The relationship between Gas and Galaxies

Allen Maurice Shone

How to cite:

Shone, Allen Maurice (2009) The relationship between Gas and Galaxies. Doctoral thesis, Durham University.

The Relationship between Gas and Galaxies

Allen Maurice Shone

A Thesis presented for the degree of
Doctor of Philosophy

Extragalactic Astronomy and Cosmology Research Group
Department of Physics
Durham University
United Kingdom
May 2009

Dedicated to those who dedicated their time to teach me Natural Philosophy

The Relationship Between Gas and Galaxies

Allen Maurice Shone

Submitted for the degree of Doctor of Philosophy
May 2009

Abstract

We have investigated the 2D 2-point correlation function, $\xi_{\rm AG}$, between low column density Lyα absorbers and galaxies at a redshift z $\sim$ 1 for the first time. 141 Lyα absorbers between redshifts z = 0.68 $\rightarrow$ 1.51 over a total redshift path length of Δz = 1.09 were collected from HST STIS E230M absorption spectra towards the quasars HE 1122-1648 (z = 2.4) and PKS 1127-145 (z = 1.187). The column density of the correlated Lyα absorbers ranged from 13.18 $\leq\log_{10}({\rm N_{HI}}\,({\rm cm^{-2}}))\leq$ 17.42, with a median column density of $\log_{10}({\rm N_{HI}}\,({\rm cm^{-2}}))=$ 13.99 $\pm$ 0.21. A total of 200 galaxy redshifts within the surrounding $6.8' \times 5.7'$ field of view of both quasars were identified in a R magnitude limited survey ($21.5 \leq {\rm R_{Vega}} \leq 24.5$) using the FORS2 spectrograph at the VLT. An upper-limit of $\xi_{\rm AG}=$ 1.83 was found when 194 Lyα absorber-galaxy pairs were binned in redshift space, in a bin of size $\Delta\sigma$ = 1.0, $\Delta\pi$ = 2.0 h_{70}^{-1}Mpc along the projected separation and line of sight distances respectively.

The upper-limit in the cross-correlation was found to be 3.2σ lower than the central peak in the galaxy auto-correlation, $\xi_{\rm GG}$, which was equal to 3.89 $\pm$ 0.65, and $\gg 5\sigma$ lower than the galaxy auto-correlation that was measured by the 2dFGRS and VVDS investigations. Thus we measured clustering amongst galaxies to be significantly stronger than the clustering between low column density Lyα absorbers and galaxies.

We then used GIMIC, a high-resolution hydrodynamical simulation, to re-create the cross-correlation. When binned with $\Delta\sigma$ = 1.0, $\Delta\pi$ = 2.0 h_{73}^{-1}Mpc in redshift space the simulations were consistent with the observations. No significant correla-

tion exists between galaxies and Lyα absorbers with $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))= 13-17$, and $\xi_{\rm AG}$ had a peak of 2.65 ± 0.78 at a redshift $z = 1.0$. The simulated $\xi_{\rm AG}$ was observed to only marginally increase when measured with absorbers of an increasing column density between $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))= 13-17$. Likewise $\xi_{\rm AG}$ increased by $< 1\sigma$ to 3.96 ± 1.21 when the cross-correlation was measured at $z = 0.5$. Hence in the models there is no significant evolution in $\xi_{\rm AG}$ with redshift.

$\xi_{\rm GG}$ from the GIMIC simulation was 27.05 ± 4.06 at $z = 1.0$, so when plotted in redshift space $\xi_{\rm GG}$ was again significantly greater than the cross-correlation.

Thus we reached a conclusion that the galaxies that inhabit the nodes and filaments of the dark matter cosmic web are embedded within, but not necessarily correlated with the low column density Lyα absorbers ($\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))\leq 17$) that loosely trace this filamentary structure. Hence the Lyα absorption lines in quasar spectra are predominantly caused by photons passing through this diffuse medium, and not because the sightline passes through a galaxy halo.

In a small side study, 47 C IV and 18 O VI absorption lines located in the UVES spectra of quasars HE 1122-1648 and PKS 1127-145 were used to calculate the oxygen and carbon metallicity of the Lyα absorbers. We found no evolution with redshift for either species. The mean carbon solar metallicity between $1.0 \leq z \leq 2.35$ was $[{\rm C/H}]= (-0.05\pm0.34)z-1.24\pm0.58$. There was a large scatter that varied from solar abundances to $[{\rm C/H}]{-3.41}^{+0.12}_{-0.09}$. The mean oxygen solar metallicity between $2.0 \leq z \leq 2.4$ was $[{\rm O/H}]= (1.01 \pm 1.54)z - 3.94 \pm 3.35$, again with a similar large scatter in the oxygen abundance.

These results are similar to previous findings by Schaye et al. (2003) and Aguirre et al. (2008), who studying C IV and O VI lines from 19 quasars at similar redshifts also found no significant evