6.2.2 Simulated spectra

A total of 17 cosmological hydrodynamical simulations were used for this work, and they are summarised in Table 6.2. The majority of them are the same as those described in Becker et al. (2011a) and were performed using a customised version of the parallel Tree-SPH code `GADGET-3`, an updated

Name	z_q	Inst.	Dates	t_{exp} (hrs)	Ref.	S/N
SDSS J1148+5251	6.42	HIRES	Jan 2005 - Feb 2005	14.2[a]	1	10.1
SDSS J1030+0524	6.31	HIRES	Feb 2005	10.0	1	6.5
SDSS J1623+3112	6.25	HIRES	Jun 2005	12.5	1	5.7
SDSS J1048+4637	6.23	HIRES	Feb 2006	5.0	5	5.8
SDSS J0818+1722	6.02	HIRES	Feb 2006	8.3	2	8.4
SDSS J1306+0356	6.02	MIKE	Feb 2007	6.7	4	6.9
SDSS J0002+2550	5.82	HIRES	Jan 2005 - Jul 2008	14.2	1,4	9.3
SDSS J0836+0054	5.81	HIRES	Jan 2005	12.5[a]	1	10.0
SDSS J1044-0125	5.80	MIKE	Feb 2007 - Mar 2008	7.5	5	8.6
SDSS J0231-0728	5.41	HIRES	Jan 2005 - Feb 2005	10.0	1	7.7
SDSS J1659+2709	5.33	HIRES	Sep 2007 - Jul 2008	11.7	3	19.2
SDSS J0915+4924	5.20	HIRES	Feb 2005	10.0	1	11.8
SDSS J1204-0021	5.09	HIRES	Jan 2005 - Feb 2005	6.7	1	8.5
SDSS J0040-0915	4.98	MIKE	Oct 2007	8.3	3	24.4
SDSS J0011+1440	4.97	HIRES	Sep 2007	6.7	3	25.0
SDSS J2225-0014	4.89	MIKE	Oct 2007	5.0	4	11.6
SDSS J1616+0501	4.88	MIKE	Mar 2008	3.3	4	11.1
SDSS J2147-0838	4.59	MIKE	Oct 2007	8.3	3	25.6
BR 0353-3820	4.59	MIKE	Nov 2006 - Oct 2007	9.2	3	43.5
BR 0006-6208	4.52	MIKE	Nov 2006	1.7	6	8.0
BR 1033-0327	4.52	MIKE	Nov 2005	4.2	6	10.5
BR 0714-6449	4.49	MIKE	Feb 2007	6.7	3	20.4
BR 0418-5726	4.48	MIKE	Nov 2005 - Nov 2006	15.8	3	28.6

Table 6.1: The list of quasars included in this Chapter. Columns give the quasar name and redshift, details of the observations and the average signal-to-noise ratio (S/N) per pixel in the $10h^{-1}$ comoving Mpc sections of Lyα forest used, after masking out skyline residuals and throwing out sections with S/N ≤ 5 (see Section 6.3.3).
[a] The present reductions include only data taken with the upgraded detector.
References: (1) Becker et al. (2006); (2) Becker et al. (2007); (3) Becker et al. (2011a); (4) Calverley et al. (2011); (5) Becker et al. (2011b); (6) This work.

version of the publicly available code `GADGET-2` (Springel, 2005). The simulations assume the cosmological parameters $(\Omega_m, \Omega_\Lambda, \Omega_b h^2, h, \sigma_8, n_s) = (0.26, 0.74, 0.023, 0.72, 0.80, 0.96)$, have both dark matter and gas components and were started at $z = 99$ with initial conditions generated with the transfer function of Eisenstein and Hu (1999). The baryons in the simulations are photoionized and heated by a spatially uniform ultraviolet background (UVB), and are in ionization equilibrium. Each simulation uses the UVB model of Haardt and Madau (2001) with contributions from both galaxies and quasars. The UVB is switched on at $z = 9$ and applied in the optically thin limit, except for models TZ12 and TZ15 where the UVB is turned on at $z = 12$ and $z = 15$ respectively. The particle mass resolution is chosen specifically to resolve the Lyα forest at high-redshift (Bolton and Becker, 2009). Star formation is included, where all particles with an overdensity $\Delta > 10^3$ and $T < 10^5$ K are converted into collisionless stars. The IGM is assumed to consist of primordial gas with a helium fraction by mass of $Y = 0.24$ (Olive and Skillman, 2004). Simulation outputs were at $1.831 \leq z \leq 6.010$, although only those at $z \geq 3.713$ were used for the analysis of the data. Gas with $\Delta \lesssim 10$ follows a tight temperature-density relation of the form $T = T_0 \Delta^{\gamma-1}$, where T_0 is the temperature at the mean density and γ is the

slope of the relation (Hui and Gnedin, 1997).

The implicit Haardt and Madau (2001) photoheating rates of the simulations, $\epsilon^{\rm HM01}$, were rescaled such that $\epsilon = \zeta\Delta^{\xi}\epsilon^{\rm HM01}$, where ζ and ξ are the constants shown in Table 6.2, in order to have a variety of thermal histories with different $T_0(z)$ and $\gamma(z)$. The fiducial γ was ~ 1.5 (models A15-G15), although flattened ($\gamma \sim 1.3$), isothermal ($\gamma \sim 1.0$), and inverted ($\gamma \sim 0.7$) temperature-density relations were also investigated. Additional models (T15S, T15F, TZ9, TZ12 and TZ15) use a density-independent scaling which varies with redshift. All feature heating due to both HI and HeII reionization, and so have thermal histories that are significantly different from that of the HM01 reference model. The thermal history due to HI reionization is the most relevant in this work. In T15S and T15F the reionization is instantaneous at $z = 9$, whilst in TZ9, TZ12 and TZ15 it is more extended, starting at a variety of redshifts. This allows us to test the impact of the thermal history on the results, and is discussed in detail in Section 6.3.7.

Simulated Lyα spectra were constructed from line-of-sight density, peculiar velocity, neutral hydrogen fraction and temperature fields, drawn for 1024 random sightlines parallel to the box boundaries (e.g. Theuns et al., 1998). Each sightline is 1024 pixels long. The spectra were convolved with a Gaussian with FWHM equal to the velocity resolution of the instrument that was being modelled (6.7 $\mathrm{km\,s^{-1}}$ for HIRES and 13.6 $\mathrm{km\,s^{-1}}$ for MIKE), before being resampled at the instrument pixel resolution (2.1 and 5.0 $\mathrm{km\,s^{-1}}$, respectively). Gaussian-distributed noise could then be added.

6.3 Method

6.3.1 The curvature statistic

Becker et al. (2011a) introduced a new statistic to measure the small-scale structure of the Lyα forest, called the curvature, κ. It is defined as

$$\kappa \equiv \frac{F''}{\left[1+(F')^2\right]^{3/2}}, \tag{6.1}$$

where the flux derivative is calculated with respect to the velocity separation between pixels, measured in $\mathrm{km\,s^{-1}}$. There are few rapid changes in the flux from pixel to pixel, so F' is normally small. The curvature can then be well approximated as $\kappa \approx F''$. The small-scale structure in the forest is quantified by computing the mean absolute curvature, $\langle|\kappa|\rangle$, among pixels within a fixed range of flux (see Section 6.3.4), for a stretch of forest $10h^{-1}$ comoving Mpc long. This is then converted into a temperature by comparison to the simulations.

The main advantages of using the curvature statistic are that it is very quick to compute for an entire stretch of forest, and does not rely upon the identification of individual absorption lines, which is difficult at $z > 4$ due to line-blending. The main drawbacks are that κ will be dominated by the presence of noise in the data, and is linearly dependent on the amplitude of the flux, and thus on the fitted continuum, which is highly uncertain at $z > 5$. Both of these issues are discussed further in Sections 6.3.3 and 6.3.4 respectively.

Model	L $[h^{-1}\,\mathrm{Mpc}]$	N	M_{gas} $[h^{-1}\,\mathrm{M}_{\odot}]$	ζ	ξ	$T_0^{z=3}$ [K]	$\gamma^{z=3}$	Ref.
A15	10	2×512^3	9.2×10^4	0.30	0.00	5100	1.52	1
B15	10	2×512^3	9.2×10^4	0.80	0.00	9600	1.54	1
C15	10	2×512^3	9.2×10^4	1.45	0.00	14000	1.54	1
C10	10	2×512^3	9.2×10^4	1.45	-1.00	13700	1.02	1
D15	10	2×512^3	9.2×10^4	2.20	0.00	18200	1.55	1
D13	10	2×512^3	9.2×10^4	2.20	-0.45	18100	1.32	1
D10	10	2×512^3	9.2×10^4	2.20	-1.00	18000	1.03	1
D07	10	2×512^3	9.2×10^4	2.20	-1.60	17900	0.71	1
E15	10	2×512^3	9.2×10^4	3.10	0.00	22500	1.55	1
E10	10	2×512^3	9.2×10^4	3.10	-1.00	22200	1.04	1
F15	10	2×512^3	9.2×10^4	4.20	0.00	27000	1.55	1
G15	10	2×512^3	9.2×10^4	5.30	0.00	31000	1.55	1
T15S	10	2×512^3	9.2×10^4	Varied	0.00	18200	1.53	1
T15F	10	2×512^3	9.2×10^4	Varied	0.00	18600	1.54	1
TZ9	10	2×512^3	9.2×10^4	Varied	0.00	17800	1.56	2
TZ12	10	2×512^3	9.2×10^4	Varied	0.00	17500	1.57	2
TZ15	10	2×512^3	9.2×10^4	Varied	0.00	17500	1.57	2

Table 6.2: The properties of the hydrodynamical simulations used in this Chapter. Columns give the model name, the size of each box in comoving Mpc, the total particle number N, the particle mass resolution, the scaling factors for the UVB photoheating rates (see the main text for details), the median volume-weighted gas temperature at the mean density, T_0, and power-law slope of the temperature-density relation, $T = T_0\Delta^{\gamma-1}$, at $z = 3$, and references to the paper in which they are first described. The values of T_0 and γ are given to three significant figures.
References: (1) Becker et al. (2011a); (2) This work.

Fig. 6.1 demonstrates the sensitivity of the curvature to the temperature of the IGM. We plot noiseless simulated spectra at $z = 4.282$ and $z = 5.260$ drawn from the B15 and D15 simulations, where the temperature at mean density is different by a factor of two, but the slope of the $T-\Delta$ relation is roughly the same. For both redshifts there is a marked difference between the colder and hotter simulations in their flux profile, with greater contrast between maxima and minima in the former. This corresponds to high curvature values in the colder simulation.

6.3.2 The peakiness statistic

To complement the temperature measurements from the curvature, we introduce a new independent statistic for measuring small-scale structure in the Lyα forest, called the 'peakiness'. For a feature within the forest that contains two maxima with fluxes $F_{\mathrm{max},1}$ and $F_{\mathrm{max},2}$ separated by x_{sep} $\mathrm{km\,s^{-1}}$, and a minimum between them with flux F_{min}, then the peakiness, ϖ, is

$$\varpi = \frac{F_{\mathrm{max,av}} - F_{\mathrm{min}}}{F_{\mathrm{max,av}}}\,, \quad \text{where} \quad F_{\mathrm{max,av}} = \frac{F_{\mathrm{max},1} + F_{\mathrm{max},2}}{2}\,. \tag{6.2}$$

It is essentially the relative depth of a minimum compared to the average height of the maxima that enclose it. The fluxes of the maxima and of the pixels inbetween must all lie within a fixed range

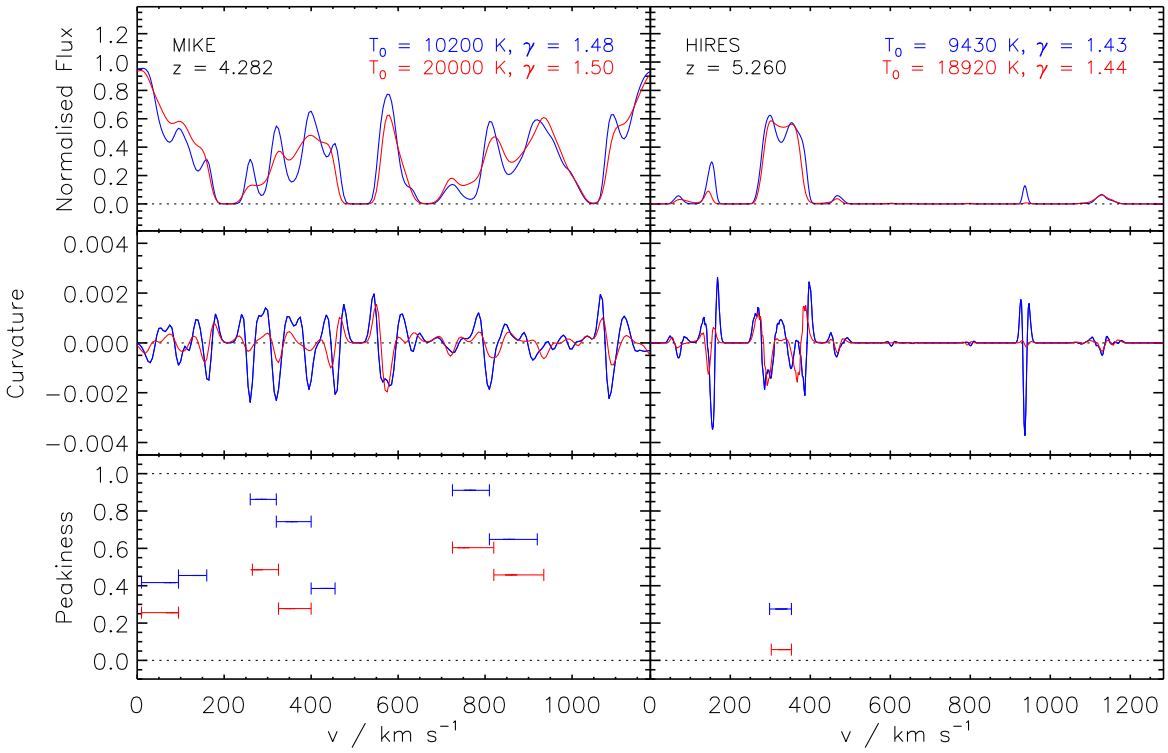


Figure 6.1: A demonstration of the curvature and peakiness statistics to the temperature of the IGM. *Top panels:* Noiseless simulated Lyα forest, $10h^{-1}$ comoving Mpc in length, drawn at $z = 4.282$ (left) and $z = 5.260$ (right). The same sightline was taken from the B15 (blue) and D15 (red) simulations, which have T_0 that differ by a factor of two, but a similar γ. *Middle panels:* The curvature for both the hot and cold spectra at each redshift, as calculated by Equation 6.1. The amplitude of the curvature is greater for the colder simulation at both redshifts. *Bottom panels:* The peakiness as measured from maximum-minimum-maximum groupings using Equation 6.2, with the length of each line representing the velocity separation between the two maxima, $x_{\rm sep}$ km s^{-1}. Suitable features for measuring this statistic becoming increasingly rare at high redshift.

of flux (see Section 6.3.4) for the feature to be used. Fig. 6.1 demonstrates how it is calculated from Lyα forest spectra. For each maximum-minimum-maximum grouping in a stretch of forest, both the separation, $x_{\rm sep}$, and peakiness, ϖ, are measured. The average values of both parameters from all the features found are then compared to simulations to convert them into a temperature. We demonstrate where individual features lie for a $x_{\rm sep}$ and ϖ measured from simulated spectra drawn through our coldest and hottest runs in the left-hand panel of Fig. 6.2. In the right-hand panel, it is clear that with a sufficient number of detected features, the mean of each parameter can be used to accurately determine a value of T_0.

Like the curvature, it does not depend upon identification of individual absorption lines, just the overall composite shape. The dependence on identification of well resolved maxima means it is optimised for use at high redshifts, where line-blending is significant. In comparison, at lower redshift, stretches of forest that are relatively flat and with fluxes very close to the continuum would dilute the sensitivity of the statistic. An important advantage of this method over the curvature is that since the peakiness measures the *relative* depth, it is not dependent upon correctly fitting the continuum level,

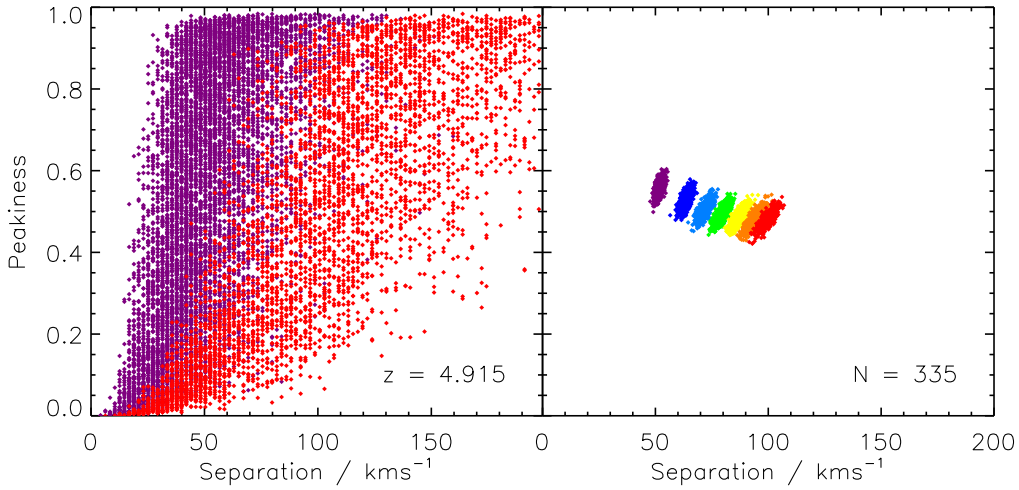


Figure 6.2: Derived values of separation and peakiness for features in the Lyα forest for different values of T_0, from noiseless simulated HIRES spectra at $z = 4.915$. *Left-hand panel:* The location of individual features in a plot of separation against peakiness for the coldest simulation, A15 (purple), and the hottest, G15 (red). For a given separation, the colder features have a higher peakiness. *Right-hand panel:* The mean separation and mean peakiness of a subset of 335 features, extracted at random 1000 times, from simulations A15 (left) to G15 (right). This is the number of features detected in the HIRES spectra at $4.6 < z < 5.0$. Such a sample should allow discrimination between the different T_0, particularly for low temperatures.

removing a source of potentially significant systematic bias at high-redshift (see Section 6.3.4). The peakiness does, however, have some disadvantages too. Unfortunately, like the curvature, it is highly susceptible to noise in the data. This is discussed further in Section 6.3.3. Also, the peakiness can only be measured for features in the Lyα forest with two maxima and a well defined minimum between them. These features become increasingly rare at high redshift, and so the majority of transmission features at $z \gtrsim 5.4$ (which have only a single, well-defined maxima) are ignored. This means temperatures are based on only a very small percentage of pixels for $z \gtrsim 5.4$, from transmission from highly biased sections of the IGM.

6.3.3 Treatment of noisy spectra

As mentioned above, both the curvature and peakiness statistics are dominated by noise, even for spectra with very high S/N. To address this issue we follow the procedure described in Becker et al. (2011a) wherein we fit a smoothly varying b-spline to the flux, and both statistics are then computed from this fit. We ensure that the first and second derivatives of the flux (and hence the curvature) are continuous across the spectrum by using cubic polynomials for the piece-wise fits. We use an adaptive b-spline fitting algorithm, with an initial break point separation of 50 kms^{-1}. Additional break points are added to sections of spectra for which the fit is poor (as determined by a high chi-squared) until the fit converges, or the break point separation matches the instrumental resolution (6.7 kms^{-1} for HIRES and 13.6 kms^{-1} for MIKE). All of the observed spectra were fit in this way, as well as simulated spectra with similar noise properties with which comparisons were made.

Fig. 6.3 demonstrates our b-spline fits to observed Lyα forest spectra. As found by Becker et al.

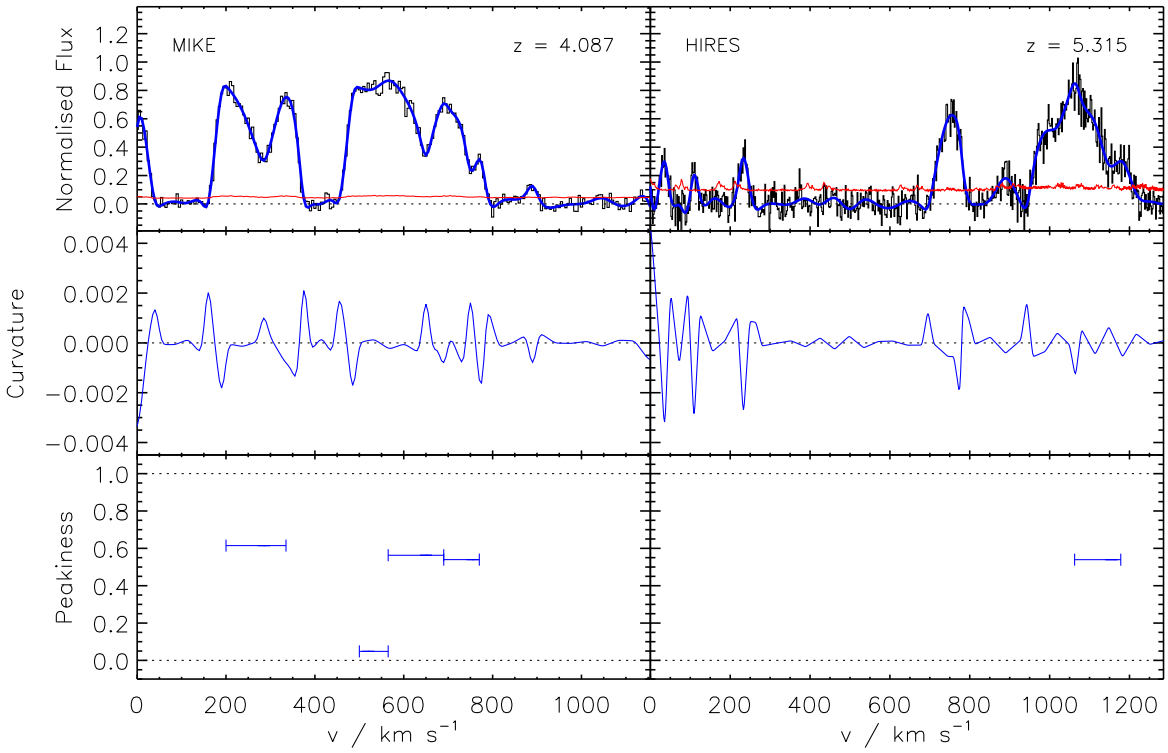


Figure 6.3: *Top panels:* Observed Lyα forest spectra at $z = 4.087$ (left) and $z = 5.315$ (right), taken with MIKE and HIRES respectively. Both sections of forest are $10h^{-1}$ comoving Mpc in length. The histograms are the normalised spectra, the red continuous line is the flux error array, and the blue continuous line is the b-spline fit. *Middle panels:* The curvature computed from the b-spline fits. *Bottom panels:* The peakiness as measured from maximum-minimum-maximum groupings in the b-spline fit.

(2011a), a distinct advantage of an adaptive b-spline fitting scheme is that both statistics are only moderately sensitive to noise, as the underlying shape is recovered well even for relatively noisy data. Despite this, Becker et al. (2011a) limited their analysis to $10h^{-1}$ comoving Mpc sections with errors in the normalised flux of < 0.06 per $2.1\mathrm{km\,s^{-1}}$ bin, in order to maximise the contrast between the curvature from different simulation runs. Such a strict limit on noise, however, would exclude the majority of our Lyα forest spectra at $z > 5$, and so to adequately sample these redshifts we must make use of the noisier data. We therefore instead use the much higher noise limit of < 0.2 per $2.1\mathrm{km\,s^{-1}}$ bin. This allows a far larger number of sightlines to contribute to our highest redshift bins, although the b-spline fit to such noisy spectra is no longer highly constrained. The main areas affected by a poor fit are near zero flux and the continuum, so these areas are masked out before measuring each statistic (see Section 6.3.4).

6.3.4 Mean flux and continuum estimation

Continuum fitting at high redshift is challenging due to the blending of individual absorption lines, leaving few pixels near the continuum. Since the curvature is linearly dependent on the normalised flux, errors in the placement of the unabsorbed continuum could lead to substantial systematic er-

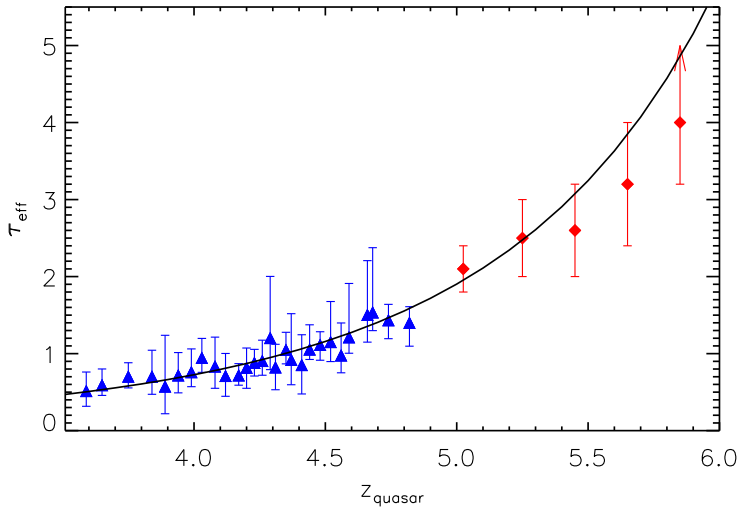


Figure 6.4: Our adopted fit to the evolution of $\tau_{\rm eff}$ with redshift. The blue triangles are from measures of the flux in the forest from Songaila (2004), whilst the red diamonds are the binned values from Fan et al. (2006b). The solid line gives the Becker et al. (2007) relation for the evolution of $\tau_{\rm eff}$ with redshift, which we adopt for this study.

rors. Becker et al. (2011a) dealt with this issue by renormalising the spectra in each $10h^{-1}$ comoving Mpc section of forest, by dividing through by the highest flux in the b-spline fit. They did this for both the real and simulated spectra, and though it lowered their sensitivity to different temperatures, it meant that any continuum placement errors were handled objectively. Unfortunately, the high saturation of the $z > 5$ Lyα forest makes this renormalisation procedure untenable at the highest redshifts. Thus, to keep as much temperature sensitivity as possible, we sacrifice this robust approach to continuum errors, and instead rely upon our initial continua. We note that since the peakiness measures the relative depth of a minimum to the maxima around it, it is unaffected by continuum placement errors. Comparison between the temperatures returned by the two methods will therefore give some indication as to how (in)correct our initial continua may be.

Since our curvature measurements are dependent upon the mean flux, we scale the ionizing background in the simulations until the mean flux matches that observed in the Lyα forest. We use the fit to the measured effective optical depths, $\tau_{\rm eff} = -\ln\langle F\rangle$, from Becker et al. (2007). The fit is reproduced along with some observed values of $\tau_{\rm eff}$ from the literature in Fig. 6.4. It assumes a simple evolution of a lognormal distribution of optical depths and matches the observed values excellently from $2 < z < 6$. It should be noted that unlike Becker et al. (2011a), we make no correction for metal lines. Their presence will lower the mean flux and bias us towards low temperatures. The absorption in the forest at $z > 4$, however, is dominated by Lyα lines, and so any bias from unidentified metal lines will be small.

In order to provide sufficient coverage at high redshift, we had to use sections of forest with S/N as low as 5. A consequence of this is that the b-spline fit to such noisy data is poorly constrained, especially around zero flux and the continuum. Often the fit will contain wide undulations (see Fig. 6.5), so both statistics could be biased to high temperatures. In order to still extract useful information, we

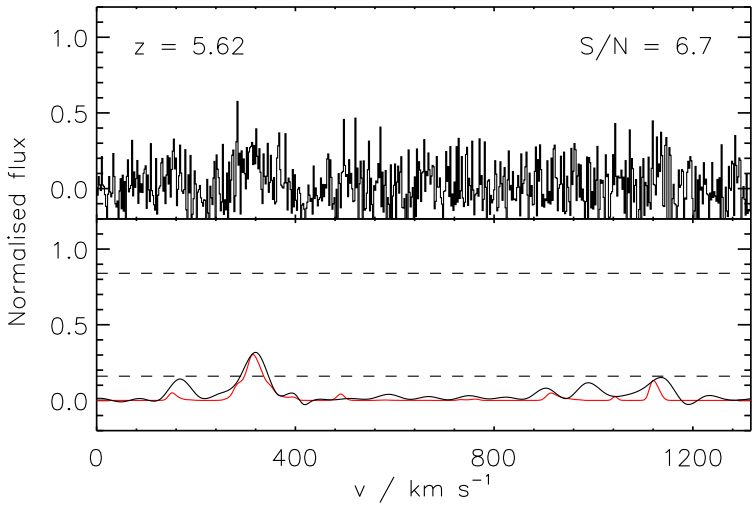


Figure 6.5: *Top panel:* A simulated spectrum taken from the B15 simulation at $z = 5.620$. Gaussian-distributed noise has been added to reduce the S/N ratio to 6.7 (i.e. $\varepsilon = 0.15$). *Bottom panel:* The b-spline fit to the noisy spectrum (black solid line) and the noiseless spectrum (red solid line). Around zero flux the b-spline fit struggles to reproduce the underlying spectrum from such noisy data, and instead contains wide undulations that would bias both the curvature and separation-peakiness statistics to high temperatures. To circumvent this issue we mask pixels with $F < \varepsilon + 0.01$ (lower dashed line) and $F > 1 - (\varepsilon + 0.01)$ (upper dashed line). Although that eliminates many of the pixels at high redshift, those that remain have b-spline fits that recover the underlying shape well. This enables us to use even relatively noisy spectra, as required for good coverage at our highest redshifts.

mask pixels from the fit with a normalised flux of $F < \varepsilon + 0.01$ and $F > 1 - (\varepsilon + 0.01)$, where ε is the rms error in the normalised flux. The additional 0.01 is chosen to provide a conservative buffer against errors in the estimates of the noise from the flux error array. This limits the impact of the b-spline fit's difficulty to recover sections of the spectrum with a very flat profile from noisy data. By comparison, Becker et al. (2011a) mask out pixels with a *renormalised* flux of $F_{\rm renorm} < 0.1$ and $F_{\rm renorm} > 0.9$. Consequently for very high S/N data, we mask fewer pixels. Although, as they point out, pixels close to the continuum or to zero flux contribute little to the curvature measurement, such pixels may contain well-resolved tall maxima or deep minima, and hence will improve the sensitivity of peakiness measurements. For the observed data, ε at each pixel was taken from the flux error array, which we estimate to be generally accurate to $10-20\%$. The flux error array was also used to identify skyline residuals, which were then masked. The mean absolute curvature, $\langle|\kappa|\rangle$, was then measured from all the unmasked pixels in each $10h^{-1}$ comoving Mpc section. The peakiness was only measured for features for which there were no masked pixels between the two maxima.

6.3.5 Identifying optimal overdensities

Care must be taken when measuring the temperature of the IGM, as it varies with density. Photoionization equilibrium of the IGM is reached after reionization, and competing heating and cooling processes produce a tight power-law $T-\Delta$ relation in low density ($\Delta \lesssim 10$) gas of the form $T(\Delta) = T_0 \Delta^{\gamma-1}$, where T_0 is the temperature at mean density (Hui and Gnedin, 1997). Long after reionization, the slope is expected to asymptotically reach $\gamma \approx 1.6$. However, during reionization, the photoheating

is roughly independent of density, and so for an instantaneous and homogeneous reionization, one would expect $\gamma \approx 1$. Reionization is likely an extended and highly inhomogeneous process, though, so γ can only be constrained to $\lesssim 1.6$, and there will be substantial spatial variations in both T_0 and γ (Gleser et al., 2005; Furlanetto and Oh, 2008b; Bolton et al., 2009). What is desired, therefore, are volume-averaged quantities that follow the thermal history of a given population of baryons in the IGM. A parcel of gas at mean density will tend to remain near the mean density (Hui and Gnedin, 1997), and so the redshift evolution of T_0 is ideal for this purpose. The Lyα forest, though, can only give meaningful information for densities that produce measurable, but non-saturated absorption. Since the Lyα optical depth for a given density scales strongly with redshift as $\tau \propto (1+z)^{4.5}$ (assuming $z > 3$ and a constant ionization rate and temperature), the forest does not trace gas at constant density, and at lower redshifts it is mild overdensities that are being probed. To convert from these temperatures to T_0 would then require a knowledge of γ, which is poorly constrained.

To compensate for this difficulty, we measure the temperature at an *optimal* overdensity, $\overline{\Delta}$, for which we are independent of γ. We determine $\overline{\Delta}$ at each redshift empirically for both statistics, following the method of Becker et al. (2011a). This approach is demonstrated in Fig. 6.6. In the upper panels we plot T_0 against $\log\langle|\kappa|\rangle$, $\langle x_{\rm sep}\rangle$ and $\langle\varpi\rangle$. Our calibration simulations (A15-G15) with $\gamma \sim 1.5$ form a well-defined relationship with temperature, whilst those with different γ lie somewhat off this line. The density is then varied until all points lie as close to this line as possible, as shown in the lower panels of Fig. 6.6. This leads to a near one-to-one relationship between $T(\overline{\Delta})$ and $\log\langle|\kappa|\rangle$, as found by Becker et al. (2011a). For the separation and peakiness the calculation was done in three-dimensional $T(\Delta) - x_{\rm sep} - \varpi$ space, but we show the two-dimensional projections on each axis for convenience. At the optimal overdensity, there is also a near one-to-one relationship between T_0 and $x_{\rm sep}$, although the fit between $T(\overline{\Delta})$ and ϖ is not substantially different to that of T_0 and ϖ. Consequently the calculation of the optimal overdensity for the separation-peakiness statistic is driven by the fit with the separation, and thus the peakiness retains some memory of the underlying γ even at the optimal overdensity. We leave it to future work as to whether the separation can be used to constrain $T(\overline{\Delta})$, and then used in conjunction with the peakiness to constrain γ. We do note, however, that even though our calculated value of $\overline{\Delta}$ is dominated by the fit with the separation, the peakiness does influence it. The lower quality of the $T(\overline{\Delta}) - \varpi$ fit compared with the $T(\overline{\Delta}) - x_{\rm sep}$ fit means that for $0.7 \lesssim \gamma \lesssim 1.5$ we find an error in the recovered $T(\overline{\Delta})$ of $\pm \sim 0.05$ dex. In comparison, the error in the recovered $T(\overline{\Delta})$ from the $T(\overline{\Delta}) - \langle|\kappa|\rangle$ fit is $\pm \sim 0.02$ dex. Our error bars in the next section do not include this effect, although at $z \sim 4$ this corresponds to ~ 700 K and so is smaller than the change we measure. Our $\overline{\Delta}$ can be seen as the characteristic overdensity probed by the forest, although both statistics are calculated from a range of densities within the IGM. Importantly though, these allow us to specify a temperature of the IGM without knowing the full $T - \Delta$ relation.

We plot our $\overline{\Delta}$ as a function of redshift for both statistics in Fig. 6.7. We calculate them from noiseless simulated spectra and spectra with S/N = 20, at both HIRES and MIKE resolution, to test the dependence on noise and spectral resolution. With the curvature we find that $\overline{\Delta}$ evolves smoothly with redshift, from mild overdensities at low redshift to roughly the mean density at $z \sim 5$, before increasingly slowly at $z > 5$. Increasing the noise or decreasing the spectral resolution introduces a small bias

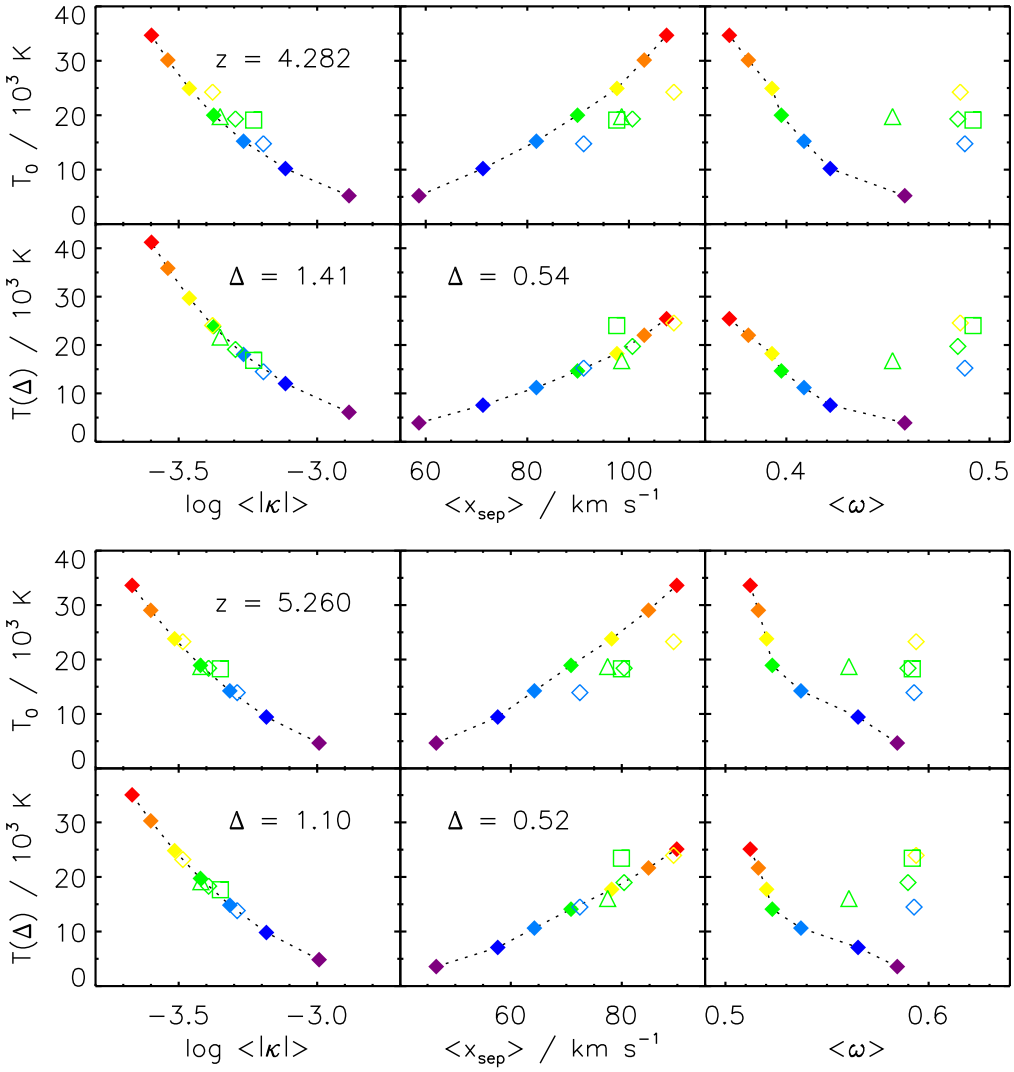


Figure 6.6: Demonstration of our approach to calculating the optimal overdensities probed by the Lyα forest. We show our approach for noiseless spectra with HIRES resolution at $z = 4.282$ (top set of panels) and $z = 5.260$ (bottom set of panels). In the upper panels of each set we show T_0 as a function of the curvature, separation and peakiness. Our calibration simulations (A15-G15) each with $\gamma \sim 1.5$ are shown as solid diamonds (coloured purple to red, respectively) and form a smooth relationship with T_0. Other values of γ are represented by other symbols. Triangles represent $\gamma \sim 1.3$, open diamonds represent $\gamma \sim 1.0$, and squares represent $\gamma \sim 0.7$. In the lower panels we measure the temperature at an overdensity Δ that has been chosen so that each point lies as close to the smooth relationship of the calibration simulations as possible, regardless of γ. For the separation-peakiness statistic, the calculation was done in three-dimensional $T(\Delta) - x_{\rm sep} - \varpi$ space, but we show the two-dimensional projections on each axis for convenience. This process produces near one-to-one relationships for the curvature and the separation, although the peakiness retains some memory of the underlying γ even at this optimal overdensity.

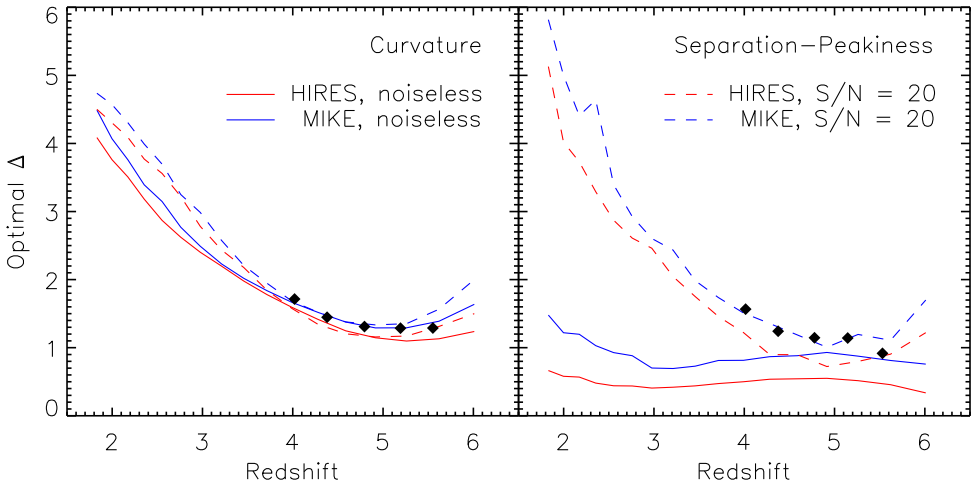


Figure 6.7: The optimal overdensity at which each statistic is independent of γ. *Left-hand panel:* The optimal overdensity for the curvature, as measured from the simulations. The solid lines represent noiseless spectra, and the dashed lines represent spectra with S/N = 20. The densities probed by HIRES are red, and those by MIKE are blue. For both lower resolution or lower S/N, higher overdensities are being probed. The black diamonds represent the densities being probed for a sample of spectra with noise and instrumental properties similar to the data in each of our redshift bins. *Right-hand panel:* The same, but for the separation-peakiness statistic. In noiseless spectra it traces the voids ($\Delta < 1$). A small amount of noise, however, masks their contribution to the small-scale structure at low redshift, and instead it becomes sensitive to the maxima with intermediate fluxes, corresponding to denser gas.

to larger $\overline{\Delta}$. The parts of the spectrum that contribute the most to the curvature signal are where the second flux derivative is varying rapidly from pixel to pixel, namely the tops of transmission peaks and the bottoms of absorption troughs. At high redshift, the peaks in transmission correspond to voids and the troughs correspond to mildly overdense filaments. The average density contributing to the curvature will therefore be close to the mean density. On the other hand, at low redshift the absorption lines correspond to dense filaments, and the transmission peaks saturate near the continuum, thus the average density increases to mild overdensities. The evolution for the separation-peakiness statistic is somewhat more complicated. As just discussed, the optimal overdensity of the statistic is driven by the fit between $T(\overline{\Delta})$ and the separation, rather than the peakiness. Thus, $\overline{\Delta}$ will be most dependent upon the identification and fluxes of the maxima. At high redshift, the most well-resolved transmission peaks will correspond to the lowest density voids. At lower redshift these voids correspond to fluxes ≈ 1. In noiseless spectra though, they can still be resolved as tiny undulations within pixels very close to the continuum. These features dominate the maximum-minimum-maximum groupings detected in low redshift spectra, and so the voids are still the main contributors. However when noise is added, the tiny undulations in pixels near the continuum are no longer distinguishable, and our approach also masks pixels near the continuum (see Section 6.3.4). This has little effect at high redshift where there are few pixels near the continuum, but at lower redshift this pushes the average flux of the maxima to < 0.9. The strong dependence of the optical depth with redshift therefore means that the range in fluxes of the maxima is higher than the fluxes of pixels contributing the most to the curvature (the maxima *and* minima of transmission and absorption features), and thus $\overline{\Delta}$ is

slightly lower. At $z \lesssim 2.5$ though, the peakiness can only be measured for deep features in the spectra whilst the curvature is dominated by shallower absorption lines, and so the peakiness actually traces denser gas than the curvature.

6.3.6 Sensitivity to changes in the gas density field

Both statistics are primarily sensitive to the shape of individual absorption and transmission features. This shape is determined by a combination of thermal broadening, turbulent broadening, and the small-scale structure of the underlying gas density distribution. If the IGM is smooth then the denser regions responsible for a particular absorption feature will be more spatially extended, and so Hubble flow broadening will be more significant (Weinberg et al., 1997). This effect of smoothing out the small-scale structure in the IGM density field is known as Jeans smoothing. Ideally most of our signal should be coming from the instantaneous temperature, as the small-scale structure of the IGM is poorly constrained. It is important, therefore, to test the sensitivity of our statistics to changes in both the instantaneous temperature, and in the small-scale structure of the gas density field.

We approach this issue with our B15 and D15 simulations by constructing spectra after swapping the gas density profiles for each sightline. The statistics are then measured for these new spectra, and converted into a temperature. We plot our results from the curvature statistic in Fig. 6.8. The red dashed line is for spectra constructed from the B15 simulation outputs, except with temperature fields from the D15 simulation. An ideal statistic that is sensitive to only the instantaneous temperature should follow the red line in Fig. 6.8, calculated for unmodified spectra from the D15 simulation. However, the greater small-scale structure in the B15 density field narrows the absorption and transmission features in the spectra in comparison to their D15 counterparts, so T_0 is underestimated. This underestimation is only minor at $z \sim 2$ since the width of absorption lines is largely determined by thermal broadening, however the discrepancy increases with redshift as the forest becomes sensitive to lower density gas. Similarly, the blue dashed line is for spectra constructed from the D15 simulation outputs, but with temperature fields from the B15 simulation. Ideally they should follow the blue line in Fig. 6.8, calculated from unmodified spectra from the B15 simulation, however the smooth D15 density fields mean that T_0 is overestimated. Again, the overestimation is small at $z \sim 2$, but increases with redshift. At high redshift therefore, these two effects become degenerate, as both a smooth density field with cold instantaneous temperature, and a coarse density field with hot instantaneous temperature have the same curvature. The redshift at which this occurs lowers with increasing noise, and decreasing spectral resolution.

Fig. 6.9 is the equivalent for T_0 calculated with the separation-peakiness statistic. This statistic is slightly more resilient to Jeans smoothing at $z > 4$ for high S/N data than the curvature, but suffers from the same problem of degeneracy at $z > 5$, particularly in noisy, low-resolution data. Both statistics are therefore not only sensitive to the instantaneous temperature, but also to the small-scale structure of the underlying density field, and therefore to any Jeans smoothing. This issue is explored further in the next section.

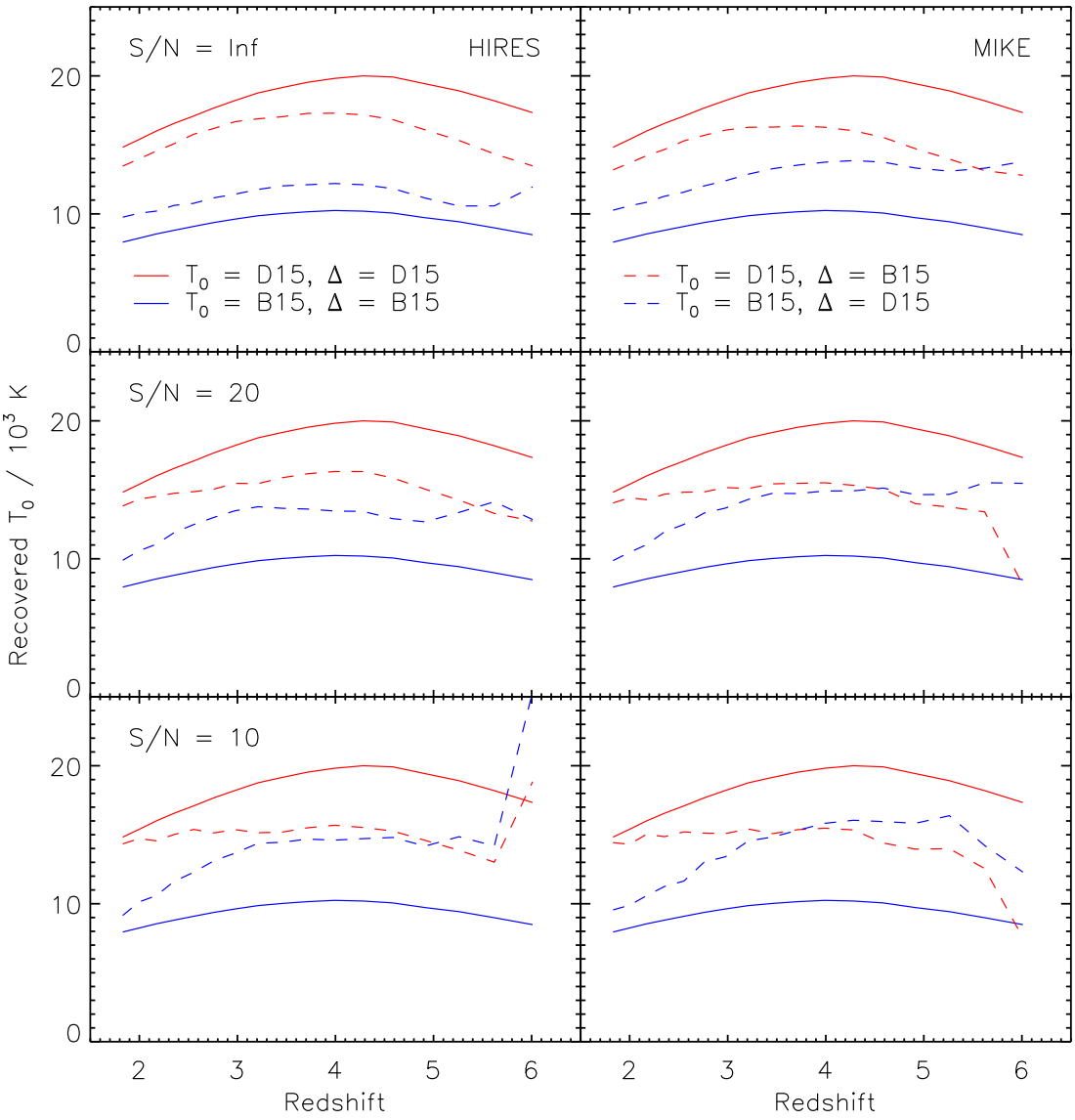


Figure 6.8: Sensitivity of the curvature to changes in the density field. *Top panels:* The temperatures returned by the curvature for a variety of combinations of instantaneous temperature and density field, for HIRES resolution (left) and MIKE resolution (right) noiseless simulated spectra. The blue line shows T_0 for the B15 simulation, and the red line is for the D15 simulation. The difference in T_0 is about a factor of 2. The blue dashed line is the returned T_0 for spectra constructed using the B15 instantaneous T_0, but the smoother D15 density field, and vice versa for the red dashed line. At high redshift the curvature becomes increasingly degenerate between high instantaneous temperatures and smooth density fields. *Middle panels:* Same as top panel but for S/N = 20. The confusion between the two effects becomes completely degenerate at $z > 5$. *Bottom panels:* Same as top panel but for S/N = 10. The two effects are essentially degenerate at $z > 3.5$.

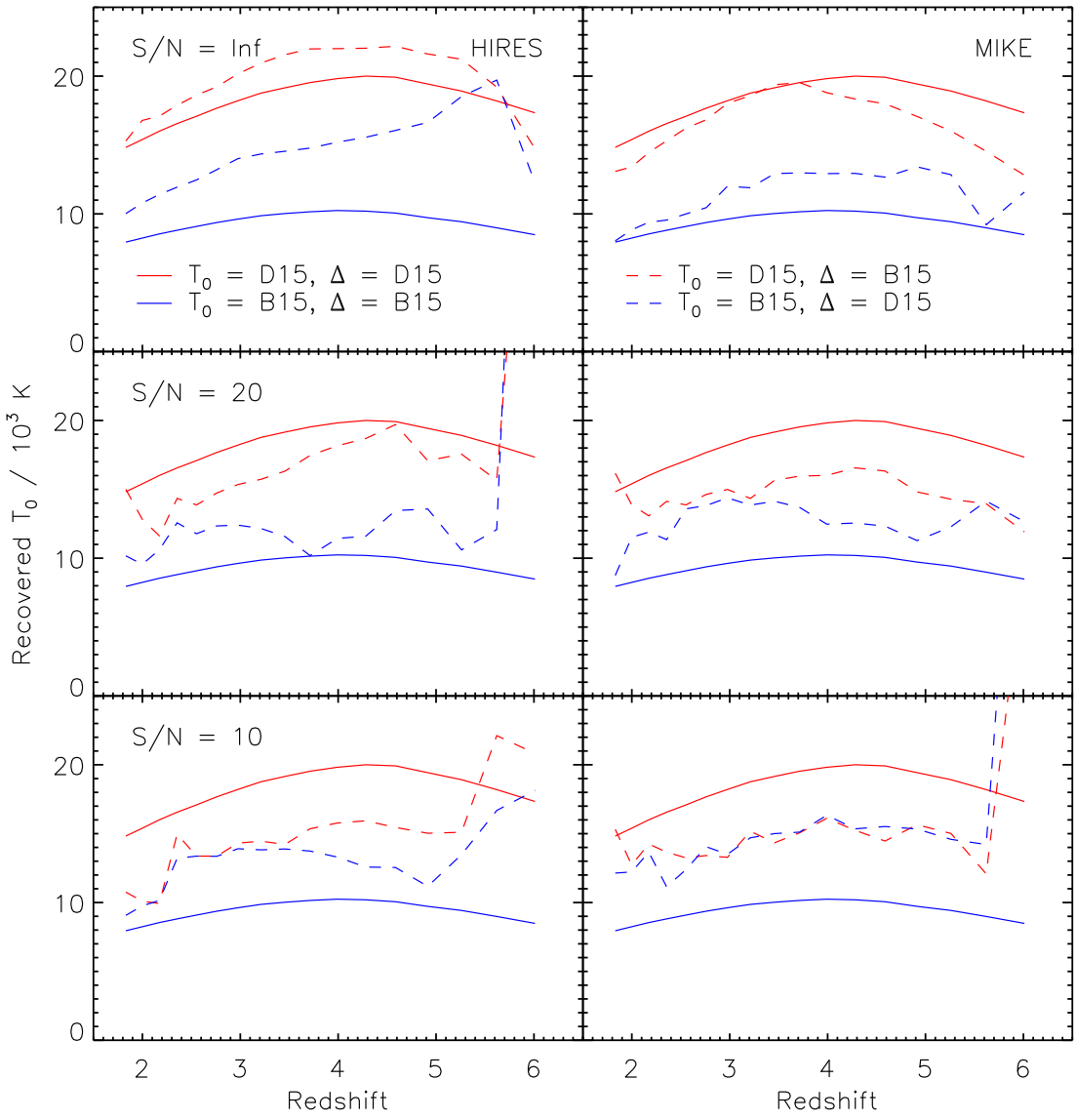


Figure 6.9: Sensitivity of the separation-peakiness statistic to changes in the density field. *Top panels:* The temperatures returned by the statistic for a variety of combinations of instantaneous temperature and density field, for HIRES resolution (left) and MIKE resolution (right) noiseless simulated spectra. The blue line shows T_0 for the B15 simulation, and the red line is for the D15 simulation. The difference in T_0 is about a factor of 2. The blue dashed line is the returned T_0 for spectra constructed using the B15 instantaneous T_0, but the smoother D15 density field, and vice versa for the red dashed line. Similar to the curvature, at high redshift the statistic becomes increasingly degenerate between high instantaneous temperatures and smooth density fields. *Middle panels:* Same as top panel but for S/N = 20. The separation-peakiness statistic maintains some distinction between the cases even at $z > 5$. *Bottom panels:* Same as top panel but for S/N = 10. The two effects are essentially degenerate at $z > 2.5$.

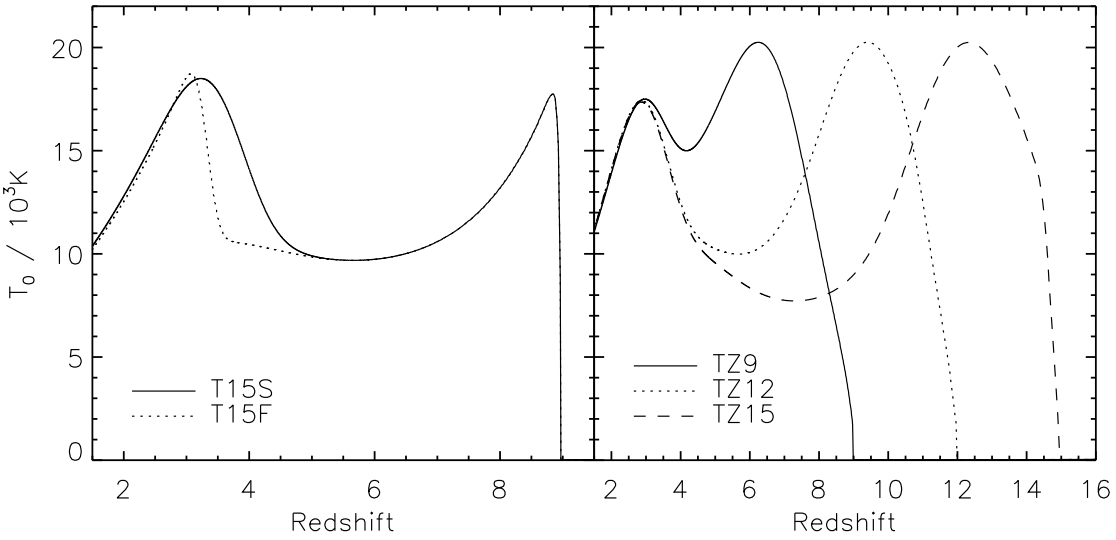


Figure 6.10: The thermal histories of the simulations that include photoheating due to hydrogen reionization. *Left-hand panel:* The simulations T15S (solid line) and T15F (dotted line). Both have instantaneous hydrogen reionization at $z = 9$, although helium reionization in T15F is much less extended. *Right-hand panel:* The simulations TZ9 (solid line), TZ12 (dotted line), and TZ15 (dotted line). They all have an extended ($\Delta z \sim 3$) hydrogen reionization beginning at $z = 9$, 12 and 15 respectively, corresponding to late, middling and early reionization scenarios. We measure both statistics from artificial spectra drawn from all these simulations to test the impact of the thermal history on measurements of the instantaneous temperature.

6.3.7 Sensitivity to different thermal histories

As we have just demonstrated, both of our statistics are sensitive to the level of Jeans smoothing in the spectra. However, for that test we assumed non-evolving 'hot' and 'cold' density fields across the whole redshift range. In practice, assuming the instantaneous temperature stays cool then the 'hot' density field would relax into the 'cold' density field over time. Similarly, if the instantaneous temperature rises significantly, then the density field will undergo Jeans smoothing accordingly, and the discrepancy between the instantaneous temperature and the small-scale structure of the density field will eventually be erased. However the time scales over which this can happen are of order a Hubble time for mean density gas, which is significantly longer than the time scales over which reionization occurs. Consequently, the level of Jeans smoothing from the photoheating associated with reionization will still be a significant consideration at $z > 4$.

To test the sensitivity of our statistics to different thermal histories, and thus different levels of Jeans smoothing, we measured them from spectra drawn from simulations with a wide variety of thermal histories, designed to mimic the effects of HI and HeII reionization. The thermal histories for these simulations are shown in Fig. 6.10. Two of them, T15S and T15F, include heating to imitate an instantaneous HI reionization at $z = 9$, and both an extended and quick HeII reionization, respectively. The other three imitate rather extended HI reionization scenarios ($\Delta z \sim 3$), representing late reionization finishing at $z \sim 6$ (TZ9), a middling reionization finishing at $z \sim 9$ (TZ12), and an early reionization finishing at $z \sim 12$ (TZ15). We hold γ to its fiducial value of ~ 1.5 in each of these runs for

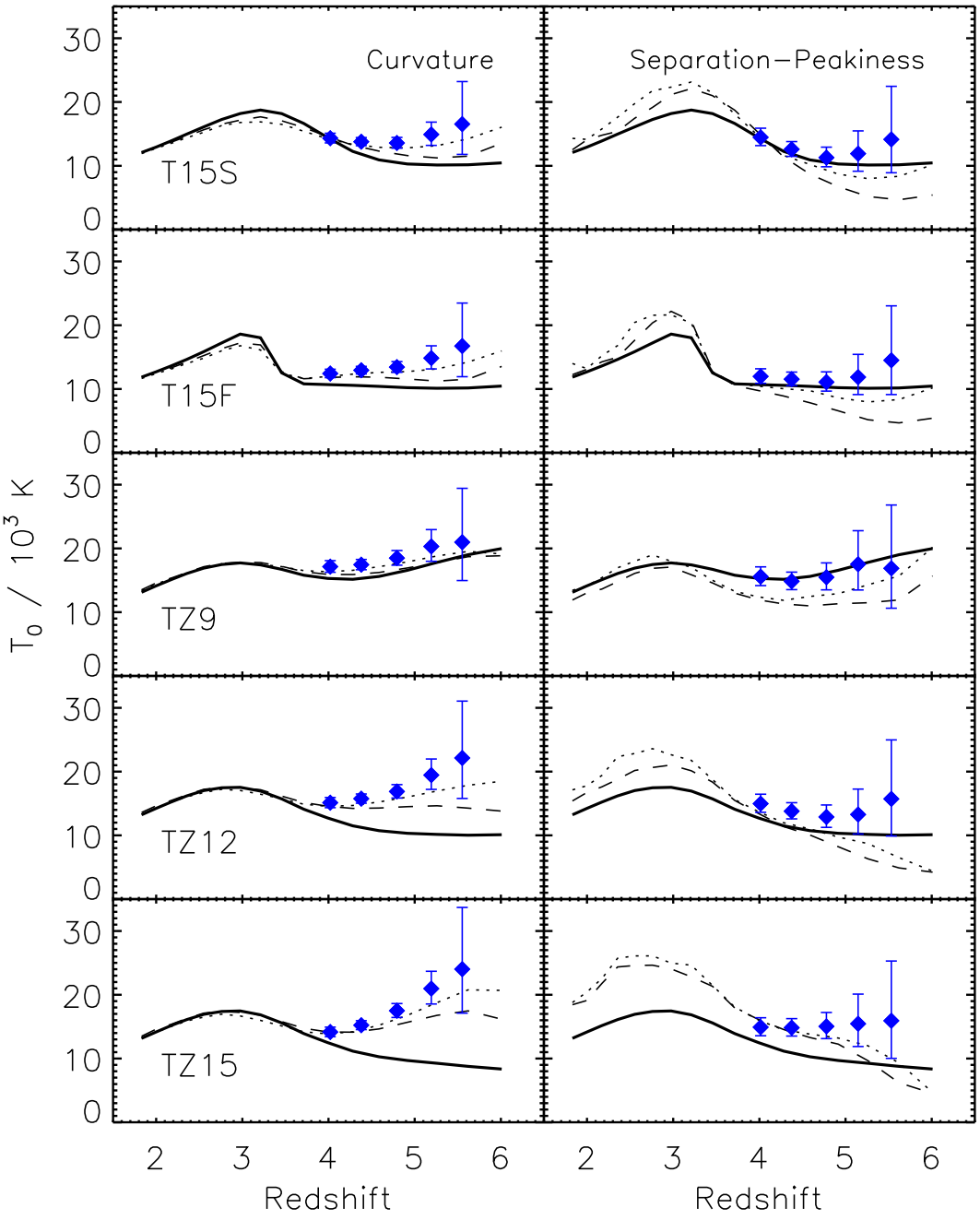


Figure 6.11: Temperature at the mean density for the fiducial $\gamma \sim 1.5$, recovered from our input thermal histories using both the curvature (left) and separation-peakiness (right) statistics. The blue diamonds give the values we measure using artificial spectra with resolution, noise properties and sample size similar to the real data. The solid lines give the actual values measured directly from the simulation. All error bars are 2σ. We also plot the T_0 that would be recovered from noiseless HIRES (dashed lines) and MIKE (dotted line) spectra. Differences in the thermal history of the test runs and the comparison simulations cause T_0 to be overestimated at $z \gtrsim 4$.

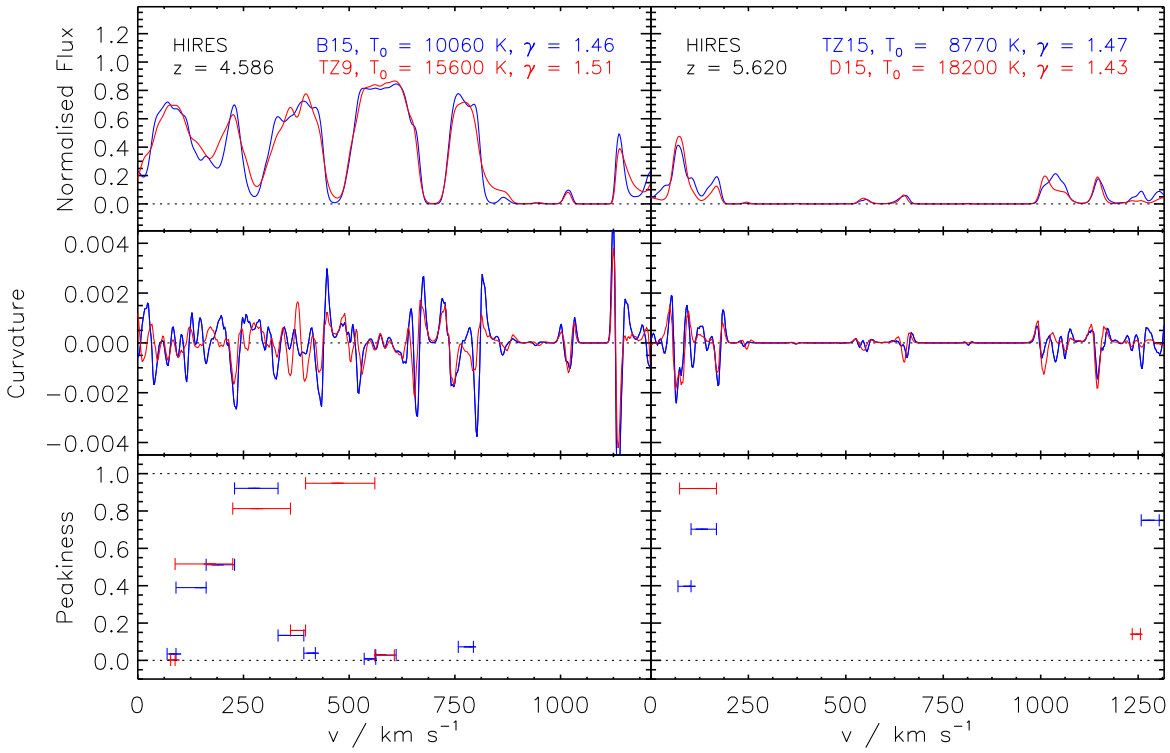


Figure 6.12: A demonstration of the degeneracy between the thermal history and instantaneous temperature. The panels are similar to those in Fig. 6.1, except that we have used slightly higher redshifts and all the spectra have HIRES resolution. Also, in the left panels we have replaced run D15 with TZ9 (which has a late reionization), and in the right panels we have replaced run B15 with TZ15 (which has an early reionization). The spectra on the left have very similar peakiness statistics despite rather different T_0, meaning it is unable to distinguish between a cool IGM and a recent reionization. Similarly, the spectra on the right have very similar curvature statistics yet T_0 varies by a factor of 2, so it is unable to distinguish between a hot IGM and an early reionization.

simplicity. Artificial spectra were then drawn from each simulation with similar redshifts, noise levels, instrumental resolutions and sample sizes to the real data, from which both statistics were measured.

Our recovered temperatures are shown in Fig. 6.11 as blue filled diamonds. The value of T_0 measured directly from the simulation is shown in each panel as a thick black line. For comparison we also show the T_0 that would be recovered from noiseless HIRES (dashed line) and MIKE (dotted line) resolution spectra. Both statistics recover well the peak in T_0 due to He II reionization, but struggle with the Jeans smoothing in the high-redshift forest, particularly at $z > 5$. The earlier the period of heating due to H I reionization, the greater the Jeans smoothing as the density structures within the IGM are less massive and easier to smooth. Similarly, a more extended H I reionization also increases the amount of Jeans smoothing. The curvature consistently overestimates T_0 at $z > 5$, and this overestimation worsens with increasing Jeans smoothing (see left panel of Fig. 6.11). This is because the transmission peaks will tend to be less sharp and spread on top of large scale undulations, and so the curvature will be lower. Additional smoothing, either due to lower spectral resolution or dealing with noise, strengthens this effect, and the overestimation increases. Also, within the instantaneous H I reionization models, the curvature underestimates the peak in T_0 from He II reionization, as found

z	Curvature					Separation-Peakiness				
	T15S	T15F	TZ9	TZ12	TZ15	T15S	T15F	TZ9	TZ12	TZ15
3.8 – 4.2	0.004	0.066	0.050	0.079	0.059	0.007	0.050	0.008	0.074	0.081
4.2 – 4.6	0.066	0.090	0.057	0.146	0.146	0.026	0.040	−0.013	0.088	0.134
4.6 – 5.0	0.111	0.116	0.057	0.207	0.247	0.029	0.032	−0.019	0.089	0.180
5.0 – 5.4	0.167	0.166	0.063	0.281	0.351	0.068	0.068	0.003	0.114	0.216
5.4 – 5.8	0.212	0.217	0.048	0.343	0.433	0.144	0.155	−0.045	0.194	0.253

Table 6.3: The offset between the returned T_0 and input T_0 for both statistics for each of the five tested thermal histories, as measured from artificial spectra with resolution, noise properties and sample size similar to the real data. All values are given in dex. These are used to correct the measured T_0 for the Jeans smoothing of each type of thermal history in Section 6.4.4 such that $\log T_0^{\rm corr} = \log T_0^{\rm meas} - \text{offset}$.

by Becker et al. (2011a). We note, however, that for all of the extended HI reionization scenarios this behaviour is not present.

By contrast, T_0 is underestimated at $z > 5$ from separation-peakiness measurements taken from the noiseless spectra. This is due to the small-scale structure that imprints itself upon the wider undulations leading to many closely spaced maxima. As such, the average separation is smaller than would be expected just from maxima of the undulations, and so low values of T_0 are returned. However, upon any smoothing due to the addition of noise, or lower resolutions, these small features are lost, leading to a wider average separation and thus higher T_0. Due to the long time required for the IGM to return to hydrostatic equilibrium, the Jeans smoothing influences the measurement of T_0 at lower redshift. Minima will be shallower at $z \lesssim 3$ from the Jeans smoothing of the filaments at higher redshift, so peakiness values will be lower and the returned T_0 higher. This overestimation at low redshift increases with the amount of Jeans smoothing.

Fig. 6.12 demonstrates the degeneracy between spectra with very different thermal histories. In the left-hand panel, we show how the peakiness struggles to differentiate between a cool IGM and a late reionization, whilst in the right-hand panel we show how the curvature fails to distinguish between a hot IGM and an early reionization. It should be noted though, that when one statistic is degenerate the other is not, and so they can together provide complementary constraints on T_0. The offsets of our recovered T_0 from the input T_0 for artificial spectra with similar properties to the data are shown (in dex) in Table 6.3. These are the shifts used to correct our measured points for Jeans smoothing in Section 6.4.4.

6.4 Results

6.4.1 Temperature at the optimal overdensity

We have made no attempt in this work to specify the whole $T - \Delta$ relation, and so our primary results are the IGM temperatures at the optimal overdensity probed by the Lyα forest (see Section 6.3.5). Our values of $T(\overline{\Delta})$ are given in Table 6.4 and plotted in the top panels of Fig. 6.13. For each redshift bin we also specify the optimal overdensity for which these temperatures were measured. All of the errors are 2σ and have been estimated from the simulations.

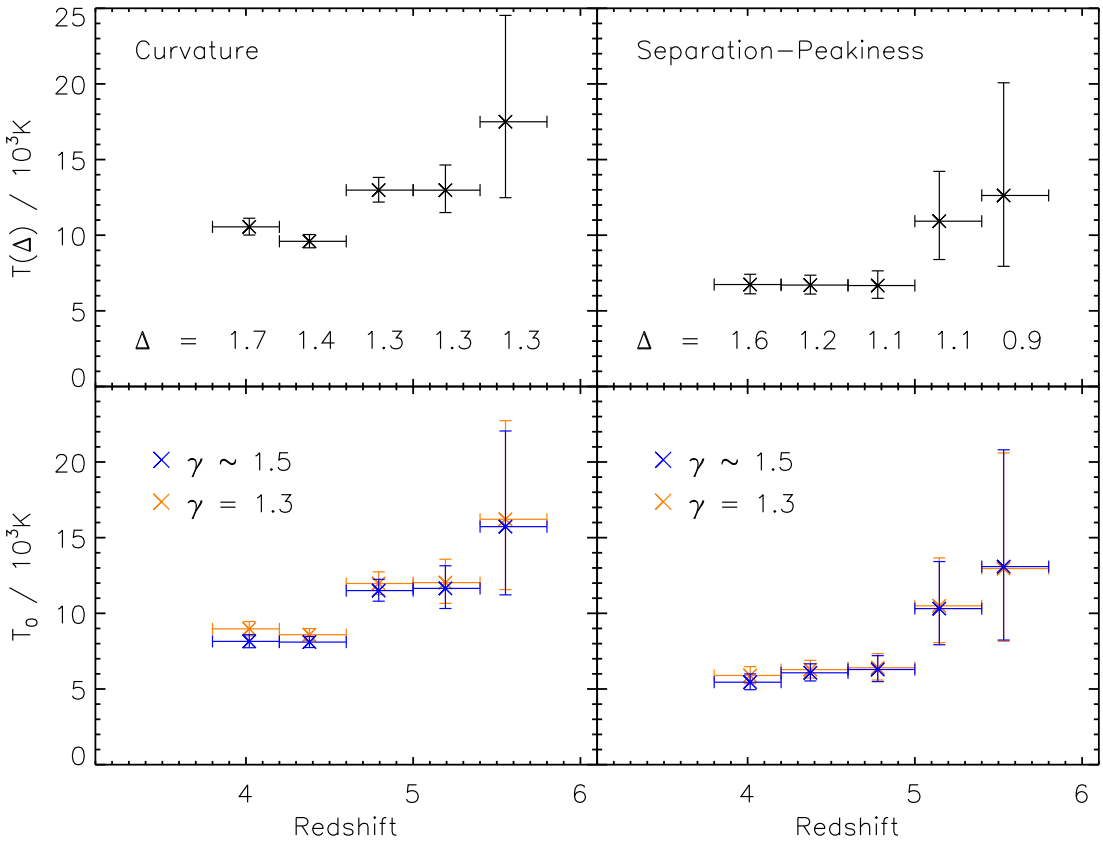


Figure 6.13: *Top panels:* The IGM temperature at the optimal density as a function of redshift as measured by the curvature statistic (left) and the separation-peakiness statistic (right). Since we do not try and specify γ these are our primary results. The optimal densities for each redshift bin are given along the bottom. The horizontal error bars specify the range in redshift covered by each bin, and the vertical error bars are 2σ estimated from simulated spectra with similar noise and instrumental properties. They do not account for errors due to incorrect continua (although this will only affect our curvature based temperatures). *Bottom panels:* The IGM temperature at mean density as a function of redshift for two assumed values of γ. The blue crosses give T_0 for the fiducial $\gamma \sim 1.5$ from the calibration simulations (listed in Table 6.4), whilst the orange crosses are for a flattened $T-\Delta$ relation with $\gamma = 1.3$. Since we are probing densities near the mean density at $z > 4$, the values of T_0 are highly insensitive to the choice of γ.

Between $z \sim 4$ and 5 the temperatures are relatively flat. At $z > 5$, however, there is a rapid rise in $T(\overline{\Delta})$ with increasing redshift. After hydrogen reionization the $T-\Delta$ relation is expected to be rather steep, as the high density gas in the filaments collapses and heats up, whilst the voids cool through adiabatic expansion. Since there is little variation in the optimal overdensity probed, this rise in temperature can not be explained simply with a steep $T-\Delta$ relation, and instead suggests that the gas at $z > 5$ was recently heated.

6.4.2 Temperature at the mean overdensity

To convert from $T(\overline{\Delta})$ to T_0 requires a knowledge of the slope of the $T-\Delta$ relation, γ. Although we do not measure γ, T_0 can be calculated based upon reasonable assumptions for γ. Long after

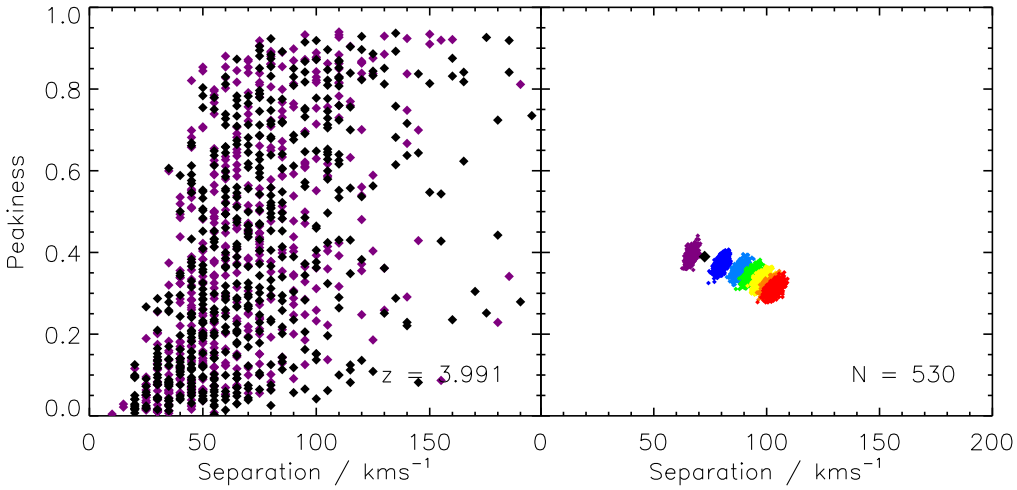


Figure 6.14: Derived values of separation and peakiness for features in the Lyα forest, from simulated MIKE spectra at $z = 3.991$ with S/N = 25, and the MIKE data with similar noise properties from the $3.8 < z < 4.2$ bin. *Left-hand panel:* the location of the 530 individual features measured in the real data (black), and a random 530 features from the best fitting simulation, A15 (purple). There is good agreement between the two. *Right-hand panel:* analogous to the right-hand panel in Fig. 6.2, but also with the mean values from the data (black diamond). The statistical error on this point is of comparable size to the simulations (coloured points). The mean values suggest an IGM that is slightly warmer than the coldest simulation (A15, with $T_0 = 5240$ K), but still cool at ~ 6000 K.

z	$\langle z \rangle$	$\overline{\Delta}$	N	$\gamma^{\rm fid}$	$\langle \varepsilon \rangle$	$T(\overline{\Delta})$	$T_0^{\rm fid}$	$T_0^{\gamma=1.3}$
$3.8-4.2$	4.02	1.72	161	1.48	0.07	10550^{+570}_{-540}	8150^{+440}_{-420}	8980^{+490}_{-460}
$4.2-4.6$	4.38	1.45	191	1.46	0.07	9600^{+450}_{-430}	8100^{+380}_{-360}	8590^{+400}_{-380}
$4.6-5.0$	4.79	1.31	135	1.45	0.08	12980^{+839}_{-790}	11500^{+740}_{-700}	11970^{+770}_{-730}
$5.0-5.4$	5.19	1.29	111	1.43	0.12	12980^{+1668}_{-1480}	11650^{+1500}_{-1330}	12030^{+1550}_{-1370}
$5.4-5.8$	5.55	1.29	57	1.42	0.12	17500^{+7032}_{-5020}	15730^{+6320}_{-4510}	16220^{+6520}_{-4650}

(a) Results from the curvature statistic.

z	$\langle z \rangle$	$\overline{\Delta}$	N	$\gamma^{\rm fid}$	$\langle \varepsilon \rangle$	$T(\overline{\Delta})$	$T_0^{\rm fid}$	$T_0^{\gamma=1.3}$
$3.8-4.2$	4.02	1.57	769	1.48	0.06	6750^{+680}_{-610}	5450^{+550}_{-500}	5900^{+590}_{-540}
$4.2-4.6$	4.38	1.24	844	1.46	0.06	6700^{+650}_{-600}	6070^{+590}_{-540}	6280^{+610}_{-560}
$4.6-5.0$	4.78	1.15	443	1.44	0.08	6670^{+970}_{-850}	6290^{+910}_{-800}	6410^{+930}_{-810}
$5.0-5.4$	5.15	1.14	135	1.43	0.10	10930^{+3290}_{-2530}	10310^{+3110}_{-2390}	10500^{+3170}_{-2430}
$5.4-5.8$	5.53	0.92	43	1.42	0.12	12630^{+7450}_{-4680}	13090^{+7720}_{-4860}	12960^{+7640}_{-4810}

(b) Results from the separation-peakiness statistic.

Table 6.4: Temperature results. Columns give the redshift interval, mean redshift of the bin, optimal overdensity total amount of data in the bin, fiducial value of γ, mean rms noise, temperature at the optimal overdensity, temperature at the mean density for the fiducial γ, and the temperature at the mean density for $\gamma = 1.3$. For the curvature statistic, $\langle z \rangle$ is the path-length-weighted mean redshift and N is the total number of $10h^{-1}$ Mpc sections included. For the separation-peakiness statistic, $\langle z \rangle$ is the mean redshift of all the individual features, N is the total number of features included, and $\langle \varepsilon \rangle$ is the mean of the rms noise values in the flux error array beneath each individual feature. All errors are 2σ, and are symmetric in log temperature.

hydrogen reionization, γ is expected to asymptote towards ~ 1.5 (Hui and Haiman, 2003). During reionization though, the inside-out nature of the process means that high density regions are heated first, later followed by the voids. Consequently the voids are only heated after the initial gas has had chance to cool, and thus the $T-\Delta$ relation flattens, with $\gamma < 1.5$. Numerical simulations suggest that $\gamma = 1.3$ may be reasonable at the end of reionization (McQuinn et al., 2009), although an isothermal ($\gamma = 1$) or even inverted ($\gamma < 1$) $T-\Delta$ relation are other possibilities (Furlanetto and Oh, 2008b, 2009; Chang et al., 2011), for which there is some observational evidence (Becker et al., 2007; Bolton et al., 2008; Viel et al., 2009, although see Lee (2011)).

We plot T_0, as measured by both statistics, for two different assumptions of γ in the bottom panels of Fig. 6.13. The first of those takes the γ from our fiducial (A15-G15) simulations, where γ increases from ~ 1.4 at $z = 5.6$ to ~ 1.5 at $z = 4.0$. The second assumes a flattened $T-\Delta$ with $\gamma = 1.3$, representative of a possible slope at the end of reionization. Both statistics, however, probe densities very close to the mean density across our redshift range, and so they are largely insensitive to our choice of γ. Our measurements of T_0 therefore avoid a significant uncertainty that could otherwise have been a substantial source of systematic bias.

Our T_0 as derived by the curvature are relatively flat for $z < 4.6$, before a large increase with increasing redshift, going from ~ 8000 K at $z = 4.4$ to $\sim 16\,000$ K at $z = 5.6$. There is a similar pattern amongst the T_0 derived by the separation-peakiness statistics, although the strong increase is only at $z > 5$. Also all the values are lower than those from the curvature by $\sim 2000-3000$ K, with a cool IGM favoured at $z \sim 4$ (see Fig. 6.14), and rising from ~ 6000 K at $z = 4.4$ to $\sim 13\,000$ K at $z = 5.6$. Both statistics detect this rise, suggesting that we are seeing the adiabatic cooling of the IGM after the completion of hydrogen reionization. However, as mentioned in Section 6.3.7, measurements of the instantaneous T_0 in this redshift range can be biased by the Jeans smoothing of an earlier epoch of heating, and so our T_0 are likely overestimates. The robustness of our increase in T_0 at $z > 5$ to a range of different thermal histories is therefore carefully investigated in Section 6.4.4.

6.4.3 Comparison to previous work

The measurements of the temperature of the IGM presented here are the first at $z > 5$ from the Lyα forest, and so there are no direct comparisons to made in this redshift range. For $3.8 \leq z \leq 5.0$ however, the most precise results are from Becker et al. (2011a), and they are used for comparison in Fig. 6.15. Despite the slight differences in our measurement of the curvature compared to theirs (no renormalisation of the stretches of forest, and pixels contributing from a slightly different ranges of fluxes), there is still very good agreement between our temperature measurements based on that statistic. We note, however, that agreement should be expected as we used the same simulations, and there is also some crossover in our data sample. Even so, at $z \sim 4$ our T_0 measurement is lower than theirs at the 2σ level. This may be due to the fact that we make no corrections for metal lines, which will bias curvature measurements to lower temperatures (Becker et al., 2011a).

More noticeably though, our T_0 at $z \sim 4.8$ is somewhat higher than that of Becker et al. (2011a). This is potentially due to a combination of two factors: poor continuum placement, and the inclu-

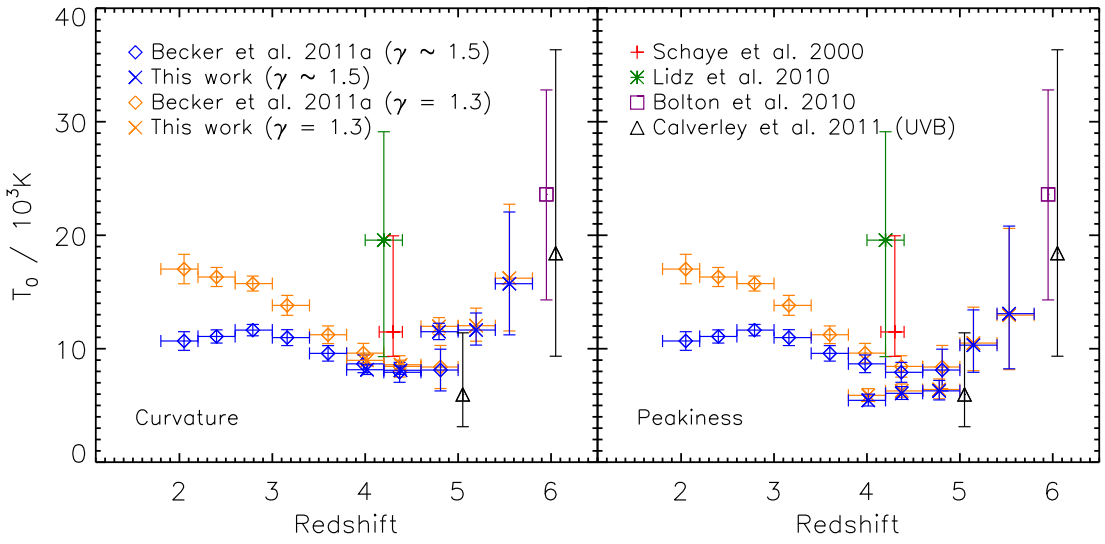


Figure 6.15: Comparison of our results to selected results from the literature. In the left-hand panel we compare our T_0 from the curvature for both our fiducial γ and $\gamma = 1.3$ to those of Becker et al. (2011a), the highest redshift bins of Lidz et al. (2010) (green star) and Schaye et al. (2000) (red cross), the temperature measured close to a quasar from Bolton et al. (2010) (purple square), and estimates of T_0 based on the UVB measured by Calverley et al. (2011) (black triangles). In the right panel we compare our T_0 from the separation-peakiness statistics to the same literature values. All error bars are 2σ, except those of Schaye et al. (2000) and those estimated from the Calverley et al. (2011) ultraviolet background (UVB) measurements.

sion of noisier stretches of forest. The renormalisation of the forest by Becker et al. (2011a) allows uncertainties in the continuum to be handled objectively and consistently, whilst here we assume the continuum to be correct, potentially making our measurements susceptible to a significant source of systematic bias. The difference in T_0 is ~ 3500 K though, which would require an error in continuum placement of $\sim 30\%$. This seems a rather large error considering that the transmission in the forest is high enough for a reasonable fraction of pixels to be near the continuum. The other possibility is that since Becker et al. (2011a) required the rms noise to be ≤ 0.06 per 2.1 km s^{-1} bin, only a very small number of sightlines contributed to their measurement, whilst our requirement of rms noise ≤ 0.20 per 2.1 km s^{-1} allows more sightlines to contribute. The Becker et al. (2011a) result should thus be more sensitive to cosmic variance than our own, such as if the sightlines with the least noise also happen to be cooler than average at that redshift. This seems to indeed be the case, as raising their noise requirement to allow noisier stretches of forest (and thus more quasar sightlines) increases the measured temperature (G. Becker, private communication). Thus the discrepancy between our $z \sim 4.8$ temperature and that of Becker et al. (2011a) can be eased if continuum placement errors bias our T_0 high, and cosmic variance bias their T_0 low.

In the right-hand panel of Fig. 6.15 we compare the results of Becker et al. (2011a) to our results from the separation-peakiness statistics. There is little variation in T_0 across $3.8 < z < 5.0$ in both. Since the peakiness is unaffected by continuum placement errors, and the Becker et al. (2011a) measurements handle the continuum in such a way as to minimise the effect of continuum placement errors, this supports the idea that our $z \sim 4.8$ measurement from the curvature is affected by contin-

uum effects. Although the shape of the evolution over $3.8 < z < 5.0$ is very similar, our temperature measurements are somewhat lower, by ~ 2000 K. Becker et al. (2011a) note, however, that their highest redshift temperatures may be overestimated by roughly this amount when Jeans smoothing by an earlier epoch of heating (such as that associated with hydrogen reionization) is taken into account. As demonstrated in Section 6.3.7, the thermal history assumed can have a large effect on where these points lie. The full impact of different thermal histories on our results is discussed in more detail in Section 6.4.4.

The Becker et al. (2011a) temperatures at $z \sim 4-5$ are lower than those at similar redshifts in the literature. Since we are lower still, the discrepancy with our results is even greater (see Fig. 6.15). Using the distribution of widths of Lyα absorption lines, Schaye et al. (2000) find that $T_0 = 11\,480^{+8470}_{-2150}$ (1σ errors) at $z \sim 4.3$, whilst Lidz et al. (2010) use wavelet analysis to derive $T_0 = 19\,570^{+9550}_{-10\,270}$ (2σ errors) at $z \sim 4.2$, although they emphasise this point is highly sensitive to errors in the mean flux. Similarly, using the flux power spectrum of the forest, Zaldarriaga et al. (2001) measure $T_0 = (2.2 \pm 0.4) \times 10^4$ K (2σ errors) at $z = 3.9$. The temperatures measured from the curvature are the lowest amongst literature values of T_0 at $z \sim 4$, and our temperatures from the separation-peakiness statistics are lower still. We leave a full comparison of different methods for measuring the small-scale structure at $z \sim 4$ for future work.

Whilst these are the first measurements of T_0 at $z > 5$ from the Lyα forest, some consistency checks can be made with other observations. Bolton et al. (2010) measured the temperature of the IGM close to a $z \sim 6$ quasar to be $T_0 = 23\,600^{+9200}_{-9300}$ (2σ errors). The hard spectrum of the quasar means that in this region He II has been reionized, whilst in the general IGM at this redshift it is likely that only H I and He I reionization has occurred. The IGM close to the quasar should therefore have received an extra boost in temperature. The amplitude of this boost is dependent on the optical thickness of the gas (Abel and Haehnelt, 1999), and the volume-averaged boost will be smaller than the boosts experienced by individual parcels of gas. Despite the unknown size of this boost, their T_0 provides an upper limit, as the general IGM should not be hotter than the gas close to a quasar. Our T_0 measured at $z > 5$ are consistent with being below this limit, which is an important check.

Finally, weak limits on the temperature of the IGM can come from measurements of the amplitude of the ionizing background, $\Gamma_{\rm bkg}$. Wyithe and Bolton (2011) used the effective optical depth of the Lyα forest to derive $\Gamma_{\rm bkg} = 0.47^{+0.3}_{-0.2} \times 10^{-12}$ s^{-1} and $\Gamma_{\rm bkg} = 0.18^{+0.18}_{-0.09} \times 10^{-12}$ s^{-1} at $z = 5$ and 6, respectively. For a given effective optical depth there is a degeneracy between $\Gamma_{\rm bkg}$ and the assumed parameters of the $T - \Delta$ relation, which for their measurements were $T_0 = 10^4$ K and $\gamma = 1.3$. The relationship between these parameters is such that the derived $\Gamma_{\rm bkg} \propto T_0^{-0.57}\gamma^{0.81}$ and $\Gamma_{\rm bkg} \propto T_0^{-0.61}\gamma^{1.07}$ at redshifts 5 and 6, respectively (Bolton and Haehnelt, 2007c). Consequently, for an independent measurement of $\Gamma_{\rm bkg}$ and an assumed value of γ, constraints can be placed on T_0. Recently, Calverley et al. (2011) used the quasar proximity effect to measure $\Gamma_{\rm bkg}$ at $z > 5$, independent of assumptions about the $T - \Delta$ relation. They find $\Gamma_{\rm bkg} = 0.71^{+0.32}_{-0.22} \times 10^{-12}$ s^{-1} at $z \sim 5$, and $\Gamma_{\rm bkg} = 0.14^{+0.07}_{-0.05} \times 10^{-12}$ s^{-1} at $z \sim 6$. Thus, assuming $\gamma = 1.5$, the results of Wyithe and Bolton (2011) can be scaled to match those of Calverley et al. (2011) for $T_0 = 5970^{+5430}_{-2840}$ K at $z = 5$, and $T_0 = 18\,420^{+17\,920}_{-9080}$ K at $z = 6$. Therefore, the redshift evolution of $\Gamma_{\rm bkg}$ as measured by Calverley et al. (2011) prefers a relatively cool IGM at $z = 5$ and

an increase in T_0 from $z \sim 5-6$, which is precisely what both the curvature and separation-peakiness temperature measurements show. This gives us confidence that the increase may be real, although, due to the large error bars associated with T_0 from $\Gamma_{\rm bkg}$, non-evolving thermal histories cannot be ruled out.

6.4.4 Implications for hydrogen reionization

After the completion of HI reionization the IGM will have cooled rapidly via a combination of adiabatic expansion and Compton cooling (Miralda-Escudé and Rees, 1994; Hui and Gnedin, 1997). The IGM only retains a relatively short-term memory of this heating and cooling, and will reach an asymptotic temperature dependent only on the spectrum of the ionizing sources (Hui and Haiman, 2003). Previous studies have used T_0 at $z \sim 4$ to try and constrain when HI reionization occurred, as a high temperature would be inconsistent with the expected asymptote (for a reasonable spectrum), and so suggest a recent reionization event.

Hui and Haiman (2003) find that the thermal asymptote can be fit to within $\sim 5\%$ for a wide range of input spectra by

$$T_0 = \left[B(1+z)^{0.9}\right]^{1/1.7} \tag{6.3}$$

where

$$B \equiv 18.8\ \mathrm{K}^{0.7} \left[\sqrt{\frac{0.14}{\Omega_{\rm m}h^2}}\frac{\Omega_{\rm b}h^2}{0.024}\right]\left(T_{\rm HI} + \frac{x_{\rm HeII}}{16}T_{\rm HeI} + \frac{5.6x_{\rm HeIII}}{16}T_{\rm HeII}\right), \tag{6.4}$$

and

$$T_i \equiv \frac{1}{k_{\rm B}}\frac{\int_{\nu_i}^{\infty}\frac{J_\nu\sigma_i(\mathrm{h}\nu-\mathrm{h}\nu_i)}{\mathrm{h}\nu}d\nu}{\int_{\nu_i}^{\infty}\frac{J_\nu\sigma_i}{\mathrm{h}\nu}d\nu}\,. \tag{6.5}$$

Here $k_{\rm B}$ is the Boltzmann constant, h is Planck's constant, J_ν is the ionizing intensity as a function of frequency (in units of energy per frequency per time per area per steradian), σ_i is the ionization cross section for the respective species (i = HI, HeI and HeII), ν_i is the ionizing threshold frequency, and x_i is the fraction of that element in that species. If $J_\nu \propto \nu^{-\alpha}$ then Equation 6.5 reduces to

$$T_i = \frac{\mathrm{h}\nu_i}{k_{\rm B}}\frac{1}{2+\alpha}\,. \tag{6.6}$$

Assuming that HeI reionizes at the same time as HI, with HeII reionized later (i.e. $x_{\rm HeII} = 1$ and $x_{\rm HeIII} = 0$), then Equation 6.3 can be recast as

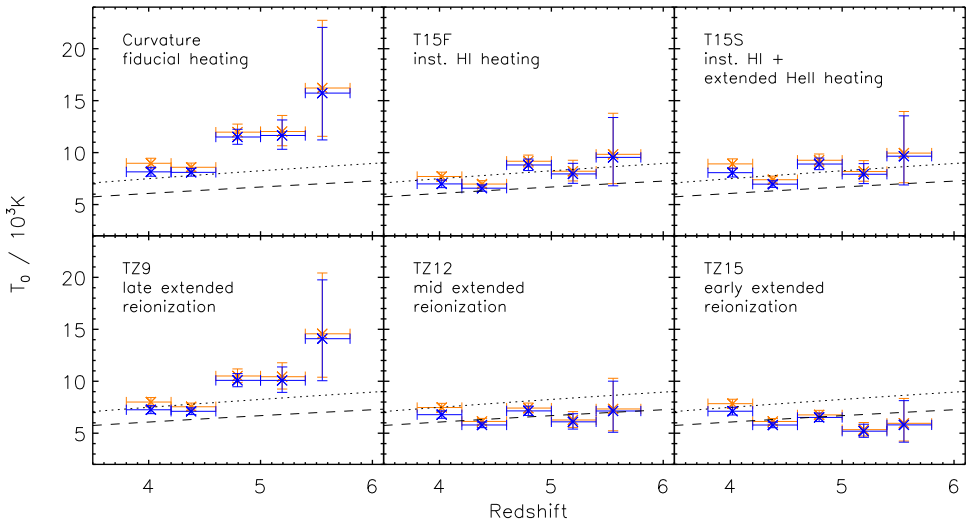
$$T_0 = 6.07\times10^3\left(\frac{2+\alpha}{5}\right)^{-1/1.7}\left(\frac{1+z}{5}\right)^{0.53}\ \mathrm{K}\,. \tag{6.7}$$

If HI reionization was carried out by Pop II stars in star-forming galaxies then $\alpha = 3$ for $\nu > \nu_{\rm HI}$ (e.g. Leitherer et al., 1999), and provided reionization happened early enough then T_0 is expected to be ~ 6000 K at $z = 4$. In comparison, for a harder, quasar-like spectrum with $\alpha = 1.5$ (e.g. Telfer et al., 2002) it would be ~ 7500 K.

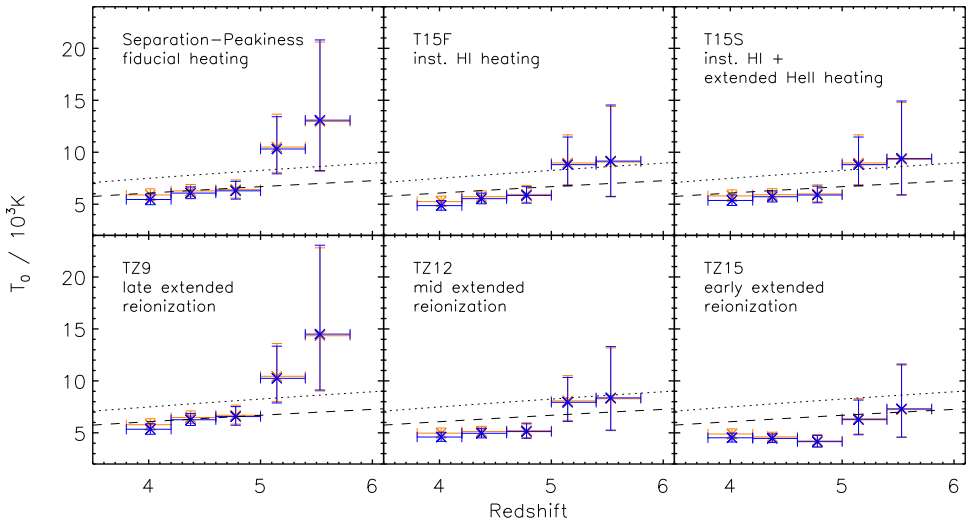
Early studies that measured the IGM temperature at $z \sim 4$ found rather high values of T_0 and

so concluded that they were observing the cooling from a recent reionization event (Theuns et al., 2002a,b; Hui and Haiman, 2003). However, Becker et al. (2011a) found lower temperatures at $4 < z < 5$, consistent with the asymptote due to a quasar-like spectrum. This suggested that T_0 had reached its asymptotic limit at $z \geq 5$, and thus there was no evidence for recent reionization. Turning to our own results, we find that the T_0 based upon curvature measurements at $4 < z < 4.6$ are consistent with the asymptote of a quasar-like spectrum, in agreement with the findings of Becker et al. (2011a). However at $z > 4.6$ we find a marked departure from this asymptote, hinting that we may be witnessing the cooling from HI reionization. As discussed in Section 6.4.3, however, our point at $z \sim 4.8$ is likely biased high due to our method's sensitivity to continuum placement errors. Our measured increase in T_0 for $z > 5$, though, is hard to explain away using similar arguments. This is further emphasised by the temperature evolution measured by the separation-peakiness statistic. Similar to Becker et al. (2011a), it finds a smooth asymptotic evolution in T_0 from $3.8 < z < 5$, but is consistent with a stellar ionizing spectrum. However, like our measurements with the curvature, it too suggests an increase in T_0 at $z > 5$. Since the peakiness is independent of continuum-fitting errors, it strongly suggests that the overall temperature evolution measured by the curvature is not caused purely by poor continuum placement.

Taken at face value, both the curvature and peakiness suggest that at $z \sim 5-6$ the IGM is cooling from HI reionization. Given the short-term thermal memory of the IGM, this implies a late reionization scenario, ending at a redshift not much higher than 6. However, as demonstrated in Section 6.3.7, Jeans smoothing due to early periods of heating can cause us to overestimate T_0, particularly at $z > 5$. Consequently, before we can make any firm conclusions about HI reionization, we must investigate how robust these initial interpretations are to the Jeans smoothing from the heating itself. Fig. 6.16 shows how the measured T_0 shift when correcting for the Jeans smoothing from a variety of thermal histories. For the T_0 based on the curvature there is a noticeable shift downwards in all bins for all thermal histories, except that of late reionization (model TZ9). This correction suppresses the apparent increase in T_0 between $z \sim 5-6$ seen in the raw measurements, and instead the results cluster around the late-time thermal asymptotes of quasar-like (dotted line) and star-like (dashed line) ionizing spectra. Models T15S and T15F show the limiting case of instantaneous reionization (at a middling redshift) with and without a heat injection at $z < 4.5$ from HeII reionization, respectively. By looking at models TZ9, TZ12 and TZ15, we can compare how late, middling and early reionization would appear in the data. If the redshift evolution of the corrected T_0 disagrees qualitatively with that required by the assumed thermal history, then that model can be tentatively discarded. Based upon the curvature-derived T_0, a middling reionization redshift is preferred (model TZ12, where reionization ends at $z \sim 9$), as the corrected temperatures are then all consistent with a single asymptote - that of Pop II stars. Instantaneous reionization, such as in models T15S and T15F, produces a smaller correction, and so shifts the points onto the thermal asymptote of a harder spectrum. However, even in the absence of a substantial population of high-redshift quasars, a hard ionizing spectrum could be possible due to a reasonable contribution from metal-free Pop III stars, or hardening of the ionizing spectrum as it is reprocessed by the IGM. As such, a short reionization epoch is difficult to rule out. Similarly, a late reionization cannot be ruled out (since it need not reach the asymptote before heating



(a) The T_0 measurements from the curvature statistic when corrected for the Jeans smoothing by a variety of thermal histories.



(b) The T_0 measurements from the separation-peakiness statistic when corrected for the Jeans smoothing by a variety of thermal histories.

Figure 6.16: Our measured T_0 after applying a correction due to the Jeans smoothing by an earlier epoch of heating, for several different thermal histories. The dotted line is the thermal asymptote expected from hard quasar-like ionizing sources with $\alpha = 1.5$, whilst the dashed line is that expected from star-like ionizing sources with $\alpha = 3.0$. All error bars are 2σ. *Top:* The corrected T_0 from the curvature statistic. Assuming only the fiducial heating of our calibration simulations (top left) the points cool from $z = 5.6$ to $z = 4.4$ and asymptote at the temperatures expected for a hard ionizing spectrum. However, correcting for the Jeans smoothing from an earlier period of heating lowers the T_0, particularly at $z > 5$ where the strong evolution can be completely suppressed. *Bottom:* The corrected T_0 from the separation-peakiness statistic. Assuming only the fiducial heating of our calibration simulations (top left) the points cool from $z = 5.6$ to $z = 4.8$ and asymptote at the temperatures expected for a soft, star-like ionizing spectrum. Similar to the curvature, correcting for the Jeans smoothing from an earlier period of heating lowers the T_0, particularly at $z > 5$, although this leads to thermal asymptotes that require very soft ionizing spectra. As such, scenarios with little Jeans smoothing are preferred, which for these thermal histories is best satisfied by a late reionization (bottom left).

again with He II reionization), nor can an early reionization (see Model C in Fig. 2 of Hui and Haiman, 2003). Overall, even though an extended ($\Delta z \gtrsim 3$) reionization completing at $z \sim 9$ is marginally preferred by the corrected T_0 from the curvature, neither the redshift evolution nor the absolute values can reject any of the other possible reionization histories.

To try and break this degeneracy between temperature and Jeans smoothing, we turn to the alternate measurements of the small-scale structure provided by the separation-peakiness statistic. As demonstrated in Fig. 6.12, the statistics suffer from somewhat different degeneracies in different temperature-Jeans smoothing regimes, so together they may provide complementary constraints. The corrections for Jeans smoothing are smaller than those based on the curvature, and so the points at $3.8 < z < 5.0$ keep their asymptotic behaviour. This provides a useful handle with which to test models. The instantaneous reionization models (T15F and T15S) shift the temperatures onto a thermal asymptote for an ionizing spectrum *softer* than a stellar source. This is the opposite problem to that faced with the corrected temperatures from the curvature. As discussed earlier, whilst several processes may harden the ionizing spectrum, softening it is difficult to achieve theoretically. We use this behaviour to argue against a short reionization. Similarly, for a middling reionization the thermal asymptote requires $\alpha \sim 4.5$, which is an implausibly soft ionizing spectrum. The T_0 corrected for an early reionization demonstrates a dip at $z \sim 5$. Whilst this behaviour does not by itself automatically eliminate the possibility (T_0 will approach the asymptote from below for early reionization scenarios), such a dip would be expected at $z \gtrsim 10$, and so it is inconsistent with that expected for an early reionization. We note too that an early reionization with a lower peak temperature than in model TZ15 would create less smoothing (and thus a smaller correction). However, in such a scenario, the measured increase in T_0 at $z \gtrsim 5$ remains, and this is again inconsistent with the expected thermal history for an early reionization. Our raw measured T_0 at $z \sim 4-5$ already imply a soft ionizing spectrum, so any thermal histories that would entail large temperature corrections due to Jeans smoothing are disfavoured. Out of our tested thermal histories, the smallest corrections are for a late reionization scenario. Within this model, the corrected T_0 at $z \sim 4-5$ lie upon the thermal asymptote expected from the ionizing spectrum of Pop II stars, and there is a factor of ~ 2 increase in T_0 from $z \sim 5.0$ to $z \sim 5.6$. This increase is consistent with that suggested by the drop in $\Gamma_{\rm bkg}$ measured by Calverley et al. (2011) over the same redshift range, and could be interpreted as the temperature evolution expected after an extended, late hydrogen reionization. However, we have not tested the sensitivity of the peakiness statistic to spatial fluctuations in the UV background, variations in σ_8, nor have we carried out any box size or resolution convergence tests. We leave it to future work to establish the importance of these parameters, although they could have a noticeable impact on the size of our error bars, and thus increase the range of permitted reionization scenarios.

Upon the completion of reionization the temperature drops rapidly due to the adiabatic expansion of the Universe, before reaching thermal equilibrium with sources of photoheating. During this initial cooling, the temperature should vary as $T_0 \propto (1+z)^2$. Fitting this model to our measured redshift evolution in T_0 at $z \geq 5$ should therefore constrain the redshift the Universe was last at a given temperature. Thus for an assumed peak temperature, this can give a rough indication of when reionization ended. Using the T_0 from the separation-peakiness statistics, corrected for the Jeans smooth-

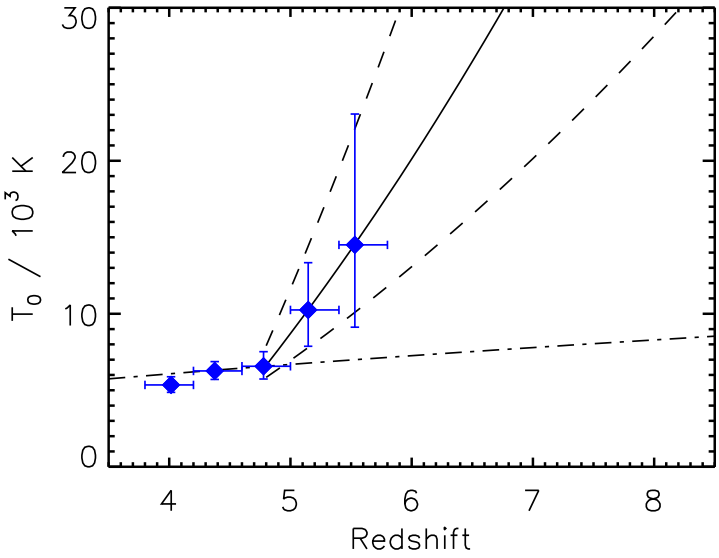


Figure 6.17: Suggested thermal evolution at higher redshift based upon the measured separation-peakiness statistics, assuming our fiducial γ and Jeans smoothing consistent with late reionization. The points at $z \gtrsim 4.8$ are well fit by a line that varies as $(1+z)^2$, suggesting adiabatic cooling. The solid line represents the best fit through the points. If T_0 peaks at a value of $2-3\times10^4$ K at the end of reionization (Furlanetto and Oh, 2009), and cools only adiabatically afterwards, then these temperatures would be reached at $z\sim6-7$, suggesting a late reionization. The dashed lines represent other allowed thermal evolutions consistent with our 2σ errors, and the dash-dotted line is the late-time thermal asymptote expected from HI and HeI reionization by a soft stellar ionizing spectrum.

ing expected from late reionization, the points at $z \gtrsim 4.8$ are well described by adiabatic cooling (see Fig. 6.17). Within a simple model where this evolution continues to higher redshift, then $T_0 = 2\times10^4$ K at $z = 6.0$ and $T_0 = 3\times10^4$ K at $z = 6.8$. Thus, for reasonable peak temperatures for HI reionization (see Furlanetto and Oh, 2009) our toy model suggests reionization ended at $6 \lesssim z \lesssim 7$, although the large error bars on our $z > 5$ points mean they are consistent with a wide range of other thermal histories. We caution that our model is highly simplified and ignores any deviations from pure adiabatic cooling after the peak T_0. Such deviations might be expected from an extended, inhomogeneous reionization (Gleser et al., 2005; McQuinn et al., 2009), and would push the peak temperature to higher redshift. Improving the precision of our $z > 5$ temperatures with more quasar sightlines is the best way to further constrain the allowed thermal histories, combined with independent measurements of the amount of Jeans smoothing using close quasar pairs (Peeples et al., 2010).

6.5 Conclusions

We have presented measurements of the temperature of the intergalactic medium (IGM) at $3.8 < z < 5.8$ using the small-scale structure in the Lyα forest. These are the first measurements of the temperature in the general IGM at $z > 5$.

The small-scale structure of the forest was quantified using two independent statistics: the curva-

ture, which has been used successfully by Becker et al. (2011a) at lower redshifts, and a new statistic called the peakiness. Importantly, the peakiness does not require accurate continuum fitting, which is a difficult procedure at $z > 5$. These parameters were measured for sections of forest $10h^{-1}$ comoving Mpc long, and then compared with simulations in order to convert them into a temperature.

We measured the temperature at an optimal overdensity, $T(\overline{\Delta})$, where the temperature is a near one-to-one function of both statistics. From this we inferred the temperature at mean density, T_0, for an assumed slope of the temperature-density relation, γ. Each statistic traces gas near the mean density, meaning that $T(\overline{\Delta}) \approx T_0$ at $z > 5$, and so we avoid systematic biases due to the large uncertainties in γ at high redshift.

Both statistics measure a decline in $T(\overline{\Delta})$ from $z = 5.6$ to 3.8. Since we are largely independent of γ, our redshift evolution of T_0 is very similar. The T_0 as measured with the curvature are roughly $\sim 2-3 \times 10^3$ K higher than those measured with the separation-peakiness statistics. The curvature measurements suggest temperatures at $z < 4.6$ that are consistent with the thermal asymptote expected if the sources responsible for HI reionization had a hard quasar-like spectrum. Those from the separation-peakiness measurements are instead consistent with the thermal asymptote of a stellar-like spectrum. The value of T_0 at $z \sim 4$ as measured with the separation-peakiness statistic is lower than any equivalent value of T_0 reported at similar redshifts in the literature.

Heating associated with HI reionization can smooth out the underlying gas density distribution. This is known as Jeans smoothing, and can introduce a substantial bias into our T_0 measurements, particularly at $z > 5$. The apparent decrease in T_0 from $z = 5.6$ to 4.8 can be reproduced by a non-evolving T_0 combined with an IGM smoothed by a period of extended heating at $z > 9$. The T_0 measurements derived from the curvature, corrected for the Jeans smoothing by a number of possible thermal histories, do not favour any one model. The measurements from the separation-peakiness statistic favour only small corrections for Jeans smoothing, as otherwise the corrected T_0 at $z \sim 4-5$ suggest an implausibly soft ionizing spectrum. Out of the models tested, this constraint is only satisfied by a late reionization scenario. In that case, we recover a thermal evolution at $z \gtrsim 5$ suggestive of the rapid adiabatic cooling expected after the end of HI reionization. This measured change in temperature from $z \sim 5$ to $z \sim 6$ by a factor of $\sim 2-3$ is consistent with the drop in the ionizing background by a factor ~ 5 measured by Calverley et al. (2011) over the same redshift range. Extrapolating this thermal evolution to higher redshifts within a simple model suggests that for a range of realistic peak IGM temperatures, HI reionization may have completed at $6 \lesssim z \lesssim 7$. The large uncertainty in our T_0 at $z > 5$ does however allow for a range of other possible reionization histories, and so we hope to improve the constraints on these points via more quasar sightlines, and independent measurements of the Jeans smoothing with close quasar pairs.

The results presented here represent the highest redshift measurements of the temperature of the IGM to date, and offer a tantalising glimpse of one of the smoking guns of hydrogen reionization. Future observations with the next generation of optical and near-infrared telescopes will be able to take many high-S/N, high-resolution spectra of faint quasars at $z \sim 5-6$, vastly improving the data set available for this kind of measurement, and so help us to formulate a more robust picture of both when and how HI reionization occurred.

All truths are easy to understand once they are discovered; the point is to discover them.

GALILEO GALILEI

7 Conclusions

QUASARS are highly luminous objects, and their spectra provide a wealth of information about the gas between them and the Earth. This gas is known as the intergalactic medium (IGM). It holds a significant fraction of baryons in the Universe, and its diffuse nature means it is an excellent tracer of large-scale structure. In this thesis, I have used high-resolution quasar spectra in conjunction with state-of-the-art hydrodynamical simulations to probe the IGM at high redshift, focusing on the ionization and thermal state of the gas. In Section 7.1 I outline the main conclusions from my thesis, and the constraints placed on hydrogen reionization. Finally, in Section 7.2 I look at where future progress may be made in answering some of the unresolved questions in the field.

7.1 Main conclusions

After reionization, the ionization state of the IGM is set by the intensity of the metagalactic ultraviolet background (UVB), the net ionizing radiation field that pervades the Universe. In the literature, the UVB intensity has been estimated using two different methods: measuring the mean flux within the Lyα forest and comparing it to simulations, and measuring the strength of the 'proximity effect' blueward of the Lyα emission line in quasar spectra. Both of these methods are investigated in this thesis.

In Chapter 3, the mean-flux method is tested against a new technique that uses the entire flux probability distribution function (PDF). With this PDF, a probability is determined for the flux of each pixel in a spectrum and a maximum-likelihood approach is used to calculate the best-fitting UVB intensity, as measured by the hydrogen photoionization rate, $\Gamma_{\rm bkg}$. This test is done on simulated

spectra of the Lyα forest at $z = 2-6$, for which the gas density distribution and thermal state of the gas are known. Comparison of the results shows that the likelihood function is driven by pixels near the mean flux, and so it is the shape of the PDF near the mean that is the most important. The mean-flux method implicitly assumes a flat PDF, so the maximum-likelihood technique can only improve upon it when this is a poor approximation of the true shape. This is true when the flux PDF is sharply peaked close to the mean, such as when the majority of pixels are near zero flux or the continuum. However, noise smoothes out these steep gradients so that for spectra with a signal-to-noise ratio S/N ≤ 20 this advantage is essentially eliminated. Consequently, the precision of estimates of $\Gamma_{\rm bkg}$ are only marginally increased for very high S/N spectra when using the entire flux PDF, in comparison to only using the mean flux. It should be noted, however, that this was only tested in the limiting case of perfect knowledge about the underlying gas density distribution and thermal state of the IGM. In practice, uncertainties in estimates of $\Gamma_{\rm bkg}$ from the mean flux are dominated by the uncertainties in these two parameters, particularly at high redshift. Importantly, the shape of the flux PDF can be used to constrain the density and temperature distributions, and thus knowledge of the entire flux PDF *can* improve the precision in observational estimates of $\Gamma_{\rm bkg}$.

The classical method of measuring $\Gamma_{\rm bkg}$ is through studies of the 'proximity effect' in quasar spectra. This effect is due to the ionizing photons emitted by the quasar, which decrease the neutral fraction in the local gas, resulting in enhanced transmission blueward of the Lyα emission line (the 'proximity region'). In Chapter 4, the strength of this effect is quantified by introducing the parameter $R_{\rm eq}$, which represents the line-of-sight distance away from the quasar at which its ionizing output equals that of the UVB. It is calculated by correcting the optical depths in the proximity region for the expected effect of the quasar, until the mean flux in the proximity region equals the observed flux in the Lyα forest. $\Gamma_{\rm bkg}$ is then estimated by combining the measured $R_{\rm eq}$ with an estimate of the luminosity at the Lyman limit, L_{ν_0}. These results are the highest-redshift proximity-effect measurements of $\Gamma_{\rm bkg}$ to date, and are consistent with estimates based upon the mean flux. There is a large decline in $\Gamma_{\rm bkg}$ with redshift, from $\log(\Gamma_{\rm bkg}) = -12.15 \pm 0.16$ at $z \sim 5$ to $\log(\Gamma_{\rm bkg}) = -12.84 \pm 0.18$ at $z \sim 6$. The measured errors are dominated by uncertainties in the systemic quasar redshifts and in the mean flux of the Lyα forest. The UVB intensity measured from individual sightlines is seen to decline smoothly with redshift over $4.6 < z < 6.4$. It shows no sign of the rapid decline which may be expected in the late stages of reionization, when there is a rapid change in the attenuation length of ionizing photons. When combined with estimates of the evolution of the mean free path of ionizing photons, the measured decline in $\Gamma_{\rm bkg}$ suggests a decline in the ionizing emissivity of roughly a factor of two from $z \sim 5$ to 6.

At $z \sim 2-4$, estimates of $\Gamma_{\rm bkg}$ from the proximity effect are systematically higher than those based upon the mean flux of the Lyα forest. This is often attributed to the environment the quasar resides in. In Chapter 5, several sources of potential systematic bias are investigated to test the sensitivity of the proximity effect to the quasar environment at $z \sim 5-6$. The most luminous quasars are thought to reside in large dark matter haloes of mass $\sim 10^{13}$ M$_\odot$ at $z \sim 6$. Such an overdensity will change the density distribution of the IGM in their vicinity. The quasar will therefore ionize a smaller volume than if it was in a region of the IGM with a density closer to the global mean density. As a result, the de-

rived value of $R_{\rm eq}$ would be underestimated, and so $\Gamma_{\rm bkg}$ overestimated. However, the bright quasars studied in this thesis all have a derived $R_{\rm eq}$ much larger than the extent of the largest overdensities at these redshifts. Consequently, the average density within the proximity region is only slightly above the mean density, and so the overestimate in $\Gamma_{\rm bkg}$ is likely to be small (~ 0.1 dex). It should be noted that the situation is somewhat different at $z \sim 2-4$, where the overdensities will have grown through hierarchical structure formation, and the UVB is higher. At these lower redshifts, $R_{\rm eq}$ will be comparable in size to the extent of the overdensity, and thus the average density in the proximity region may be noticeably more than the global mean density. This leads to an overestimation in $\Gamma_{\rm bkg}$, and hence the observed discrepancy with independent measures of the UVB from the mean flux. The contribution of low-luminosity ionizing sources clustered around quasars is negligible at $z \sim 5-6$. Parcels of optically-thick neutral gas in the vicinity of a quasar, known as Lyman limit systems (LLSs), will modify the $1/r^2$ drop-off in ionizing flux assumed by the proximity effect. This again will truncate the size of the proximity region, and so $\Gamma_{\rm bkg}$ will be overestimated. However, even for a rather generous estimate of the frequency of LLSs, the overestimate in $\Gamma_{\rm bkg}$ is small (~ 0.1 dex). Finally, the hard ionizing photons from the quasar will locally ionize He II and thus boost the temperature in the proximity region. This increases the transmission, so $R_{\rm eq}$ is *overestimated* and $\Gamma_{\rm bkg}$ *underestimated*. Based on reasonable assumptions about the size of this temperature boost, the underestimate in $\Gamma_{\rm bkg}$ is comparable to the combined effect of the other two parameters (~ 0.2 dex) but in the opposite sense. The proximity effect should therefore give relatively unbiased estimates of $\Gamma_{\rm bkg}$ at $z > 5$, in contrast to lower redshifts, and so differences with other independent measures are unlikely to be due to the quasar environment.

The statistical properties of the Lyα forest are determined by both cosmological and astrophysical parameters. Out of the latter, the two most important are the ionization and thermal state of the gas. Whilst Chapters 3–5 concentrate on the ionization state, Chapter 6 investigates the thermal state of the gas. This is intricately linked to reionization, since photoionization heats the gas. Photoionization equilibrium of the IGM is reached after reionization, and competing heating and cooling processes produce a tight power-law temperature-density ($T-\Delta$) relation in low-density gas of the form $T(\Delta) = T_0 \Delta^{\gamma-1}$, where T_0 is the temperature at mean density. Varying T_0 or γ will affect the level of small-scale structure in the Lyα forest, so measuring the observed level can constrain the $T-\Delta$ relation. In this chapter, the small-scale structure is quantified using two independent statistics: the 'curvature', and the 'peakiness'. The former is essentially the second derivative of the flux and has been used successfully at lower redshifts. The latter, meanwhile, is newly developed for analysing the high-redshift Lyα forest and measures the relative depth of a minimum between two maxima in the flux spectrum. Both statistics are measured for short sections of forest, and then compared with simulations in order to convert them into a temperature. These are the first measurements of the temperature in the general IGM at $z > 5$. The temperature is measured at a well-chosen overdensity, $\overline{\Delta}$, for which $T(\overline{\Delta})$ is independent of γ, and T_0 is then calculated for reasonable assumptions for γ. Since $\overline{\Delta}$ is very close to the mean density, $T(\overline{\Delta}) \approx T_0$. Both statistics measure a decline in T_0 from $z = 5.8$ to 3.8, although those returned by the peakiness are $\sim 2000-3000$ K lower than those from the curvature, and are lower at $z \sim 4$ than any other T_0 in the literature. The curvature measurements suggest

temperatures at $z < 4.6$ that are consistent with the thermal asymptote expected if the sources responsible for H I reionization have a hard quasar-like spectrum. Those from the peakiness measurements are instead consistent with the thermal asymptote of a stellar-like spectrum. The measured change in temperature from $z \sim 5$ to 6 by a factor of $\sim 2-3$ is consistent with the drop in $\Gamma_{\rm bkg}$ by a factor ~ 5 as presented in Chapter 4. This is calculated by scaling the assumed T_0 used in estimates of $\Gamma_{\rm bkg}$ derived from the mean flux, until those estimates match the $\Gamma_{\rm bkg}$ derived from the proximity effect, since any difference between the two is not likely to be due to the quasar environment. The curvature and peakiness measurements over $z = 5.6$ to 4.8 can alternatively be reproduced by a non-evolving T_0 if the IGM is smoothed by a period of extended heating at $z > 9$. The T_0 measurements derived from the curvature, corrected for the Jeans smoothing by a number of possible thermal histories, do not favour any one model. The measurements from the peakiness statistics favour only small corrections for Jeans smoothing, as otherwise the corrected T_0 at $z \sim 4-5$ suggest an implausibly soft ionizing spectrum. Out of the models tested, this constraint is only satisfied by a late reionization scenario. In that case, the recovered thermal evolution at $z > 5$ is suggestive of the rapid adiabatic cooling expected after the end of hydrogen reionization.

Implications for hydrogen reionization

The work presented in this thesis in Chapters 4 and 6 provides valuable constraints on hydrogen reionization. In Chapter 4, a smooth decline in $\Gamma_{\rm bkg}$, as measured from individual sightlines, is seen over $4.6 < z < 6.4$ (see Fig. 4.7). Since the UVB intensity is proportional to the mean free path of ionizing photons, this implies that the mean free path also evolves smoothly over the same redshift range. At the very end of reionization, the mean free path is expected to evolve rapidly as the H II bubbles begin to overlap. Consequently, the smooth redshift evolution of $\Gamma_{\rm bkg}$ suggests that overlap happens at $z > 6.4$. However, when combined with estimates of the mean free path based upon the abundance of LLSs, the drop in $\Gamma_{\rm bkg}$ from $z \sim 5-6$ by a factor of ~ 5 corresponds to a drop in the ionizing emissivity by a factor of ~ 2. The emissivity is already very low at $z = 6$ and so this evolution cannot continue to much higher redshifts without reionization failing to complete by $z = 6$. Thus, unless there is a substantial increase in the emissivity at very high redshifts, the end of reionization, while potentially before $z = 6.4$, appears unlikely to occur much earlier.

In Chapter 6, T_0 is measured to increase by about a factor of two from $z = 4.4$ to 5.6 (see Fig. 6.13), regardless of the statistic used to determine the temperature. This result could possibly be due to a smoother underlying gas density distribution in the IGM than we have used in the simulations. However, large corrections for this effect lower the temperatures at $z \sim 4-5$, as measured by the peakiness, to values that require an implausibly soft ionizing spectrum. Therefore, small corrections are preferred. There are two main impacts of this constraint on the size of the corrections. First, corrections are small if reionization is late, very rapid, or has a cool peak temperature, or if any combination of these factors are true. Secondly, if the corrections are small, then the rough factor of two change in T_0 at $z > 5$ is still detected. Since the IGM retains a relatively short-term memory of the heating by reionization, the measured redshift evolution of T_0 implies a recent heating event, which peaked

in temperature at $z \gtrsim 6$. During reionization the peak temperature is expected close to the end of overlap, so this could be taken as evidence for a late reionization, with overlap at $z \gtrsim 6$. However, a tentative upper limit can be placed on when overlap occurred, based upon assumptions about the redshift evolution of T_0 and the peak temperature. In the simplest model of the thermal history spanning reionization, after reaching a peak temperature the IGM cools adiabatically due to the expansion of the Universe. The temperatures returned by the peakiness statistic at $z \geq 4.8$, after correcting for the smoothing of the gas density distribution by a late reionization, are well fit by an adiabatic cooling curve. Assuming this continues to higher redshift, and assuming a peak temperature during overlap of $2-3\times10^4$ K, then within the framework of this simple model. overlap may have been at $6 \lesssim z \lesssim 7$, although the range $5.5 \lesssim z \lesssim 8$ is consistent with the 2σ errors on the measured T_0 (see Fig. 6.17). Finally, the corrected temperatures at $z \sim 4-5$ are consistent with the dominant ionizing sources responsible for hydrogen reionization being Pop II stars within star-forming galaxies, rather than sources with hard spectra such as active galactic nuclei (AGN).

The key conclusions from this thesis, with regard to hydrogen reionization, are therefore the following:

- A smooth redshift evolution in $\Gamma_{\rm bkg}$ suggests that the end of reionization is at $z > 6.4$.
- The corresponding drop in the ionizing emissivity implies, however, that the end of reionization can not be much higher than 6.4, otherwise it would fail to complete.
- An increase in the temperature of the IGM by about a factor of two between $z = 4.4-5.6$ may be the first detection of the adiabatic cooling expected after the end of reionization. If so, then for reasonable peak temperatures, overlap may have occurred at $6 \lesssim z \lesssim 7$.
- The temperature of the IGM at $z \sim 4-5$ implies that hydrogen reionization was carried out by ionizing sources with soft spectra, such as Pop II stars.

The overall picture, therefore, is of a late, extended reionization, with most of the ionizing photons coming from high-redshift, star-forming galaxies. Together, the redshift evolution of the UVB intensity and the IGM temperature suggest that reionization possibly ended at $6.5 \lesssim z \lesssim 7$. Although the temperature evolution would be consistent with overlap occurring at higher redshifts, this would be difficult to achieve considering the evolution in the ionizing emissivity. It should be emphasised that this upper limit on the redshift at the end of reionization requires several simplifying assumptions about the temperature evolution of the IGM after reionization, and ignores the possibility of a substantial increase in the emissivity at very high redshifts. Modifying either of these assumptions can accommodate overlap at higher redshifts. However, there are reasons to believe that overlap at $6.5 \lesssim z \lesssim 7$ is not unreasonable. The recent measurement of the neutral fraction, $x_{\rm HI} > 0.1$, from the near zone of the $z \sim 7.1$ quasar by Mortlock et al. (2011) suggests that, at least along that sightline, overlap happened at $z < 7.1$. The truncated near zone, however, could be due to a LLS very close to the quasar, and although this happens only rarely in simulations (Bolton et al., 2011), more quasars need to be observed in this redshift range to verify this result. Similarly, theoretical models designed

to match the observed star formation rate history also suggest that hydrogen reionization ended close to redshift seven (Haardt and Madau, 2011; Shull et al., 2011). Overall, it appears that observational evidence for the redshift of overlap could soon be within our grasp.

7.2 Outlook

The work in this thesis has raised several possible avenues for future research. Measurements of the proximity effect in the spectra of more quasars at $z \sim 5-6$ will decrease the statistical errors on the UVB, and the accuracy of these measurements can be improved through better modelling of the proximity region. The most pressing work, however, is on the thermal state of the IGM. As demonstrated in Chapter 6, at $z > 5$ there is a strong degeneracy between the instantaneous temperature, as measured by both the curvature and peakiness statistics, and the smoothness of the gas density distribution. On the other hand, this redshift range is also where a strong rise is detected in T_0. The robustness of this result, therefore, depends on determining the smoothness of the gas. As shown in that chapter, large corrections for Jeans smoothing would lower the measured temperatures at $z \sim 4-5$ onto the thermal asymptote for an implausibly soft ionizing spectrum. However, this is only for temperatures measured from the peakiness, and there may be systematic errors that are unaccounted for that would alleviate this discrepancy. Consequently, an independent measure of the Jeans smoothing at $z > 5$ is required. Close quasar pairs are a possible way of approaching this, as if their Lyα forest spectra are similar at a given redshift, then the Jeans length is greater than the transverse separation between the sightlines (Peeples et al., 2010). To probe the expected scales, though, requires sightlines separated by $\Delta\theta \lesssim 15$ arcsec. Unfortunately, there are no known quasar pairs at $z > 5$ with this separation (for example, the quasars used by Gallerani et al., 2008a, to claim a detection of the transverse proximity effect at $z = 5.65$ were separated by 109 arcsec). However, an early reionization will still leave the IGM substantially smoother at $z \sim 4-5$ than a late reionization (see Fig. 6.11), so measurements of the Jeans smoothing at these redshifts may be our best prospect, since there is a greater chance of finding a sufficiently close quasar pair. Constraints on the Jeans smoothing may also come from the detection of excess power in the flux power spectrum of the Lyα forest on scales of the Jeans length, which is expected to be larger than the thermal broadening scale.

The coming decade is likely to be full of very exciting advances in our understanding of hydrogen reionization and the high-redshift Universe. In the previous section I specified the redshift range $6.5 \lesssim z \lesssim 7$ as potentially containing the end of reionization, and in the next few years we should start to have a much clearer picture as to whether or not that is true. Quasar spectra will be one of the best probes of this period, as quasars can be bright enough for high-S/N spectroscopy even in this redshift range. Although the Gunn-Peterson trough saturates for a neutral fraction of order $\sim 10^{-3}$, only a substantially more neutral IGM (neutral fraction of order $\sim 10^{-1}$) will lead to a damping wing profile blueward of Lyα (see Fig. 7.1 and Miralda-Escudé, 1998). Consequently, detection of such a damping wing within quasar spectra would be strong evidence that the IGM at that redshift is being observed before the end of reionization. Bright quasars become increasingly rare at high redshift, however, and at $z > 6.5$ the Lyα emission line moves into the infrared. The discovery of quasars in this

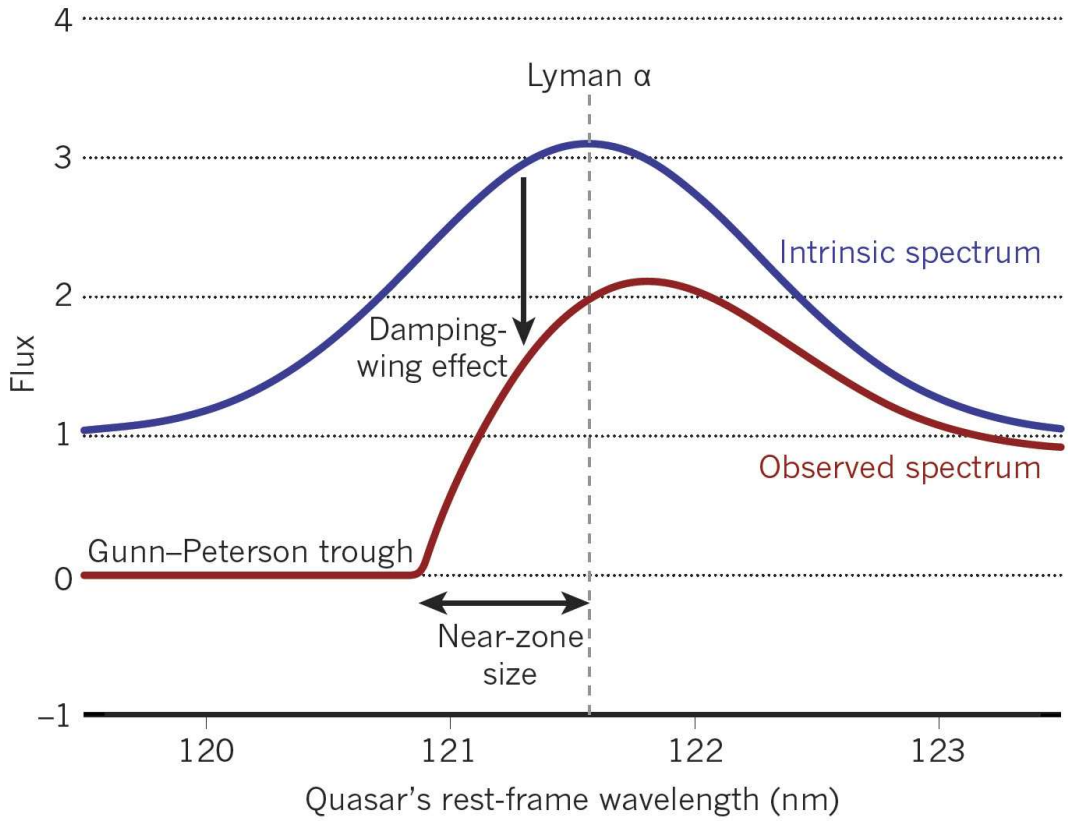


Figure 7.1: The Lyα emission line spectrum of a quasar embedded in a highly neutral IGM. The blue line is the intrinsic emission, whilst the red line shows the observed spectrum after absorption by the IGM. Although the Gunn-Peterson trough can become saturated in a highly ionized IGM, the presence of a damping wing is indicative of a substantially more neutral IGM. Detection of such a damping wing in several quasars would be strong evidence that reionization was not complete by the quasar redshifts. Figure taken from Willott (2011).

epoch therefore requires large-area, near-infrared (NIR) surveys. Two such surveys that are currently ongoing are the Large Area Survey (LAS) with the UK Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS)[1] and the Visible and Infrared Survey Telescope for Astronomy (VISTA)[2] Kilo-Degree Infrared Galaxy Survey (VIKING). The former was responsible for the discovery of the $z \sim 7.1$ quasar by Mortlock et al. (2011), where the presence of a damping wing has been used to suggest that we are seeing the first quasar found during the epoch of reionization. Very recently, it was also announced that the latter has discovered two quasars at $z \sim 6.8$ and $z \sim 6.9$ (Venemans et al., in prep.). NIR surveys such as these are dependent upon deep optical data in order to aid in the identification of low-redshift galaxies and stars that can mimic high-redshift quasars in their NIR photometry. Many current NIR surveys can be supplemented with optical data from the SDSS, although several new large-area optical-NIR surveys are planned (or already underway). One of the most notable is the Panoramic Survey Telescope and Rapid Response System[3] (Pan-STARRS) survey, which will consist of four 1.8-m telescopes, the first of which began full time science observations in May 2010. When complete, it will survey an area equivalent to the entire sky visible from Hawaii four times a month,

[1] http://www.ukidss.org/
[2] http://www.vista.ac.uk/
[3] http://pan-starrs.ifa.hawaii.edu/public/home.html

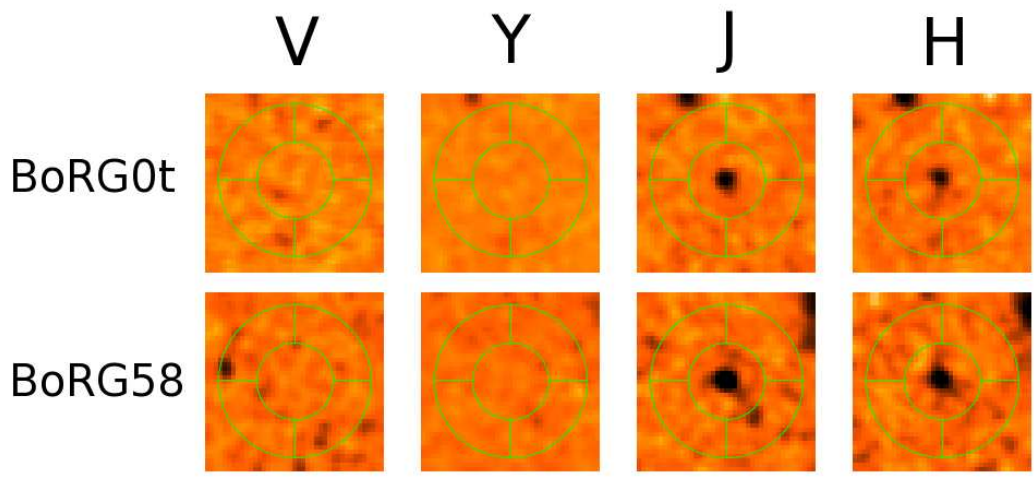


Figure 7.2: The most robust candidates for high-redshift galaxies from the first results of the Brightest of Reionizing Galaxies (BoRG) survey. They were found using the Lyman-break technique as Y-dropouts, which would correspond to $z \sim 8$. BoRG58 (bottom panels) is bright enough to allow spectroscopic follow-up, and so makes a good target for learning about the galaxy population potentially responsible for reionization. Figure taken from Trenti et al. (2011).

to two magnitudes deeper and longer wavelengths than the SDSS. Looking further ahead, the Large Synoptic Survey Telescope[4] (LSST) is an 8.4-m telescope due to begin surveying the sky before the end of the decade to even fainter magnitudes than Pan-STARRS. Also, planned space-based instruments, such as the Wide Field Infrared Survey Telescope[5] (WFIRST), will provide large-area deep NIR surveys that, together with the optical data from the LSST, should find many high-redshift quasar candidates. More quasars at $z > 6.5$, found with surveys such as these, will help establish if reionization does indeed end at $6.5 \lesssim z \lesssim 7$.

At higher redshifts, substantial progress has been made in identifying galaxies at $z \sim 7-10$ using the Lyman-break technique. If reionization ends close to redshift 7, then these Lyman-break galaxies (LBGs) are possibly the population responsible for reionizing the Universe. Unfortunately, most of the photometrically-selected candidates are faint, making it difficult to spectroscopically confirm their redshifts. Lehnert et al. (2010) claim to spectroscopically confirm a galaxy at $z = 8.56$ via its Lyα emission line, but this detection is controversial. The Brightest of Reionizing Galaxies (BoRG) survey[6] aims to find LBGs that are bright enough for secure spectroscopic confirmation. Fig.7.2 shows the two most robust Y-dropout ($z \sim 8$) candidates found from the first results of the survey (Trenti et al., 2011). Importantly, BoRG58 (bottom panel) is roughly an order of magnitude brighter than the galaxy observed by Lehnert et al. (2010), making it an ideal target for spectroscopic follow-up. Surveys such as these will allow us to securely identify galaxies that are potentially deep within the epoch of reionization.

Much of the progress in finding these galaxies has come from the leap in sensitivity in the infrared from the installation of the Wide Field Camera 3 (WFCAM3) on the Hubble Space Telescope (HST).

[4] http://www.lsst.org/lsst
[5] http://wfirst.gsfc.nasa.gov/
[6] http://casa.colorado.edu/ trenti/BoRG/index.html

However, the James Webb Space Telescope (JWST), successor to the HST and due to launch in 2018, will provide a similar leap again in sensitivity in the infrared. With JWST, spectroscopy of $z \sim 8-10$ galaxies should become more routine, and some galaxies may even be identified via photometry out to $z \sim 14$. The next generation of ground-based Extremely Large Telescopes (ELTs), such as the Giant Magellan Telescope (GMT), the Thirty Meter Telescope (TMT), and the European Extremely Large Telescope (E-ELT), all with mirror diameters in excess of $\sim$ 25-m, will complement much of the work done by JWST. Their sensitivity and resolution will enable us to study the high-redshift Universe in unprecedented detail, so in the next few decades we stand to learn a great deal about the population of galaxies that reionized the Universe.

Finally, JWST and the ELTs are fundamentally limited to only luminous objects, and so the last word on reionization, and the Dark Ages that preceeded it, may instead come from 21 cm radiation. This is emitted (or absorbed) by the neutral hydrogen gas itself, and so serves as a direct tracer of neutral gas in the Universe. Radio telescopes currently being built, such as the Low Frequency Array (LOFAR), may detect the signal from pre-reionization neutral hydrogen before the end of the current decade. Future radio telescopes such as the Square Kilometer Array (SKA) may even be able to observe out to higher redshifts than JWST, perhaps giving the first glimpses of the Dark Ages. With instruments such as these, many of the questions surrounding reionization may finally be answered.

BIBLIOGRAPHY

Aarseth, S. J.: 1963, Dynamical evolution of clusters of galaxies, I, *MNRAS* **126**, 223

Abel, T. and Haehnelt, M. G.: 1999, Radiative Transfer Effects during Photoheating of the Intergalactic Medium, *ApJ* **520**, L13

Adelberger, K. L., Steidel, C. C., Shapley, A. E., and Pettini, M.: 2003, Galaxies and Intergalactic Matter at Redshift $z \sim 3$: Overview, *ApJ* **584**, 45

Alpher, R. A., Bethe, H., and Gamow, G.: 1948, The Origin of Chemical Elements, *Physical Review* **73**, 803

Alpher, R. A. and Herman, R.: 1948, Evolution of the Universe, *Nature* **162**, 774

Antonucci, R.: 1993, Unified models for active galactic nuclei and quasars, *ARA&A* **31**, 473

Arons, J.: 1972, Low-Mass Protogalaxies and Absorption Lines in Quasi-Stellar Objects., *ApJ* **172**, 553

Arons, J. and McCray, R.: 1970, Photo-Ionization of Intergalactic Hydrogen by Quasars, *Astrophys. Lett.* **5**, 123

Arons, J. and Wingert, D. W.: 1972, Theoretical Models of Photoionized Intergalactic Hydrogen, *ApJ* **177**, 1

Aykutalp, A. and Spaans, M.: 2011, The Complexity that the First Stars Brought to the Universe: Fragility of Metal-enriched Gas in a Radiation Field, *ApJ* **737**, 63

Baade, W.: 1944, The Resolution of Messier 32, NGC 205, and the Central Region of the Andromeda Nebula., *ApJ* **100**, 137

Bahcall, J. N. and Salpeter, E. E.: 1965, On the Interaction of Radiation from Distant Sources with the Intervening Medium., *ApJ* **142**, 1677

Bajtlik, S., Duncan, R. C., and Ostriker, J. P.: 1988, Quasar ionization of Lyman-alpha clouds - The proximity effect, a probe of the ultraviolet background at high redshift, *ApJ* **327**, 570

Barkana, R. and Loeb, A.: 2001, In the beginning: the first sources of light and the reionization of the universe, *Phys. Rep.* **349**, 125

Barkana, R. and Loeb, A.: 2005, A Method for Separating the Physics from the Astrophysics of High-Redshift 21 Centimeter Fluctuations, *ApJ* **624**, L65

Barkana, R. and Loeb, A.: 2007, The physics and early history of the intergalactic medium, *Reports on Progress in Physics* **70**, 627

Bechtold, J., Weymann, R. J., Lin, Z., and Malkan, M. A.: 1987, The integrated ultraviolet radiation field from QSOs, *ApJ* **315**, 180

Becker, G. D., Bolton, J. S., Haehnelt, M. G., and Sargent, W. L. W.: 2011a, Detection of extended He II reionization in the temperature evolution of the intergalactic medium, *MNRAS* **410**, 1096

Becker, G. D., Rauch, M., and Sargent, W. L. W.: 2007, The Evolution of Optical Depth in the Lyα Forest: Evidence Against Reionization at $z \sim 6$, *ApJ* **662**, 72

Becker, G. D., Sargent, W. L. W., Rauch, M., and Calverley, A. P.: 2011b, High-redshift Metals. II. Probing Reionization Galaxies with Low-ionization Absorption Lines at Redshift Six, *ApJ* **735**, 93

Becker, G. D., Sargent, W. L. W., Rauch, M., and Simcoe, R. A.: 2006, Discovery of Excess O I Absorption toward the $z = 6.42$ QSO SDSS J1148+5251, *ApJ* **640**, 69

Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., and York, D. G.: 2001, Evidence for Reionization at $z \sim 6$: Detection of a Gunn-Peterson Trough in a $z = 6.28$ Quasar, *AJ* **122**, 2850

Benson, A. J. and Bower, R.: 2010, Galaxy formation spanning cosmic history, *MNRAS* **405**, 1573

Berger, E.: 2009, The Host Galaxies of Short-Duration Gamma-Ray Bursts: Luminosities, Metallicities, and Star-Formation Rates, *ApJ* **690**, 231

Bernstein, R., Shectman, S. A., Gunnels, S. M., Mochnacki, S., and Athey, A. E.: 2003, in M. Iye and A. F. M. Moorwood (eds.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4841 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp 1694–1704

Bertschinger, E.: 1998, Simulations of Structure Formation in the Universe, *ARA&A* **36**, 599

Bi, H.: 1993, Lyman-alpha absorption spectrum of the primordial intergalactic medium, *ApJ* **405**, 479

Bland-Hawthorn, J. and Maloney, P. R.: 1999, The Escape of Ionizing Photons from the Galaxy, *ApJ* **510**, L33

Blumenthal, G. R., Faber, S. M., Primack, J. R., and Rees, M. J.: 1984, Formation of galaxies and large-scale structure with cold dark matter, *Nature* **311**, 517

Bolton, J. S. and Becker, G. D.: 2009, Resolving the high redshift Lyα forest in smoothed particle hydrodynamics simulations, *MNRAS* **398**, L26

Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., and Sargent, W. L. W.: 2010, A first direct measurement of the intergalactic medium temperature around a quasar at $z = 6$, *MNRAS* **406**, 612

Bolton, J. S. and Haehnelt, M. G.: 2007a, A closer look at using quasar near-zones as a probe of neutral hydrogen in the intergalactic medium, *MNRAS* **381**, L35

Bolton, J. S. and Haehnelt, M. G.: 2007b, The nature and evolution of the highly ionized near-zones in the absorption spectra of $z \simeq 6$ quasars, *MNRAS* **374**, 493

Bolton, J. S. and Haehnelt, M. G.: 2007c, The observed ionization rate of the intergalactic medium and the ionizing emissivity at $z \geq 5$: evidence for a photon-starved and extended epoch of reionization, *MNRAS* **382**, 325

Bolton, J. S., Haehnelt, M. G., Viel, M., and Carswell, R. F.: 2006, Spatial fluctuations in the spectral shape of the ultraviolet background at $2 < z < 3$ and the reionization of helium, *MNRAS* **366**, 1378

Bolton, J. S., Haehnelt, M. G., Viel, M., and Springel, V.: 2005, The Lyman α forest opacity and the metagalactic hydrogen ionization rate at $z \sim 2-4$, *MNRAS* **357**, 1178

Bolton, J. S., Haehnelt, M. G., Warren, S. J., Hewett, P. C., Mortlock, D. J., Venemans, B. P., McMahon, R. G., and Simpson, C.: 2011, How neutral is the intergalactic medium surrounding the redshift $z = 7.085$ quasar ULAS J1120+0641?, *MNRAS* **416**, L70

Bolton, J. S., Oh, S. P., and Furlanetto, S. R.: 2009, Photoheating and the fate of hard photons during the reionization of HeII by quasars, *MNRAS* **395**, 736

Bolton, J. S., Viel, M., Kim, T.-S., Haehnelt, M. G., and Carswell, R. F.: 2008, Possible evidence for an

inverted temperature-density relation in the intergalactic medium from the flux distribution of the Lyα forest, *MNRAS* **386**, 1131

Bond, J. R., Kofman, L., and Pogosyan, D.: 1996, How filaments of galaxies are woven into the cosmic web, *Nature* **380**, 603

Bond, J. R. and Szalay, A. S.: 1983, The collisionless damping of density fluctuations in an expanding universe, *ApJ* **274**, 443

Bondi, H. and Gold, T.: 1948, The Steady-State Theory of the Expanding Universe, *MNRAS* **108**, 252

Bonoli, S., Marulli, F., Springel, V., White, S. D. M., Branchini, E., and Moscardini, L.: 2009, Modelling the cosmological co-evolution of supermassive black holes and galaxies - II. The clustering of quasars and their dark environment, *MNRAS* **396**, 423

Boutsia, K., Grazian, A., Giallongo, E., Fontana, A., Pentericci, L., Castellano, M., Zamorani, G., Mignoli, M., Vanzella, E., Fiore, F., Lilly, S. J., Gallozzi, S., Testa, V., Paris, D., and Santini, P.: 2011, A Low Escape Fraction of Ionizing Photons of $L > L_*$ Lyman Break Galaxies at $z = 3.3$, *ApJ* **736**, 41

Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., and Franx, M.: 2006, Galaxies at $z \sim 6$: The UV Luminosity Function and Luminosity Density from 506 HUDF, HUDF Parallel ACS Field, and GOODS i-Dropouts, *ApJ* **653**, 53

Bouwens, R. J., Illingworth, G. D., Franx, M., Chary, R., Meurer, G. R., Conselice, C. J., Ford, H., Giavalisco, M., and van Dokkum, P.: 2009, UV Continuum Slope and Dust Obscuration from $z \sim 6$ to $z \sim 2$: The Star Formation Rate Density at High Redshift, *ApJ* **705**, 936

Bouwens, R. J., Illingworth, G. D., Franx, M., and Ford, H.: 2008, $z \sim 7-10$ Galaxies in the HUDF and GOODS Fields: UV Luminosity Functions, *ApJ* **686**, 230

Bouwens, R. J., Illingworth, G. D., Labbe, I., Oesch, P. A., Trenti, M., Carollo, C. M., van Dokkum, P. G., Franx, M., Stiavelli, M., González, V., Magee, D., and Bradley, L.: 2011a, A candidate redshift $z \approx 10$ galaxy and rapid changes in that population at an age of 500Myr, *Nature* **469**, 504

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Labbé, I., Trenti, M., van Dokkum, P., Franx, M., Stiavelli, M., Carollo, C. M., Magee, D., and Gonzalez, V.: 2011b, Ultraviolet Luminosity Functions from 132 $z \sim 7$ and $z \sim 8$ Lyman-break Galaxies in the Ultra-deep HUDF09 and Wide-area Early Release Science WFC3/IR Observations, *ApJ* **737**, 90

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Trenti, M., Labbe, I., Franx, M., Stiavelli, M., Carollo, C. M., van Dokkum, P., and Magee, D.: 2011c, Lower-Luminosity Galaxies could reionize the Universe: Very Steep Faint-End Slopes to the UV Luminosity Functions at $z \geq 5-8$ from the HUDF09 WFC3/IR Observations, *ArXiv e-prints*

Bouwens, R. J., Thompson, R. I., Illingworth, G. D., Franx, M., van Dokkum, P. G., Fan, X., Dickinson, M. E., Eisenstein, D. J., and Rieke, M. J.: 2004, Galaxies at $z \sim 7-8$: z_{850}-Dropouts in the Hubble Ultra Deep Field, *ApJ* **616**, L79

Bowman, J. D. and Rogers, A. E. E.: 2010, A lower limit of $\Delta z > 0.06$ for the duration of the reionization epoch, *Nature* **468**, 796

Broderick, A. E., Chang, P., and Pfrommer, C.: 2011, The Cosmological Impact of Luminous TeV Blazars I: Implications of Plasma Instabilities for the Intergalactic Magnetic Field and Extragalactic Gamma-Ray Background, *ArXiv e-prints*

Bromm, V., Coppi, P. S., and Larson, R. B.: 2002, The Formation of the First Stars. I. The Primordial Star-forming Cloud, *ApJ* **564**, 23

Bromm, V., Kudritzki, R. P., and Loeb, A.: 2001, Generic Spectrum and Ionization Efficiency of a Heavy Initial Mass Function for the First Stars, *ApJ* **552**, 464

Bromm, V. and Larson, R. B.: 2004, The First Stars, *ARA&A* **42**, 79

Bromm, V. and Loeb, A.: 2003, The formation of the first low-mass stars from gas with low carbon and oxygen abundances, *Nature* **425**, 812

Bromm, V. and Loeb, A.: 2007, in S. Immler, K. Weiler, and R. McCray (eds.), *Supernova 1987A: 20 Years After: Supernovae and Gamma-Ray Bursters*, Vol. 937 of *American Institute of Physics Conference Series*, pp 532–541

Bromm, V. and Yoshida, N.: 2011, The First Galaxies, *ARA&A* **49**, 373

Bromm, V., Yoshida, N., Hernquist, L., and McKee, C. F.: 2009, The formation of the first stars and galaxies, *Nature* **459**, 49

Bruns, Jr., L. R., Wyithe, J. S. B., Bland-Hawthorn, J., and Dijkstra, M.: 2011, Clustering of Ly-alpha emitters around luminous quasars at $z = 2-3$: an alternative probe of reionization on galaxy formation, *ArXiv e-prints*

Bunker, A., Stanway, E., Ellis, R., McMahon, R., Eyles, L., and Lacy, M.: 2006, Star forming galaxies at $z \approx 6$ and reionization, *New Astronomy Review* **50**, 94

Bunker, A. J., Stanway, E. R., Ellis, R. S., and McMahon, R. G.: 2004, The star formation rate of the Universe at $z \sim 6$ from the Hubble Ultra-Deep Field, *MNRAS* **355**, 374

Bunker, A. J., Wilkins, S., Ellis, R. S., Stark, D. P., Lorenzoni, S., Chiu, K., Lacy, M., Jarvis, M. J., and Hickey, S.: 2010, The contribution of high-redshift galaxies to cosmic reionization: new results from deep WFC3 imaging of the Hubble Ultra Deep Field, *MNRAS* **409**, 855

Burbidge, G. R.: 1956, On Synchrotron Radiation from Messier 87., *ApJ* **124**, 416

Calverley, A. P., Becker, G. D., Haehnelt, M. G., and Bolton, J. S.: 2011, Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect, *MNRAS* **412**, 2543

Carilli, C. L.: 2006, in *IAU Joint Discussion*, Vol. 12 of *IAU Joint Discussion*

Carilli, C. L., Furlanetto, S., Briggs, F., Jarvis, M., Rawlings, S., and Falcke, H.: 2004, Probing the dark ages with the Square Kilometer Array, *New Astronomy Review* **48**, 1029

Carilli, C. L., Wang, R., Fan, X., Walter, F., Kurk, J., Riechers, D., Wagg, J., Hennawi, J., Jiang, L., Menten, K. M., Bertoldi, F., Strauss, M. A., and Cox, P.: 2010, Ionization Near Zones Associated with Quasars at $z \sim 6$, *ApJ* **714**, 834

Carswell, R. F., Webb, J. K., Baldwin, J. A., and Atwood, B.: 1987, High-redshift QSO absorbing clouds and the background ionizing source, *ApJ* **319**, 709

Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A., and Tytler, D.: 1982, Observations of the spectra of Q0122-380 and Q1101-264, *MNRAS* **198**, 91

Cen, R.: 2003a, The Implications of Wilkinson Microwave Anisotropy Probe Observations for Population III Star Formation Processes, *ApJ* **591**, L5

Cen, R.: 2003b, The Universe Was Reionized Twice, *ApJ* **591**, 12

Cen, R.: 2010, The State of Star Formation and the Intergalactic Medium at $z \sim 6$, *ApJ* **725**, 115

Cen, R. and McDonald, P.: 2002, Evolution of the Ionizing Radiation Background and Star Formation in the Aftermath of Cosmological Reionization, *ApJ* **570**, 457

Cen, R., McDonald, P., Trac, H., and Loeb, A.: 2009, Probing the Epoch of Reionization with the Lyα Forest at $z \sim 4-5$, *ApJ* **706**, L164

Cen, R., Miralda-Escudé, J., Ostriker, J. P., and Rauch, M.: 1994, Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest, *ApJ* **437**, L9

Chang, P., Broderick, A. E., and Pfrommer, C.: 2011, The Cosmological Impact of Luminous TeV Blazars II: Rewriting the Thermal History of the Intergalactic Medium, *ArXiv e-prints*

Chen, H.-W., Perley, D. A., Pollack, L. K., Prochaska, J. X., Bloom, J. S., Dessauges-Zavadsky, M., Pettini, M., Lopez, S., Dall'aglio, A., and Becker, G. D.: 2009, High-Redshift Starbursting Dwarf Galaxies Revealed by γ-Ray Burst Afterglows, *ApJ* **691**, 152

Choudhury, T. R.: 2009, Analytical Models of the Intergalactic Medium and Reionization, *Current Science* **97**, 841

Choudhury, T. R. and Ferrara, A.: 2005, Experimental constraints on self-consistent reionization models, *MNRAS* **361**, 577

Choudhury, T. R., Ferrara, A., and Gallerani, S.: 2008, On the minimum mass of reionization sources, *MNRAS* **385**, L58

Choudhury, T. R., Haehnelt, M. G., and Regan, J.: 2009, Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Lyα forest data, *MNRAS* **394**, 960

Ciardi, B. and Ferrara, A.: 2005, The First Cosmic Structures and Their Effects, *Space Science Reviews* **116**, 625

Clark, P. C., Glover, S. C. O., Klessen, R. S., and Bromm, V.: 2011, Gravitational Fragmentation in Turbulent Primordial Gas and the Initial Mass Function of Population III Stars, *ApJ* **727**, 110

Clément, B., Cuby, J. ., Courbin, F., Fontana, A., Freudling, W., Fynbo, J., Gallego, J., Hibon, P., Kneib, J. ., Le Fèvre, O., Lidman, C., McMahon, R., Milvang-Jensen, B., Moller, P., Moorwood, A., Nilsson, K. K., Pentericci, L., Venemans, B., Villar, V., and Willis, J.: 2011, Evolution of the Ly-alpha luminosity function from $z = 6.5$ to $z = 7.7$: evidence for the epoch of reionization?, *ArXiv e-prints*

Cole, S., Lacey, C. G., Baugh, C. M., and Frenk, C. S.: 2000, Hierarchical galaxy formation, *MNRAS* **319**, 168

Colgate, S. A.: 1968, Prompt gamma rays and X-rays from supernovae., *Canadian Journal of Physics* **46**, 476

Cooke, A. J., Espey, B., and Carswell, R. F.: 1997, Evolution of the ionizing background at high redshifts, *MNRAS* **284**, 552

Couchman, H. M. P. and Rees, M. J.: 1986, Pregalactic evolution in cosmologies with cold dark matter, *MNRAS* **221**, 53

Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., and Savaglio, S.: 1995, The space distribution of the Lyman alpha clouds in the line of sight to the $z = 3.66$ QSO 0055-269, *MNRAS* **273**, 1016

Croft, R. A. C.: 2004, Ionizing Radiation Fluctuations and Large-Scale Structure in the Lyα Forest, *ApJ* **610**, 642

Cucchiara, A., Levan, A. J., Fox, D. B., Tanvir, N. R., Ukwatta, T. N., Berger, E., Krühler, T., Küpcü

Yoldaş, A., Wu, X. F., Toma, K., Greiner, J., Olivares, F. E., Rowlinson, A., Amati, L., Sakamoto, T., Roth, K., Stephens, A., Fritz, A., Fynbo, J. P. U., Hjorth, J., Malesani, D., Jakobsson, P., Wiersema, K., O'Brien, P. T., Soderberg, A. M., Foley, R. J., Fruchter, A. S., Rhoads, J., Rutledge, R. E., Schmidt, B. P., Dopita, M. A., Podsiadlowski, P., Willingale, R., Wolf, C., Kulkarni, S. R., and D'Avanzo, P.: 2011, A Photometric Redshift of $z \sim 9.4$ for GRB 090429B, *ApJ* **736**, 7

da Ângela, J., Shanks, T., Croom, S. M., Weilbacher, P., Brunner, R. J., Couch, W. J., Miller, L., Myers, A. D., Nichol, R. C., Pimbblet, K. A., de Propris, R., Richards, G. T., Ross, N. P., Schneider, D. P., and Wake, D.: 2008, The 2dF-SDSS LRG and QSO survey: QSO clustering and the $L-z$ degeneracy, *MNRAS* **383**, 565

Dall'Aglio, A. and Gnedin, N. Y.: 2010, Analysis of Methods for Detecting the Proximity Effect in Quasar Spectra, *ApJ* **722**, 699

Dall'Aglio, A., Wisotzki, L., and Worseck, G.: 2008, An unbiased measurement of the UV background and its evolution via the proximity effect in quasar spectra, *A&A* **491**, 465

Dall'Aglio, A., Wisotzki, L., and Worseck, G.: 2009, The UV background photoionization rate at $2.3 < z < 4.6$ as measured from the Sloan Digital Sky Survey, *ArXiv e-prints*

de Souza, R. S., Yoshida, N., and Ioka, K.: 2011, Populations III.1 and III.2 gamma-ray bursts: constraints on the event rate for future radio and X-ray surveys, *A&A* **533**, A32+

Diemand, J., Moore, B., and Stadel, J.: 2005, Earth-mass dark-matter haloes as the first structures in the early Universe, *Nature* **433**, 389

Dijkstra, M., Lidz, A., and Hui, L.: 2004, Beyond Lyα: Constraints and Consistency Tests from the Lyβ Forest, *ApJ* **605**, 7

Dijkstra, M., Mesinger, A., and Wyithe, J. S. B.: 2011, The detectability of Lyα emission from galaxies during the epoch of reionization, *MNRAS* **414**, 2139

Dijkstra, M., Wyithe, J. S. B., and Haiman, Z.: 2007, Luminosity functions of Lyα emitting galaxies and cosmic reionization of hydrogen, *MNRAS* **379**, 253

D'Odorico, V., Bruscoli, M., Saitta, F., Fontanot, F., Viel, M., Cristiani, S., and Monaco, P.: 2008, The quasar proximity effect at redshift $\langle z \rangle \simeq 2.6$ with the From Lines to Overdensities approach, *MNRAS* **389**, 1727

Donahue, M. and Shull, J. M.: 1987, Can quasars ionize the intergalactic medium?, *ApJ* **323**, L13

Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E., and Archer, S.: 1959, A survey of radio sources at a frequency of 159 Mc/s., *MmRAS* **68**, 37

Einstein, A.: 1915, Die Feldgleichungen der Gravitation, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* pp 844–847

Einstein, A.: 1917, Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* pp 142–152

Eisenstein, D. J. and Hu, W.: 1999, Power Spectra for Cold Dark Matter and Its Variants, *ApJ* **511**, 5

Espey, B. R.: 1993, The influence of QSO emission-line velocity shifts on estimates of the intergalactic background intensity, *ApJ* **411**, L59

Fan, X., Carilli, C. L., and Keating, B.: 2006a, Observational Constraints on Cosmic Reionization, *ARA&A* **44**, 415

Fan, X., Hennawi, J. F., Richards, G. T., Strauss, M. A., Schneider, D. P., Donley, J. L., Young, J. E., Annis, J., Lin, H., Lampeitl, H., Lupton, R. H., Gunn, J. E., Knapp, G. R., Brandt, W. N., Anderson, S., Bahcall, N. A., Brinkmann, J., Brunner, R. J., Fukugita, M., Szalay, A. S., Szokoly, G. P., and York, D. G.: 2004, A Survey of $z > 5.7$ Quasars in the Sloan Digital Sky Survey. III. Discovery of Five Additional Quasars, *AJ* **128**, 515

Fan, X., Narayanan, V. K., Lupton, R. H., Strauss, M. A., Knapp, G. R., Becker, R. H., White, R. L., Pentericci, L., Leggett, S. K., Haiman, Z., Gunn, J. E., Ivezić, Ž., Schneider, D. P., Anderson, S. F., Brinkmann, J., Bahcall, N. A., Connolly, A. J., Csabai, I., Doi, M., Fukugita, M., Geballe, T., Grebel, E. K., Harbeck, D., Hennessy, G., Lamb, D. Q., Miknaitis, G., Munn, J. A., Nichol, R., Okamura, S., Pier, J. R., Prada, F., Richards, G. T., Szalay, A., and York, D. G.: 2001, 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 $z \sim 6$, *AJ* **122**, 2833

Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., and Fukugita, M.: 2006b, Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars, *AJ* **132**, 117

Fan, X., Strauss, M. A., Richards, G. T., Hennawi, J. F., Becker, R. H., White, R. L., Diamond-Stanic, A. M., Donley, J. L., Jiang, L., Kim, J. S., Vestergaard, M., Young, J. E., Gunn, J. E., Lupton, R. H., Knapp, G. R., Schneider, D. P., Brandt, W. N., Bahcall, N. A., Barentine, J. C., Brinkmann, J., Brewington, H. J., Fukugita, M., Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., Snedden, S. A., and Voges, W.: 2006c, A Survey of $z > 5.7$ Quasars in the Sloan Digital Sky Survey. IV. Discovery of Seven Additional Quasars, *AJ* **131**, 1203

Fan, X., Strauss, M. A., Schneider, D. P., Becker, R. H., White, R. L., Haiman, Z., Gregg, M., Pentericci, L., Grebel, E. K., Narayanan, V. K., Loh, Y., Richards, G. T., Gunn, J. E., Lupton, R. H., Knapp, G. R., Ivezić, Ž., Brandt, W. N., Collinge, M., Hao, L., Harbeck, D., Prada, F., Schaye, J., Strateva, I., Zakamska, N., Anderson, S., Brinkmann, J., Bahcall, N. A., Lamb, D. Q., Okamura, S., Szalay, A., and York, D. G.: 2003, A Survey of $z > 5.7$ Quasars in the Sloan Digital Sky Survey. II. Discovery of Three Additional Quasars at $z > 6$, *AJ* **125**, 1649

Fardal, M. A., Giroux, M. L., and Shull, J. M.: 1998, The High-Redshift He II Gunn-Peterson Effect: Implications and Future Prospects, *AJ* **115**, 2206

Faucher-Giguère, C., Lidz, A., Hernquist, L., and Zaldarriaga, M.: 2008a, Evolution of the Intergalactic Opacity: Implications for the Ionizing Background, Cosmic Star Formation, and Quasar Activity, *ApJ* **688**, 85

Faucher-Giguère, C., Lidz, A., Zaldarriaga, M., and Hernquist, L.: 2008b, The Line-of-Sight Proximity Effect and the Mass of Quasar Host Halos, *ApJ* **673**, 39

Feng, L., Bi, H., Liu, J., and Fang, L.-Z.: 2008, Lyα leaks and reionization, *MNRAS* **383**, 1459

Fernandez, E. R. and Shull, J. M.: 2011, The Effect of Galactic Properties on the Escape Fraction of Ionizing Photons, *ApJ* **731**, 20

Field, G. B.: 1959, An Attempt to Observe Neutral Hydrogen Between the Galaxies., *ApJ* **129**, 525

Fontanot, F., Monaco, P., Cristiani, S., and Tozzi, P.: 2006, The effect of stellar feedback and quasar winds on the active galactic nucleus population, *MNRAS* **373**, 1173

Frenk, C. S., White, S. D. M., Davis, M., and Efstathiou, G.: 1988, The formation of dark halos in a universe dominated by cold dark matter, *ApJ* **327**, 507

Friedmann, A.: 1922, Über die Krümmung des Raumes, *Zeitschrift fur Physik* **10**, 377

Friedrich, M. M., Mellema, G., Alvarez, M. A., Shapiro, P. R., and Iliev, I. T.: 2011, Topology and sizes of H II regions during cosmic reionization, *MNRAS* **413**, 1353

Furlanetto, S. R.: 2006, The global 21-centimeter background from high redshifts, *MNRAS* **371**, 867

Furlanetto, S. R., Haiman, Z., and Oh, S. P.: 2008, Fossil Ionized Bubbles around Dead Quasars during Reionization, *ApJ* **686**, 25

Furlanetto, S. R. and Loeb, A.: 2005, Is Double Reionization Physically Plausible?, *ApJ* **634**, 1

Furlanetto, S. R. and Mesinger, A.: 2009, The ionizing background at the end of reionization, *MNRAS* **394**, 1667

Furlanetto, S. R. and Oh, S. P.: 2005, Taxing the rich: recombinations and bubble growth during reionization, *MNRAS* **363**, 1031

Furlanetto, S. R. and Oh, S. P.: 2008a, The History and Morphology of Helium Reionization, *ApJ* **681**, 1

Furlanetto, S. R. and Oh, S. P.: 2008b, Inhomogeneous Helium Reionization and the Equation of State of the Intergalactic Medium, *ApJ* **682**, 14

Furlanetto, S. R. and Oh, S. P.: 2009, The Temperature-Density Relation of the Intergalactic Medium after Hydrogen Reionization, *ApJ* **701**, 94

Furlanetto, S. R., Oh, S. P., and Briggs, F. H.: 2006a, Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe, *Phys. Rep.* **433**, 181

Furlanetto, S. R., Sokasian, A., and Hernquist, L.: 2004a, Observing the reionization epoch through 21-centimetre radiation, *MNRAS* **347**, 187

Furlanetto, S. R., Zaldarriaga, M., and Hernquist, L.: 2004b, The Growth of H II Regions During Reionization, *ApJ* **613**, 1

Furlanetto, S. R., Zaldarriaga, M., and Hernquist, L.: 2006b, The effects of reionization on Lyα galaxy surveys, *MNRAS* **365**, 1012

Gallerani, S., Ferrara, A., Fan, X., and Choudhury, T. R.: 2008a, Glimpsing through the high-redshift neutral hydrogen fog, *MNRAS* **386**, 359

Gallerani, S., Salvaterra, R., Ferrara, A., and Choudhury, T. R.: 2008b, Testing reionization with gamma-ray burst absorption spectra, *MNRAS* **388**, L84

Gamow, G.: 1948, The Origin of Elements and the Separation of Galaxies, *Physical Review* **74**, 505

Geil, P. M. and Wyithe, J. S. B.: 2008, The impact of a percolating IGM on redshifted 21-cm observations of quasar HII regions, *MNRAS* **386**, 1683

Ghosh, A., Bharadwaj, S., Ali, S. S., and Chengalur, J. N.: 2011, GMRT observation towards detecting the post-reionization 21-cm signal, *MNRAS* **411**, 2426

Gleser, L., Nusser, A., Benson, A. J., Ohno, H., and Sugiyama, N.: 2005, Patchy He II reionization and the physical state of the intergalactic medium, *MNRAS* **361**, 1399

Gnedin, N. Y.: 2000, Cosmological Reionization by Stellar Sources, *ApJ* **535**, 530

Gnedin, N. Y.: 2004, Reionization, Sloan, and WMAP: Is the Picture Consistent?, *ApJ* **610**, 9

Gnedin, N. Y. and Fan, X.: 2006, Cosmic Reionization Redux, *ApJ* **648**, 1

Gnedin, N. Y., Kravtsov, A. V., and Chen, H.-W.: 2008, Escape of Ionizing Radiation from High-Redshift Galaxies, *ApJ* **672**, 765

Gnedin, N. Y. and Ostriker, J. P.: 1997, Reionization of the Universe and the Early Production of Metals, *ApJ* **486**, 581

Goto, T., Utsumi, Y., Hattori, T., Miyazaki, S., and Yamauchi, C.: 2011, A Gunn-Peterson test with a QSO at $z = 6.4$, *MNRAS* **415**, L1

Granato, G. L., De Zotti, G., Silva, L., Bressan, A., and Danese, L.: 2004, A Physical Model for the Coevolution of QSOs and Their Spheroidal Hosts, *ApJ* **600**, 580

Greenstein, J. L. and Schmidt, M.: 1964, The Quasi-Stellar Radio Sources 3C 48 and 3C 273., *ApJ* **140**, 1

Greif, T. H., Johnson, J. L., Klessen, R. S., and Bromm, V.: 2008, The first galaxies: assembly, cooling and the onset of turbulence, *MNRAS* **387**, 1021

Greif, T. H., Johnson, J. L., Klessen, R. S., and Bromm, V.: 2009, The observational signature of the first HII regions, *MNRAS* **399**, 639

Greiner, J., Krühler, T., Fynbo, J. P. U., Rossi, A., Schwarz, R., Klose, S., Savaglio, S., Tanvir, N. R., McBreen, S., Totani, T., Zhang, B. B., Wu, X. F., Watson, D., Barthelmy, S. D., Beardmore, A. P., Ferrero, P., Gehrels, N., Kann, D. A., Kawai, N., Yoldaş, A. K., Mészáros, P., Milvang-Jensen, B., Oates, S. R., Pierini, D., Schady, P., Toma, K., Vreeswijk, P. M., Yoldaş, A., Zhang, B., Afonso, P., Aoki, K., Burrows, D. N., Clemens, C., Filgas, R., Haiman, Z., Hartmann, D. H., Hasinger, G., Hjorth, J., Jehin, E., Levan, A. J., Liang, E. W., Malesani, D., Pyo, T.-S., Schulze, S., Szokoly, G., Terada, K., and Wiersema, K.: 2009, GRB 080913 at Redshift 6.7, *ApJ* **693**, 1610

Griffiths, L. M., Barbosa, D., and Liddle, A. R.: 1999, Cosmic microwave background constraints on the epoch of reionization, *MNRAS* **308**, 854

Guimarães, R., Petitjean, P., Rollinde, E., de Carvalho, R. R., Djorgovski, S. G., Srianand, R., Aghaee, A., and Castro, S.: 2007, Evidence for overdensity around $z_{\rm em} > 4$ quasars from the proximity effect, *MNRAS* **377**, 657

Gunn, J. E. and Peterson, B. A.: 1965, On the Density of Neutral Hydrogen in Intergalactic Space., *ApJ* **142**, 1633

Guth, A. H.: 1981, Inflationary universe: A possible solution to the horizon and flatness problems, *Phys. Rev. D* **23**, 347

Haardt, F. and Madau, P.: 1996, Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background, *ApJ* **461**, 20

Haardt, F. and Madau, P.: 2001, in N. D. M. and T. J. T. V (eds.), *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*

Haardt, F. and Madau, P.: 2011, Radiative transfer in a clumpy universe: IV. New synthesis models of the cosmic UV/X-ray background, *ArXiv e-prints*

Haehnelt, M. G. and Steinmetz, M.: 1998, Probing the thermal history of the intergalactic medium with Lyα absorption lines, *MNRAS* **298**, L21

Haiman, Z. and Bryan, G. L.: 2006, Was Star Formation Suppressed in High-Redshift Minihalos?, *ApJ* **650**, 7

Haiman, Z. and Holder, G. P.: 2003, The Reionization History at High Redshifts. I. Physical Models and New Constraints from Cosmic Microwave Background Polarization, *ApJ* **595**, 1

Haiman, Z., Thoul, A. A., and Loeb, A.: 1996, Cosmological Formation of Low-Mass Objects, *ApJ* **464**, 523

Hansen, S. H. and Haiman, Z.: 2004, Do We Need Stars to Reionize the Universe at High Redshifts? Early Reionization by Decaying Heavy Sterile Neutrinos, *ApJ* **600**, 26

Hernquist, L., Katz, N., Weinberg, D. H., and Miralda-Escudé, J.: 1996, The Lyman-Alpha Forest in the Cold Dark Matter Model, *ApJ* **457**, L51+

Holmberg, E.: 1941, On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure., *ApJ* **94**, 385

Hopkins, P. F., Richards, G. T., and Hernquist, L.: 2007, An Observational Determination of the Bolometric Quasar Luminosity Function, *ApJ* **654**, 731

Horne, K.: 1986, An optimal extraction algorithm for CCD spectroscopy, *PASP* **98**, 609

Hoyle, F.: 1948, A New Model for the Expanding Universe, *MNRAS* **108**, 372

Hubble, E. P.: 1925, NGC 6822, a remote stellar system., *ApJ* **62**, 409

Hubble, E. P.: 1929, A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae, *Proceedings of the National Academy of Science* **15**, 168

Hui, L. and Gnedin, N. Y.: 1997, Equation of state of the photoionized intergalactic medium, *MNRAS* **292**, 27

Hui, L. and Haiman, Z.: 2003, The Thermal Memory of Reionization History, *ApJ* **596**, 9

Jakobsson, P., Levan, A., Fynbo, J. P. U., Priddey, R., Hjorth, J., Tanvir, N., Watson, D., Jensen, B. L., Sollerman, J., Natarajan, P., Gorosabel, J., Castro Cerón, J. M., Pedersen, K., Pursimo, T., Árnadóttir, A. S., Castro-Tirado, A. J., Davis, C. J., Deeg, H. J., Fiuza, D. A., Mykolaitis, S., and Sousa, S. G.: 2006, A mean redshift of 2.8 for Swift gamma-ray bursts, *A&A* **447**, 897

Jena, T., Norman, M. L., Tytler, D., Kirkman, D., Suzuki, N., Chapman, A., Melis, C., Paschos, P., O'Shea, B., So, G., Lubin, D., Lin, W., Reimers, D., Janknecht, E., and Fechner, C.: 2005, A concordance model of the Lyman α forest at $z = 1.95$, *MNRAS* **361**, 70

Johnson, J. L.: 2011, Formation of the First Galaxies: Theory and Simulations, *ArXiv e-prints*

Johnson, J. L., Greif, T. H., and Bromm, V.: 2008, Occurrence of metal-free galaxies in the early Universe, *MNRAS* **388**, 26

Johnson, J. L. and Khochfar, S.: 2011, The Contribution of Supernovae to Cosmic Reionization, *ArXiv e-prints*

Karlsson, T., Bromm, V., and Bland-Hawthorn, J.: 2011, Pre-galactic metal enrichment - The chemical signatures of the first stars, *ArXiv e-prints*

Kashikawa, N., Shimasaku, K., Malkan, M. A., Doi, M., Matsuda, Y., Ouchi, M., Taniguchi, Y., Ly, C., Nagao, T., Iye, M., Motohara, K., Murayama, T., Murozono, K., Nariai, K., Ohta, K., Okamura, S., Sasaki, T., Shioya, Y., and Umemura, M.: 2006, The End of the Reionization Epoch Probed by Lyα Emitters at $z = 6.5$ in the Subaru Deep Field, *ApJ* **648**, 7

Kashikawa, N., Shimasaku, K., Matsuda, Y., Egami, E., Jiang, L., Nagao, T., Ouchi, M., Malkan, M. A., Hattori, T., Ota, K., Taniguchi, Y., Okamura, S., Ly, C., Iye, M., Furusawa, H., Shioya, Y., Shibuya, T.,

Ishizaki, Y., and Toshikawa, J.: 2011, Completing the Census of Lyα Emitters at the Reionization Epoch, *ApJ* **734**, 119

Kawai, N., Kosugi, G., Aoki, K., Yamada, T., Totani, T., Ohta, K., Iye, M., Hattori, T., Aoki, W., Furusawa, H., Hurley, K., Kawabata, K. S., Kobayashi, N., Komiyama, Y., Mizumoto, Y., Nomoto, K., Noumaru, J., Ogasawara, R., Sato, R., Sekiguchi, K., Shirasaki, Y., Suzuki, M., Takata, T., Tamagawa, T., Terada, H., Watanabe, J., Yatsu, Y., and Yoshida, A.: 2006, An optical spectrum of the afterglow of a γ-ray burst at a redshift of $z = 6.295$, *Nature* **440**, 184

Kelson, D. D.: 2003, Optimal Techniques in Two-dimensional Spectroscopy: Background Subtraction for the 21st Century, *PASP* **115**, 688

Kim, Y.-R. and Croft, R. A. C.: 2008, Constraining quasar host halo masses with the strength of nearby Lyα forest absorption, *MNRAS* **387**, 377

Kirkman, D. and Tytler, D.: 2008, The transverse proximity effect in the $z \sim 2$ Lyα forest suggests quasi-stellar object episodic lifetimes of ~ 1 Myr, *MNRAS* **391**, 1457

Kirkman, D., Tytler, D., Suzuki, N., Melis, C., Hollywood, S., James, K., So, G., Lubin, D., Jena, T., Norman, M. L., and Paschos, P.: 2005, The HI opacity of the intergalactic medium at redshifts $1.6 < z < 3.2$, *MNRAS* **360**, 1373

Klebesadel, R. W., Strong, I. B., and Olson, R. A.: 1973, Observations of Gamma-Ray Bursts of Cosmic Origin, *ApJ* **182**, L85+

Kogut, A., Spergel, D. N., Barnes, C., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Wollack, E., and Wright, E. L.: 2003, First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature-Polarization Correlation, *ApJS* **148**, 161

Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Limon, M., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Meyer, S. S., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L.: 2009, Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation, *ApJS* **180**, 330

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L.: 2011, Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation, *ApJS* **192**, 18

Kriss, G. A., Shull, J. M., Oegerle, W., Zheng, W., Davidsen, A. F., Songaila, A., Tumlinson, J., Cowie, L. L., Deharveng, J.-M., Friedman, S. D., Giroux, M. L., Green, R. F., Hutchings, J. B., Jenkins, E. B., Kruk, J. W., Moos, H. W., Morton, D. C., Sembach, K. R., and Tripp, T. M.: 2001, Resolving the Structure of Ionized Helium in the Intergalactic Medium with the Far Ultraviolet Spectroscopic Explorer, *Science* **293**, 1112

Krug, H., Veilleux, S., Tilvi, V., Malhotra, S., Rhoads, J., Hibon, P., Swaters, R., Probst, R., Dey, A., Dickinson, M., and Jannuzi, B.: 2011, Searching for High-Redshift Lyman Alpha Emitters in the COSMOS Field with NEWFIRM, *ArXiv e-prints*

Kulkarni, S. R., Djorgovski, S. G., Odewahn, S. C., Bloom, J. S., Gal, R. R., Koresko, C. D., Harrison, F. A., Lubin, L. M., Armus, L., Sari, R., Illingworth, G. D., Kelson, D. D., Magee, D. K., van Dokkum, P. G.,

Frail, D. A., Mulchaey, J. S., Malkan, M. A., McClean, I. S., Teplitz, H. I., Koerner, D., Kirkpatrick, D., Kobayashi, N., Yadigaroglu, I.-A., Halpern, J., Piran, T., Goodrich, R. W., Chaffee, F. H., Feroci, M., and Costa, E.: 1999, The afterglow, redshift and extreme energetics of the γ-ray burst of 23 January 1999, *Nature* **398**, 389

Kulkarni, S. R., Djorgovski, S. G., Ramaprakash, A. N., Goodrich, R., Bloom, J. S., Adelberger, K. L., Kundic, T., Lubin, L., Frail, D. A., Frontera, F., Feroci, M., Nicastro, L., Barth, A. J., Davis, M., Filippenko, A. V., and Newman, J.: 1998, Identification of a host galaxy at redshift $z = 3.42$ for the γ-ray burst of 14 December 1997, *Nature* **393**, 35

Kulkarni, V. P. and Fall, S. M.: 1993, The proximity effect and the mean intensity of ionizing radiation at low redshifts, *ApJ* **413**, L63

Lacey, C. G., Baugh, C. M., Frenk, C. S., and Benson, A. J.: 2011, The evolution of Lyman-break galaxies in the cold dark matter model, *MNRAS* **412**, 1828

Lee, K.-G.: 2011, Systematic Continuum Errors in the Lyman-Alpha Forest and The Measured Temperature-Density Relation, *ArXiv e-prints*

Lehnert, M. D., Nesvadba, N. P. H., Cuby, J.-G., Swinbank, A. M., Morris, S., Clément, B., Evans, C. J., Bremer, M. N., and Basa, S.: 2010, Spectroscopic confirmation of a galaxy at redshift $z = 8.6$, *Nature* **467**, 940

Leitherer, C., Schaerer, D., Goldader, J. D., González Delgado, R. M., Robert, C., Kune, D. F., de Mello, D. F., Devost, D., and Heckman, T. M.: 1999, Starburst99: Synthesis Models for Galaxies with Active Star Formation, *ApJS* **123**, 3

Lemaître, G.: 1927, Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques, *Annales de la Societe Scietifique de Bruxelles* **47**, 49

Lemaître, G.: 1931, Contributions to a British Association Discussion on the Evolution of the Universe., *Nature* **128**, 704

Lidz, A., Faucher-Giguère, C.-A., Dall'Aglio, A., McQuinn, M., Fechner, C., Zaldarriaga, M., Hernquist, L., and Dutta, S.: 2010, A Measurement of Small-scale Structure in the $2.2 \leq z \leq 4.2$ Lyα Forest, *ApJ* **718**, 199

Lidz, A., Heitmann, K., Hui, L., Habib, S., Rauch, M., and Sargent, W. L. W.: 2006a, Tightening Constraints from the Lyα Forest with the Flux Probability Distribution Function, *ApJ* **638**, 27

Lidz, A., McQuinn, M., Zaldarriaga, M., Hernquist, L., and Dutta, S.: 2007, Quasar Proximity Zones and Patchy Reionization, *ApJ* **670**, 39

Lidz, A., Oh, S. P., and Furlanetto, S. R.: 2006b, Have We Detected Patchy Reionization in Quasar Spectra?, *ApJ* **639**, L47

Liske, J. and Williger, G. M.: 2001, The proximity effect in a close group of QSOs, *MNRAS* **328**, 653

Liu, J., Bi, H., and Fang, L.-Z.: 2007, Lyα Leaks in the Absorption Spectra of High-Redshift Quasi-stellar Objects, *ApJ* **671**, L89

Loeb, A. and Barkana, R.: 2001, The Reionization of the Universe by the First Stars and Quasars, *ARA&A* **39**, 19

Loeb, A. and Eisenstein, D. J.: 1995, Probing Early Clustering with Lyα Absorption Lines beyond the

Quasar Redshift, *ApJ* **448**, 17

Lorenzoni, S., Bunker, A. J., Wilkins, S. M., Stanway, E. R., Jarvis, M. J., and Caruana, J.: 2011, Star-forming galaxies at $z \approx 8-9$ from Hubble Space Telescope/WFC3: implications for reionization, *MNRAS* **414**, 1455

Lu, L., Sargent, W. L. W., Womble, D. S., and Takada-Hidai, M.: 1996, The Lyman-Alpha Forest at $z \approx 4$: Keck HIRES Observations of Q0000-26, *ApJ* **472**, 509

Lu, Y. and Yu, Q.: 2011, On Probing the Properties of QSOs through Their Proximity Effects on the Intergalactic Medium, *ApJ* **736**, 49

Lunnan, R., Vogelsberger, M., Frebel, A., Hernquist, L., Lidz, A., and Boylan-Kolchin, M.: 2011, The Effects of Patchy Reionization on Satellite Galaxies of the Milky Way, *ArXiv e-prints*

Lynden-Bell, D.: 1969, Galactic Nuclei as Collapsed Old Quasars, *Nature* **223**, 690

Lynds, R.: 1971, The Absorption-Line Spectrum of 4C 05.34, *ApJ* **164**, L73+

Madau, P., Haardt, F., and Rees, M. J.: 1999, Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources, *ApJ* **514**, 648

Madau, P. and Meiksin, A.: 1994, The He II Lyman-alpha opacity of the universe, *ApJ* **433**, L53

Madau, P., Meiksin, A., and Rees, M. J.: 1997, 21 Centimeter Tomography of the Intergalactic Medium at High Redshift, *ApJ* **475**, 429

Madau, P., Rees, M. J., Volonteri, M., Haardt, F., and Oh, S. P.: 2004, Early Reionization by Miniquasars, *ApJ* **604**, 484

Maio, U., Koopmans, L. V. E., and Ciardi, B.: 2011, The impact of primordial supersonic flows on early structure formation, reionization and the lowest-mass dwarf galaxies, *MNRAS* **412**, L40

Malhotra, S. and Rhoads, J. E.: 2004, Luminosity Functions of Lyα Emitters at Redshifts $z = 6.5$ and $z = 5.7$: Evidence against Reionization at $z \leq 6.5$, *ApJ* **617**, L5

Mapelli, M., Ferrara, A., and Pierpaoli, E.: 2006, Impact of dark matter decays and annihilations on reionization, *MNRAS* **369**, 1719

Marigo, P., Chiosi, C., and Kudritzki, R.-P.: 2003, Zero-metallicity stars. II. Evolution of very massive objects with mass loss, *A&A* **399**, 617

Marigo, P., Girardi, L., Chiosi, C., and Wood, P. R.: 2001, Zero-metallicity stars. I. Evolution at constant mass, *A&A* **371**, 152

Maselli, A., Ferrara, A., and Gallerani, S.: 2009, Interpreting the transmission windows of distant quasars, *MNRAS* **395**, 1925

Mather, J. C., Cheng, E. S., Eplee, Jr., R. E., Isaacman, R. B., Meyer, S. S., Shafer, R. A., Weiss, R., Wright, E. L., Bennett, C. L., Boggess, N. W., Dwek, E., Gulkis, S., Hauser, M. G., Janssen, M., Kelsall, T., Lubin, P. M., Moseley, Jr., S. H., Murdock, T. L., Silverberg, R. F., Smoot, G. F., and Wilkinson, D. T.: 1990, A preliminary measurement of the cosmic microwave background spectrum by the Cosmic Background Explorer (COBE) satellite, *ApJ* **354**, L37

Matthews, T. A. and Sandage, A. R.: 1963, Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects., *ApJ* **138**, 30

McDonald, P. and Miralda-Escudé, J.: 2001, The Lyα Forest Flux Distribution at $z \sim 5.2$ and the Evolution of the Ionizing Background, *ApJ* **549**, L11

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., and Cen, R.: 2001, A Measurement of the Temperature-Density Relation in the Intergalactic Medium Using a New Lyα Absorption-Line Fitting Method, *ApJ* **562**, 52

McGreer, I. D., Mesinger, A., and Fan, X.: 2011, The first (nearly) model-independent constraint on the neutral hydrogen fraction at $z \sim 5-6$, *MNRAS* **415**, 3237

McKee, C. F. and Tan, J. C.: 2008, The Formation of the First Stars. II. Radiative Feedback Processes and Implications for the Initial Mass Function, *ApJ* **681**, 771

McLure, R. J., Dunlop, J. S., Cirasuolo, M., Koekemoer, A. M., Sabbi, E., Stark, D. P., Targett, T. A., and Ellis, R. S.: 2010, Galaxies at $z = 6-9$ from the WFC3/IR imaging of the Hubble Ultra Deep Field, *MNRAS* **403**, 960

McLure, R. J., Dunlop, J. S., de Ravel, L., Cirasuolo, M., Ellis, R. S., Schenker, M., Robertson, B. E., Koekemoer, A. M., Stark, D. P., and Bowler, R. A. A.: 2011, A robust sample of galaxies at redshifts $6.0 < z < 8.7$: stellar populations, star-formation rates and stellar masses, *ArXiv e-prints*

McQuinn, M., Hernquist, L., Zaldarriaga, M., and Dutta, S.: 2007, Studying reionization with Lyα emitters, *MNRAS* **381**, 75

McQuinn, M., Lidz, A., Zaldarriaga, M., Hernquist, L., and Dutta, S.: 2008, Probing the neutral fraction of the IGM with GRBs during the epoch of reionization, *MNRAS* **388**, 1101

McQuinn, M., Lidz, A., Zaldarriaga, M., Hernquist, L., Hopkins, P. F., Dutta, S., and Faucher-Giguère, C.-A.: 2009, He II Reionization and its Effect on the Intergalactic Medium, *ApJ* **694**, 842

McQuinn, M., Oh, S. P., and Faucher-Giguere, C.-A.: 2011, On Lyman-limit Systems and the Evolution of the Intergalactic Ionizing Background, *ArXiv e-prints*

Meiksin, A.: 2005, Constraints on the ionization sources of the high-redshift intergalactic medium, *MNRAS* **356**, 596

Meiksin, A.: 2006, The possible detection of high-redshift Type II QSOs in deep fields, *MNRAS* **365**, 833

Meiksin, A. and Madau, P.: 1993, On the photoionization of the intergalactic medium by quasars at high redshift, *ApJ* **412**, 34

Meiksin, A. and White, M.: 2003, Constraints on the ultraviolet metagalactic emissivity using the Lyα forest, *MNRAS* **342**, 1205

Meiksin, A. and White, M.: 2004, The effects of ultraviolet background correlations on Lyα forest flux statistics, *MNRAS* **350**, 1107

Meiksin, A. A.: 2009, The physics of the intergalactic medium, *Reviews of Modern Physics* **81**, 1405

Mellema, G., Iliev, I. T., Pen, U.-L., and Shapiro, P. R.: 2006, Simulating cosmic reionization at large scales - II. The 21-cm emission features and statistical signals, *MNRAS* **372**, 679

Mesinger, A.: 2010, Was reionization complete by $z \sim 5-6$?, *MNRAS* **407**, 1328

Mesinger, A. and Furlanetto, S.: 2009, The inhomogeneous ionizing background following reionization, *MNRAS* **400**, 1461

Mesinger, A. and Furlanetto, S. R.: 2008a, Lyα damping wing constraints on inhomogeneous reionization, *MNRAS* **385**, 1348

Mesinger, A. and Furlanetto, S. R.: 2008b, Lyα emitters during the early stages of reionization, *MNRAS*

386, 1990

Messier, C.: 1781, *Catalogue des Nébuleuses & des amas d'Étoiles*, Connoissance des Temps for 1784, Bureau des longitudes, Paris

Metzger, M. R., Djorgovski, S. G., Kulkarni, S. R., Steidel, C. C., Adelberger, K. L., Frail, D. A., Costa, E., and Frontera, F.: 1997, Spectral constraints on the redshift of the optical counterpart to the γ-ray burst of 8 May 1997, *Nature* **387**, 878

Minkowski, R.: 1960, A New Distant Cluster of Galaxies., *ApJ* **132**, 908

Mirabel, I. F., Dijkstra, M., Laurent, P., Loeb, A., and Pritchard, J. R.: 2011, Stellar black holes at the dawn of the universe, *A&A* **528**, A149+

Miralda-Escudé, J.: 1998, Reionization of the Intergalactic Medium and the Damping Wing of the Gunn-Peterson Trough, *ApJ* **501**, 15

Miralda-Escudé, J.: 2003, On the Evolution of the Ionizing Emissivity of Galaxies and Quasars Required by the Hydrogen Reionization, *ApJ* **597**, 66

Miralda-Escudé, J., Haehnelt, M., and Rees, M. J.: 2000, Reionization of the Inhomogeneous Universe, *ApJ* **530**, 1

Miralda-Escudé, J. and Rees, M. J.: 1994, Reionization and thermal evolution of a photoionized intergalactic medium., *MNRAS* **266**, 343

Miralda-Escudé, J. and Rees, M. J.: 1998, Searching for the Earliest Galaxies Using the Gunn-Peterson Trough and the Ly alpha Emission Line, *ApJ* **497**, 21

Modjaz, M.: 2011, Stellar forensics with the supernova-GRB connection, *Astronomische Nachrichten* **332**, 434

Morales, M. F. and Wyithe, J. S. B.: 2010, Reionization and Cosmology with 21-cm Fluctuations, *ARA&A* **48**, 127

Morgan, W. W., Keenan, P. C., and Kellman, E.: 1943, *An atlas of stellar spectra, with an outline of spectral classification*, The University of Chicago press, Chicago, Ill

Mortlock, D. J., Warren, S. J., Venemans, B. P., Patel, M., Hewett, P. C., McMahon, R. G., Simpson, C., Theuns, T., Gonzáles-Solares, E. A., Adamson, A., Dye, S., Hambly, N. C., Hirst, P., Irwin, M. J., Kuiper, E., Lawrence, A., and Röttgering, H. J. A.: 2011, A luminous quasar at a redshift of $z = 7.085$, *Nature* **474**, 616

Murdoch, H. S., Hunstead, R. W., Pettini, M., and Blades, J. C.: 1986, Absorption spectrum of the $z = 3.78$ QSO 2000-330. II - The redshift and equivalent width distributions of primordial hydrogen clouds, *ApJ* **309**, 19

Nakamura, E., Inoue, A. K., Hayashino, T., Horie, M., Kousai, K., Fujii, T., and Matsuda, Y.: 2011, Lyα emitters at $z = 6.5$ in the SSA22 field: an area more neutral or void at the end of the re-ionization epoch, *MNRAS* **412**, 2579

Nestor, D. B., Shapley, A. E., Steidel, C. C., and Siana, B.: 2011, Narrowband Imaging of Escaping Lyman-continuum Emission in the SSA22 Field, *ApJ* **736**, 18

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Carollo, C. M., Franx, M., Labbé, I., Magee, D., Stiavelli, M., Trenti, M., and van Dokkum, P. G.: 2010, $z \sim 7$ Galaxies in the HUDF: First Epoch WFC3/IR Results, *ApJ* **709**, L16

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbe, I., Trenti, M., Gonzalez, V., Carollo, C. M., Franx, M., van Dokkum, P. G., and Magee, D.: 2011, Expanded Search for $z \sim 10$ Galaxies from HUDF09, ERS, and CANDELS Data: Evidence for Accelerated Evolution at $z > 8$?, *ArXiv e-prints*

Oh, S. P.: 2001, Reionization by Hard Photons. I. X-Rays from the First Star Clusters, *ApJ* **553**, 499

Oh, S. P. and Furlanetto, S. R.: 2005, How Universal is the Gunn-Peterson Trough at $z \sim 6$?: A Closer Look at the Quasar SDSS J1148+5251, *ApJ* **620**, L9

Oke, J. B. and Korycansky, D. G.: 1982, Absolute spectrophotometry of very large redshift quasars, *ApJ* **255**, 11

Olive, K. A. and Skillman, E. D.: 2004, A Realistic Determination of the Error on the Primordial Helium Abundance: Steps toward Nonparametric Nebular Helium Abundances, *ApJ* **617**, 29

Ono, Y., Ouchi, M., Shimasaku, K., Dunlop, J., Farrah, D., McLure, R., and Okamura, S.: 2010, Stellar Populations of Lyα Emitters at $z \sim 6-7$: Constraints on the Escape Fraction of Ionizing Photons from Galaxy Building Blocks, *ApJ* **724**, 1524

Ouchi, M., Mobasher, B., Shimasaku, K., Ferguson, H. C., Fall, S. M., Ono, Y., Kashikawa, N., Morokuma, T., Nakajima, K., Okamura, S., Dickinson, M., Giavalisco, M., and Ohta, K.: 2009, Large Area Survey for $z = 7$ Galaxies in SDF and GOODS-N: Implications for Galaxy Formation and Cosmic Reionization, *ApJ* **706**, 1136

Ouchi, M., Shimasaku, K., Furusawa, H., Saito, T., Yoshida, M., Akiyama, M., Ono, Y., Yamada, T., Ota, K., Kashikawa, N., Iye, M., Kodama, T., Okamura, S., Simpson, C., and Yoshida, M.: 2010, Statistics of 207 Lyα Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models, *ApJ* **723**, 869

Ouyed, R., Pudritz, R. E., and Jaikumar, P.: 2009, Quark-Novae, Cosmic Reionization, and Early r-Process Element Production, *ApJ* **702**, 1575

Paardekooper, J.-P., Pelupessy, F. I., Altay, G., and Kruip, C. J. H.: 2011, The escape of ionising radiation from high-redshift dwarf galaxies, *A&A* **530**, A87+

Paciga, G., Chang, T.-C., Gupta, Y., Nityanada, R., Odegova, J., Pen, U.-L., Peterson, J. B., Roy, J., and Sigurdson, K.: 2011, The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at $z \approx 8.6$, *MNRAS* **413**, 1174

Page, L., Hinshaw, G., Komatsu, E., Nolta, M. R., Spergel, D. N., Bennett, C. L., Barnes, C., Bean, R., Doré, O., Dunkley, J., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Peiris, H. V., Tucker, G. S., Verde, L., Weiland, J. L., Wollack, E., and Wright, E. L.: 2007, Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis, *ApJS* **170**, 335

Partl, A. M., Müller, V., Yepes, G., and Gottlöber, S.: 2011, Large-scale environmental bias of the high-redshift quasar line-of-sight proximity effect, *MNRAS* **415**, 3851

Partridge, R. B. and Peebles, P. J. E.: 1967, Are Young Galaxies Visible?, *ApJ* **147**, 868

Pascarelle, S. M., Lanzetta, K. M., Chen, H., and Webb, J. K.: 2001, Discovery of the Galaxy Proximity Effect and Implications for Measurements of the Ionizing Background Radiation at Low Redshifts, *ApJ* **560**, 101

Paschos, P., Norman, M. L., Bordner, J. O., and Harkness, R.: 2007, Late Reheating of the IGM by

Quasars: A Radiation Hydrodynamical Simulation of Helium II Reionization, *ArXiv e-prints*

Peebles, P. J. E.: 1982, Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations, *ApJ* **263**, L1

Peeples, M. S., Weinberg, D. H., Davé, R., Fardal, M. A., and Katz, N.: 2010, Pressure support versus thermal broadening in the Lyman α forest - II. Effects of the equation of state on transverse structure, *MNRAS* **404**, 1295

Pentericci, L., Fan, X., Rix, H., Strauss, M. A., Narayanan, V. K., Richards, G. T., Schneider, D. P., Krolik, J., Heckman, T., Brinkmann, J., Lamb, D. Q., and Szokoly, G. P.: 2002, VLT Optical and Near-Infrared Observations of the $z = 6.28$ Quasar SDSS J1030+0524, *AJ* **123**, 2151

Penzias, A. A. and Wilson, R. W.: 1965, A Measurement of Excess Antenna Temperature at 4080 Mc/s., *ApJ* **142**, 419

Percival, W. J., Reid, B. A., Eisenstein, D. J., Bahcall, N. A., Budavari, T., Frieman, J. A., Fukugita, M., Gunn, J. E., Ivezić, Ž., Knapp, G. R., Kron, R. G., Loveday, J., Lupton, R. H., McKay, T. A., Meiksin, A., Nichol, R. C., Pope, A. C., Schlegel, D. J., Schneider, D. P., Spergel, D. N., Stoughton, C., Strauss, M. A., Szalay, A. S., Tegmark, M., Vogeley, M. S., Weinberg, D. H., York, D. G., and Zehavi, I.: 2010, Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample, *MNRAS* **401**, 2148

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J., and The Supernova Cosmology Project: 1999, Measurements of Omega and Lambda from 42 High-Redshift Supernovae, *ApJ* **517**, 565

Péroux, C., McMahon, R. G., Storrie-Lombardi, L. J., and Irwin, M. J.: 2003, The evolution of $\Omega_{\rm HI}$ and the epoch of formation of damped Lyman α absorbers, *MNRAS* **346**, 1103

Petitjean, P. and Vergani, S. D.: 2011, Gamma-ray bursts as probes of the distant Universe, *Comptes Rendus Physique* **12**, 288

Powell, L. C., Slyz, A., and Devriendt, J.: 2011, The impact of supernova-driven winds on stream-fed protogalaxies, *MNRAS* **414**, 3671

Press, W. H. and Schechter, P.: 1974, Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation, *ApJ* **187**, 425

Pritchard, J. R., Loeb, A., and Wyithe, J. S. B.: 2010, Constraining reionization using 21-cm observations in combination with CMB and Lyα forest data, *MNRAS* **408**, 57

Prochaska, J. X., O'Meara, J. M., and Worseck, G.: 2010, A Definitive Survey for Lyman Limit Systems at $z \sim 3.5$ with the Sloan Digital Sky Survey, *ApJ* **718**, 392

Puchwein, E., Pfrommer, C., Springel, V., Broderick, A. E., and Chang, P.: 2011, The Lyman-alpha forest in a blazar-heated Universe, *ArXiv e-prints*

Raičević, M., Theuns, T., and Lacey, C.: 2011, The galaxies that reionized the Universe, *MNRAS* **410**, 775

Rauch, M.: 1998, The Lyman Alpha Forest in the Spectra of QSOs, *ARA&A* **36**, 267

Rauch, M., Becker, G. D., Haehnelt, M. G., Gauthier, J.-R., Ravindranath, S., and Sargent, W. L. W.: 2011, Filamentary Infall of Cold Gas and Escape of Lyman Alpha and Hydrogen Ionizing Radiation from an Interacting High-Redshift Galaxy, *ArXiv e-prints*

Rauch, M., Miralda-Escude, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., and Ostriker, J. P.: 1997, The Opacity of the Lyα Forest and Implications for Ω_b and the Ionizing Background, *ApJ* **489**, 7

Razoumov, A. O. and Sommer-Larsen, J.: 2010, Ionizing Radiation from $z = 4 - 10$ Galaxies, *ApJ* **710**, 1239

Rees, M. J.: 1998, The Universe at $z > 5$: When and How did the "Dark Age" End?, *Proceedings of the National Academy of Science* **95**, 47

Rees, M. J. and Ostriker, J. P.: 1977, Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters, *MNRAS* **179**, 541

Rees, M. J. and Setti, G.: 1970, Absorption and Scattering of Ultraviolet and X Ray Photons by Intergalactic Gas, *A&A* **8**, 410

Richard, J., Pelló, R., Schaerer, D., Le Borgne, J., and Kneib, J.: 2006, Constraining the population of $6 \lesssim z \lesssim 10$ star-forming galaxies with deep near-IR images of lensing clusters, *A&A* **456**, 861

Richard, J., Stark, D. P., Ellis, R. S., George, M. R., Egami, E., Kneib, J., and Smith, G. P.: 2008, A Hubble and Spitzer Space Telescope Survey for Gravitationally Lensed Galaxies: Further Evidence for a Significant Population of Low-Luminosity Galaxies beyond $z = 7$, *ApJ* **685**, 705

Richards, G. T., Strauss, M. A., Fan, X., Hall, P. B., Jester, S., Schneider, D. P., Vanden Berk, D. E., Stoughton, C., Anderson, S. F., Brunner, R. J., Gray, J., Gunn, J. E., Ivezić, Ž., Kirkland, M. K., Knapp, G. R., Loveday, J., Meiksin, A., Pope, A., Szalay, A. S., Thakar, A. R., Yanny, B., York, D. G., Barentine, J. C., Brewington, H. J., Brinkmann, J., Fukugita, M., Harvanek, M., Kent, S. M., Kleinman, S. J., Krzesiński, J., Long, D. C., Lupton, R. H., Nash, T., Neilsen, Jr., E. H., Nitta, A., Schlegel, D. J., and Snedden, S. A.: 2006, The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3, *AJ* **131**, 2766

Richards, G. T., Vanden Berk, D. E., Reichard, T. A., Hall, P. B., Schneider, D. P., SubbaRao, M., Thakar, A. R., and York, D. G.: 2002, Broad Emission-Line Shifts in Quasars: An Orientation Measure for Radio-Quiet Quasars?, *AJ* **124**, 1

Ricotti, M.: 2004, in H. J. G. L. M. Lamers, L. J. Smith, and A. Nota (eds.), *The Formation and Evolution of Massive Young Star Clusters*, Vol. 322 of *Astronomical Society of the Pacific Conference Series*, pp 509–+

Ricotti, M., Gnedin, N. Y., and Shull, J. M.: 2000, The Evolution of the Effective Equation of State of the Intergalactic Medium, *ApJ* **534**, 41

Ricotti, M. and Ostriker, J. P.: 2004a, Reionization, chemical enrichment and seed black holes from the first stars: is Population III important?, *MNRAS* **350**, 539

Ricotti, M. and Ostriker, J. P.: 2004b, X-ray pre-ionization powered by accretion on the first black holes - I. A model for the WMAP polarization measurement, *MNRAS* **352**, 547

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schom-

mer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J.: 1998, Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, *AJ* **116**, 1009

Riess, A. G., Macri, L., Casertano, S., Sosey, M., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W., Chornock, R., and Sarkar, D.: 2009, A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder, *ApJ* **699**, 539

Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J., and Stark, D. P.: 2010, Early star-forming galaxies and the reionization of the Universe, *Nature* **468**, 49

Robertson, H. P.: 1935, Kinematics and World-Structure, *ApJ* **82**, 284

Rollinde, E., Srianand, R., Theuns, T., Petitjean, P., and Chand, H.: 2005, The density structure around quasars from optical depth statistics, *MNRAS* **361**, 1015

Rubin, V. C., Ford, W. K. J., and . Thonnard, N.: 1980, Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 $R = 4$ kpc to UGC 2885 $R = 122$ kpc, *ApJ* **238**, 471

Salvaterra, R., Della Valle, M., Campana, S., Chincarini, G., Covino, S., D'Avanzo, P., Fernández-Soto, A., Guidorzi, C., Mannucci, F., Margutti, R., Thöne, C. C., Antonelli, L. A., Barthelmy, S. D., de Pasquale, M., D'Elia, V., Fiore, F., Fugazza, D., Hunt, L. K., Maiorano, E., Marinoni, S., Marshall, F. E., Molinari, E., Nousek, J., Pian, E., Racusin, J. L., Stella, L., Amati, L., Andreuzzi, G., Cusumano, G., Fenimore, E. E., Ferrero, P., Giommi, P., Guetta, D., Holland, S. T., Hurley, K., Israel, G. L., Mao, J., Markwardt, C. B., Masetti, N., Pagani, C., Palazzi, E., Palmer, D. M., Piranomonte, S., Tagliaferri, G., and Testa, V.: 2009, GRB090423 at a redshift of $z \sim 8.1$, *Nature* **461**, 1258

Sandage, A.: 1965, The Existence of a Major New Constituent of the Universe: the Quasistellar Galaxies., *ApJ* **141**, 1560

Santoro, F. and Shull, J. M.: 2006, Critical Metallicity and Fine-Structure Emission of Primordial Gas Enriched by the First Stars, *ApJ* **643**, 26

Sargent, W. L. W., Young, P. J., Boksenberg, A., and Tytler, D.: 1980, The distribution of Lyman-alpha absorption lines in the spectra of six QSOs - Evidence for an intergalactic origin, *ApJS* **42**, 41

Schaerer, D.: 2002, On the properties of massive Population III stars and metal-free stellar populations, *A&A* **382**, 28

Schaerer, D. and Charbonnel, C.: 2011, A new perspective on globular clusters, their initial mass function and their contribution to the stellar halo and the cosmic reionization, *MNRAS* **413**, 2297

Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., and Sargent, W. L. W.: 2000, The thermal history of the intergalactic medium, *MNRAS* **318**, 817

Schirber, M., Miralda-Escudé, J., and McDonald, P.: 2004, The Transverse Proximity Effect: A Probe to the Environment, Anisotropy, and Megayear Variability of QSOs, *ApJ* **610**, 105

Schmidt, M.: 1963, 3C 273 : A Star-Like Object with Large Red-Shift, *Nature* **197**, 1040

Schmidt, M.: 1965, Large Redshifts of Five Quasi-Stellar Sources., *ApJ* **141**, 1295

Schneider, R., Ferrara, A., Natarajan, P., and Omukai, K.: 2002, First Stars, Very Massive Black Holes, and Metals, *ApJ* **571**, 30

Scott, D., Rees, M. J., and Sciama, D. W.: 1991, Dark matter decay, reionization and microwave background anisotropies, *A&A* **250**, 295

Scott, J., Bechtold, J., Dobrzycki, A., and Kulkarni, V. P.: 2000, A Uniform Analysis of the Lyα Forest at $z = 0-5$. II. Measuring the Mean Intensity of the Extragalactic Ionizing Background Using the Proximity Effect, *ApJS* **130**, 67

Scott, J., Bechtold, J., Morita, M., Dobrzycki, A., and Kulkarni, V. P.: 2002, A Uniform Analysis of the Lyα Forest at $z = 0-5$. V. The Extragalactic Ionizing Background at Low Redshift, *ApJ* **571**, 665

Scott, J. E., Kriss, G. A., Brotherton, M., Green, R. F., Hutchings, J., Shull, J. M., and Zheng, W.: 2004, A Composite Extreme-Ultraviolet QSO Spectrum from FUSE, *ApJ* **615**, 135

Shapiro, P. R.: 1986, Cosmological H II regions and the photoionization of the intergalactic medium, *PASP* **98**, 1014

Shapiro, P. R. and Giroux, M. L.: 1987, Cosmological H II regions and the photoionization of the intergalactic medium, *ApJ* **321**, L107

Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., and Erb, D. K.: 2006, The Direct Detection of Lyman Continuum Emission from Star-forming Galaxies at $z \sim 3$, *ApJ* **651**, 688

Shull, J. M. and Venkatesan, A.: 2008, Constraints on First-Light Ionizing Sources from Optical Depth of the Cosmic Microwave Background, *ApJ* **685**, 1

Shull, M., Harness, A., Trenti, M., and Smith, B.: 2011, Critical Star-Formation Rates for Reionization: Full Reionization occurs at $z = 7$, *ArXiv e-prints*

Sijacki, D., Springel, V., and Haehnelt, M. G.: 2009, Growing the first bright quasars in cosmological simulations of structure formation, *MNRAS* **400**, 100

Silk, J.: 1977, On the fragmentation of cosmic gas clouds. I - The formation of galaxies and the first generation of stars, *ApJ* **211**, 638

Slipher, V. M.: 1913, The radial velocity of the Andromeda Nebula, *Lowell Observatory Bulletin* **2**, 56

Slipher, V. M.: 1915, Spectrographic Observations of Nebulae, *Popular Astronomy* **23**, 21

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F., Tenorio, L., Weiss, R., and Wilkinson, D. T.: 1992, Structure in the COBE differential microwave radiometer first-year maps, *ApJ* **396**, L1

Sokasian, A., Yoshida, N., Abel, T., Hernquist, L., and Springel, V.: 2004, Cosmic reionization by stellar sources: population III stars, *MNRAS* **350**, 47

Somerville, R. S., Bullock, J. S., and Livio, M.: 2003, The Epoch of Reionization in Models with Reduced Small-Scale Power, *ApJ* **593**, 616

Songaila, A.: 2004, The Evolution of the Intergalactic Medium Transmission to Redshift 6, *AJ* **127**, 2598

Songaila, A. and Cowie, L. L.: 2010, The Evolution of Lyman Limit Absorption Systems to Redshift Six, *ApJ* **721**, 1448

Songaila, A., Hu, E. M., Cowie, L. L., and McMahon, R. G.: 1999, Limits on the Gunn-Peterson Effect at $z = 5$, *ApJ* **525**, L5

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E.,

and Wright, E. L.: 2003, First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters, *ApJS* **148**, 175

Springel, V.: 2005, The cosmological simulation code GADGET-2, *MNRAS* **364**, 1105

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F.: 2005, Simulations of the formation, evolution and clustering of galaxies and quasars, *Nature* **435**, 629

Srbinovsky, J. A. and Wyithe, J. S. B.: 2007, Constraining the quasar contribution to the reionization of cosmic hydrogen, *MNRAS* **374**, 627

Srbinovsky, J. A. and Wyithe, J. S. B.: 2010, The Fraction of Ionizing Photons Escaping from High-Redshift Galaxies, *Publications of the Astronomical Society of Australia* **27**, 110

Stacy, A., Bromm, V., and Loeb, A.: 2011, Effect of Streaming Motion of Baryons Relative to Dark Matter on the Formation of the First Stars, *ApJ* **730**, L1+

Stacy, A., Greif, T. H., and Bromm, V.: 2010, The first stars: formation of binaries and small multiple systems, *MNRAS* **403**, 45

Stark, D. P., Ellis, R. S., Richard, J., Kneib, J., Smith, G. P., and Santos, M. R.: 2007, A Keck Survey for Gravitationally Lensed Lyα Emitters in the Redshift Range $8.5 < z < 10.4$: New Constraints on the Contribution of Low-Luminosity Sources to Cosmic Reionization, *ApJ* **663**, 10

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., and Pettini, M.: 1999, Lyman-Break Galaxies at $z \gtrsim 4$ and the Evolution of the Ultraviolet Luminosity Density at High Redshift, *ApJ* **519**, 1

Stiavelli, M., Fall, S. M., and Panagia, N.: 2004, The Possible Detection of Cosmological Reionization Sources, *ApJ* **610**, L1

Stiavelli, M., Mather, J., Clampin, M., Doyon, R., Flanagan, K., Franx, M., Gardner, J., Greenhouse, M., Hammel, H., Hutchings, J., Jakobsen, P., Lilly, S., McCaughrean, M., Mountain, M., Rieke, G., Sonneborn, G., Windhorst, R., and Wright, G.: 2009, in *astro2010: The Astronomy and Astrophysics Decadal Survey*, Vol. 2010 of *Astronomy*, pp 287–+

Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J., and Hazard, C.: 1994, Evolution of Lyman-limit absorption systems over the redshift range $0.40 < z < 4.69$, *ApJ* **427**, L13

Strömgren, B.: 1939, The Physical State of Interstellar Hydrogen., *ApJ* **89**, 526

Sundman, K. F.: 1913, Mémoire sur le problème des trois corps, *Acta Math.* **36**, 105

Tanvir, N. R., Fox, D. B., Levan, A. J., Berger, E., Wiersema, K., Fynbo, J. P. U., Cucchiara, A., Krühler, T., Gehrels, N., Bloom, J. S., Greiner, J., Evans, P. A., Rol, E., Olivares, F., Hjorth, J., Jakobsson, P., Farihi, J., Willingale, R., Starling, R. L. C., Cenko, S. B., Perley, D., Maund, J. R., Duke, J., Wijers, R. A. M. J., Adamson, A. J., Allan, A., Bremer, M. N., Burrows, D. N., Castro-Tirado, A. J., Cavanagh, B., de Ugarte Postigo, A., Dopita, M. A., Fatkhullin, T. A., Fruchter, A. S., Foley, R. J., Gorosabel, J., Kennea, J., Kerr, T., Klose, S., Krimm, H. A., Komarova, V. N., Kulkarni, S. R., Moskvitin, A. S., Mundell, C. G., Naylor, T., Page, K., Penprase, B. E., Perri, M., Podsiadlowski, P., Roth, K., Rutledge, R. E., Sakamoto, T., Schady, P., Schmidt, B. P., Soderberg, A. M., Sollerman, J., Stephens, A. W., Stratta, G., Ukwatta, T. N., Watson, D., Westra, E., Wold, T., and Wolf, C.: 2009, A γ-ray burst at a redshift of $z \sim 8.2$,

Nature **461**, 1254

Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T., and Palla, F.: 1997, How Small Were the First Cosmological Objects?, *ApJ* **474**, 1

Telfer, R. C., Zheng, W., Kriss, G. A., and Davidsen, A. F.: 2002, The Rest-Frame Extreme-Ultraviolet Spectral Properties of Quasi-stellar Objects, *ApJ* **565**, 773

Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., and Thomas, P. A.: 1998, P^3M-SPH simulations of the Lyα forest, *MNRAS* **301**, 478

Theuns, T., Schaye, J., Zaroubi, S., Kim, T.-S., Tzanavaris, P., and Carswell, B.: 2002a, Constraints on Reionization from the Thermal History of the Intergalactic Medium, *ApJ* **567**, L103

Theuns, T. and Zaroubi, S.: 2000, A wavelet analysis of the spectra of quasi-stellar objects, *MNRAS* **317**, 989

Theuns, T., Zaroubi, S., Kim, T.-S., Tzanavaris, P., and Carswell, R. F.: 2002b, Temperature fluctuations in the intergalactic medium, *MNRAS* **332**, 367

Thöne, C. C., Campana, S., Lazzati, D., de Ugarte Postigo, A., Fynbo, J. P. U., Christensen, L., Levan, A. J., Aloy, M. A., Hjorth, J., Jakobsson, P., Levesque, E. M., Malesani, D., Milvang-Jensen, B., Roming, P. W. A., Tanvir, N. R., Wiersema, K., Gladders, M., Wuyts, E., and Dahle, H.: 2011, Variable Lyα sheds light on the environment surrounding GRB 090426, *MNRAS* **414**, 479

Tilvi, V., Rhoads, J. E., Hibon, P., Malhotra, S., Wang, J., Veilleux, S., Swaters, R., Probst, R., Krug, H., Finkelstein, S. L., and Dickinson, M.: 2010, The Luminosity Function of Lyα Emitters at Redshift $z = 7.7$, *ApJ* **721**, 1853

Tittley, E. R. and Meiksin, A.: 2007, Reionization scenarios and the temperature of the intergalactic medium, *MNRAS* **380**, 1369

Totani, T., Kawai, N., Kosugi, G., Aoki, K., Yamada, T., Iye, M., Ohta, K., and Hattori, T.: 2006, Implications for Cosmic Reionization from the Optical Afterglow Spectrum of the Gamma-Ray Burst 050904 at $z = 6.3$, *PASJ* **58**, 485

Trac, H., Cen, R., and Loeb, A.: 2008, Imprint of Inhomogeneous Hydrogen Reionization on the Temperature Distribution of the Intergalactic Medium, *ApJ* **689**, L81

Trac, H. and Gnedin, N. Y.: 2009, Computer Simulations of Cosmic Reionization, *ArXiv e-prints*

Trenti, M., Bradley, L. D., Stiavelli, M., Oesch, P., Treu, T., Bouwens, R. J., Shull, J. M., MacKenty, J. W., Carollo, C. M., and Illingworth, G. D.: 2011, The Brightest of Reionizing Galaxies Survey: Design and Preliminary Results, *ApJ* **727**, L39+

Tseliakhovich, D., Barkana, R., and Hirata, C. M.: 2011, Suppression and spatial variation of early galaxies and minihaloes, *MNRAS* pp 1501–+

Tumlinson, J.: 2006, Chemical Evolution in Hierarchical Models of Cosmic Structure. I. Constraints on the Early Stellar Initial Mass Function, *ApJ* **641**, 1

Tumlinson, J. and Shull, J. M.: 2000, Zero-Metallicity Stars and the Effects of the First Stars on Reionization, *ApJ* **528**, L65

Tumlinson, J., Shull, J. M., and Venkatesan, A.: 2003, Cosmological Effects of the First Stars: Evolving Spectra of Population III, *ApJ* **584**, 608

Turk, M. J., Abel, T., and O'Shea, B.: 2009, The Formation of Population III Binaries from Cosmological

Initial Conditions, *Science* **325**, 601

Tytler, D.: 1982, QSO Lyman limit absorption, *Nature* **298**, 427

Tytler, D.: 1987, The redshift distribution of QSO Lyman-alpha absorbption systems, *ApJ* **321**, 69

Tytler, D., Kirkman, D., O'Meara, J. M., Suzuki, N., Orin, A., Lubin, D., Paschos, P., Jena, T., Lin, W., Norman, M. L., and Meiksin, A.: 2004, Cosmological Parameters σ_8, the Baryon Density Ω_b, the Vacuum Energy Density Ω_Λ, the Hubble Constant and the UV Background Intensity from a Calibrated Measurement of HI Lyα Absorption at $z = 1.9$, *ApJ* **617**, 1

Utsumi, Y., Goto, T., Kashikawa, N., Miyazaki, S., Komiyama, Y., Furusawa, H., and Overzier, R.: 2010, A Large Number of $z > 6$ Galaxies Around a QSO at $z = 6.43$: Evidence for a Protocluster?, *ApJ* **721**, 1680

Vanzella, E., Pentericci, L., Fontana, A., Grazian, A., Castellano, M., Boutsia, K., Cristiani, S., Dickinson, M., Gallozzi, S., Giallongo, E., Giavalisco, M., Maiolino, R., Moorwood, A., Paris, D., and Santini, P.: 2011, Spectroscopic Confirmation of Two Lyman Break Galaxies at Redshift Beyond 7, *ApJ* **730**, L35+

Venkatesan, A., Giroux, M. L., and Shull, J. M.: 2001, Heating and Ionization of the Intergalactic Medium by an Early X-Ray Background, *ApJ* **563**, 1

Venkatesan, A., Tumlinson, J., and Shull, J. M.: 2003, Evolving Spectra of Population III Stars: Consequences for Cosmological Reionization, *ApJ* **584**, 621

Viel, M.: 2009, The Intergalactic Medium as a Cosmological Tool, *Nuclear Physics B Proceedings Supplements* **194**, 156

Viel, M., Bolton, J. S., and Haehnelt, M. G.: 2009, Cosmological and astrophysical constraints from the Lyman α forest flux probability distribution function, *MNRAS* **399**, L39

Vogt, S. S., Allen, S. L., Bigelow, B. C., Bresee, L., Brown, B., Cantrall, T., Conrad, A., Couture, M., Delaney, C., Epps, H. W., Hilyard, D., Hilyard, D. F., Horn, E., Jern, N., Kanto, D., Keane, M. J., Kibrick, R. I., Lewis, J. W., Osborne, J., Pardeilhan, G. H., Pfister, T., Ricketts, T., Robinson, L. B., Stover, R. J., Tucker, D., Ward, J., and Wei, M. Z.: 1994, in D. L. Crawford and E. R. Craine (eds.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 2198 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp 362–+

Volonteri, M.: 2010, Formation of supermassive black holes, *A&A Rev.* **18**, 279

von Hoerner, S.: 1960, Die numerische Integration des n-Körper-Problemes für Sternhaufen. I, *ZAp* **50**, 184

von Hoerner, S.: 1963, Die numerische Integration des n-Körper-Problems für Sternhaufen, II., *ZAp* **57**, 47

Walker, A. G.: 1935, On the formal comparison of Milne's kinematical system with the systems of general relativity, *MNRAS* **95**, 263

Wang, Q.-D.: 1991, The global solution of the n-body problem, *Celestial Mechanics and Dynamical Astronomy* **50**, 73

Weinberg, D. H., Burles, S., Croft, R. A. C., Dave', R., Gomez, G., Hernquist, L., Katz, N., Kirkman, D., Liu, S., Miralda-Escudé, J., Pettini, M., Phillips, J., Tytler, D., and Wright, J.: 1999, in A. J. Banday, R. K. Sheth, and L. N. da Costa (eds.), *Evolution of Large Scale Structure : From Recombination to*

Garching, pp 346–+

Weinberg, D. H., Hernsquit, L., Katz, N., Croft, R., and Miralda-Escudé, J.: 1997, in P. Petitjean and S. Charlot (eds.), *Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System*, pp 133–+

Weymann, R. J., Carswell, R. F., and Smith, M. G.: 1981, Absorption lines in the spectra of quasistellar objects, *ARA&A* **19**, 41

White, R. L., Becker, R. H., Fan, X., and Strauss, M. A.: 2005, Hubble Space Telescope Advanced Camera for Surveys Observations of the $z = 6.42$ Quasar SDSS J1148+5251: A Leak in the Gunn-Peterson Trough, *AJ* **129**, 2102

White, S. D. M. and Frenk, C. S.: 1991, Galaxy formation through hierarchical clustering, *ApJ* **379**, 52

White, S. D. M. and Rees, M. J.: 1978, Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering, *MNRAS* **183**, 341

Wien, W.: 1894, Temperatur und Entropie der Strahlung, *Annalen der Physik* **288**, 132

Williger, G. M., Baldwin, J. A., Carswell, R. F., Cooke, A. J., Hazard, C., Irwin, M. J., McMahon, R. G., and Storrie-Lombardi, L. J.: 1994, Lyman-alpha absorption in the spectrum of the $z = 4.5$ QSO BR 1033-0327, *ApJ* **428**, 574

Willott, C. J.: 2011, A Monster in the early Universe, *Nature* **474**, 583

Wise, J. H. and Abel, T.: 2007, Resolving the Formation of Protogalaxies. I. Virialization, *ApJ* **665**, 899

Wise, J. H. and Abel, T.: 2008, Resolving the Formation of Protogalaxies. III. Feedback from the First Stars, *ApJ* **685**, 40

Wise, J. H. and Cen, R.: 2009, Ionizing Photon Escape Fractions From High-Redshift Dwarf Galaxies, *ApJ* **693**, 984

Wolfe, A. M., Turnshek, D. A., Smith, H. E., and Cohen, R. D.: 1986, Damped Lyman-alpha absorption by disk galaxies with large redshifts. I - The Lick survey, *ApJS* **61**, 249

Woosley, S. E. and Bloom, J. S.: 2006, The Supernova Gamma-Ray Burst Connection, *ARA&A* **44**, 507

Worseck, G., Prochaska, J. X., McQuinn, M., Dall'Aglio, A., Fechner, C., Hennawi, J. F., Reimers, D., Richter, P., and Wisotzki, L.: 2011, The End of Helium Reionization at $z \simeq 2.7$ Inferred from Cosmic Variance in HST/COS He II Lyα Absorption Spectra, *ApJ* **733**, L24+

Wyithe, J. S. B. and Bolton, J. S.: 2011, Near-zone sizes and the rest-frame extreme ultraviolet spectral index of the highest redshift quasars, *MNRAS* **412**, 1926

Wyithe, J. S. B., Bolton, J. S., and Haehnelt, M. G.: 2008, Reionization bias in high-redshift quasar near-zones, *MNRAS* **383**, 691

Wyithe, J. S. B., Hopkins, A. M., Kistler, M. D., Yüksel, H., and Beacom, J. F.: 2010, Determining the escape fraction of ionizing photons during reionization with the GRB-derived star formation rate, *MNRAS* **401**, 2561

Wyithe, J. S. B. and Loeb, A.: 2003a, Reionization of Hydrogen and Helium by Early Stars and Quasars, *ApJ* **586**, 693

Wyithe, J. S. B. and Loeb, A.: 2003b, Was the Universe Reionized by Massive Metal-free Stars?, *ApJ* **588**, L69

Wyithe, J. S. B. and Loeb, A.: 2004, A characteristic size of ~ 10 Mpc for the ionized bubbles at the end

of cosmic reionization, *Nature* **432**, 194

Wyithe, J. S. B. and Loeb, A.: 2008, Fluctuations in 21-cm emission after reionization, *MNRAS* **383**, 606

Wyithe, J. S. B. and Loeb, A.: 2011, Extrapolating the evolution of galaxy sizes to the epoch of reionization, *MNRAS* **413**, L38

Wyithe, J. S. B., Loeb, A., and Barnes, D. G.: 2005, Prospects for Redshifted 21 cm Observations of Quasar H II Regions, *ApJ* **634**, 715

Yajima, H., Choi, J.-H., and Nagamine, K.: 2011, Escape fraction of ionizing photons from high-redshift galaxies in cosmological SPH simulations, *MNRAS* **412**, 411

Yan, H. and Windhorst, R. A.: 2004, The Major Sources of the Cosmic Reionizing Background at $z \simeq 6$, *ApJ* **600**, L1

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C., Hull, C., Husby, D., Ichikawa, S., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Neilsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sandford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S., Watanabe, M., Weinberg, D. H., Yanny, B., and Yasuda, N.: 2000, The Sloan Digital Sky Survey: Technical Summary, *AJ* **120**, 1579

Yoshida, M., Shimasaku, K., Kashikawa, N., Ouchi, M., Okamura, S., Ajiki, M., Akiyama, M., Ando, H., Aoki, K., Doi, M., Furusawa, H., Hayashino, T., Iwamuro, F., Iye, M., Karoji, H., Kobayashi, N., Kodaira, K., Kodama, T., Komiyama, Y., Malkan, M. A., Matsuda, Y., Miyazaki, S., Mizumoto, Y., Morokuma, T., Motohara, K., Murayama, T., Nagao, T., Nariai, K., Ohta, K., Sasaki, T., Sato, Y., Sekiguchi, K., Shioya, Y., Tamura, H., Taniguchi, Y., Umemura, M., Yamada, T., and Yasuda, N.: 2006, Luminosity Functions of Lyman Break Galaxies at $z \sim 4$ and $z \sim 5$ in the Subaru Deep Field, *ApJ* **653**, 988

Yoshida, N., Abel, T., Hernquist, L., and Sugiyama, N.: 2003, Simulations of Early Structure Formation:

Primordial Gas Clouds, *ApJ* **592**, 645

Zaldarriaga, M.: 2002, Searching for Fluctuations in the Intergalactic Medium Temperature Using the Lyα Forest, *ApJ* **564**, 153

Zaldarriaga, M., Hui, L., and Tegmark, M.: 2001, Constraints from the Lyα Forest Power Spectrum, *ApJ* **557**, 519

Zhang, Y., Anninos, P., and Norman, M. L.: 1995, A Multispecies Model for Hydrogen and Helium Absorbers in Lyman-Alpha Forest Clouds, *ApJ* **453**, L57+

Zwicky, F.: 1929, On the Red Shift of Spectral Lines through Interstellar Space, *Proceedings of the National Academy of Science* **15**, 773

Zwicky, F.: 1933, Die Rotverschiebung von extragalaktischen Nebeln, *Helvetica Physica Acta* **6**, 110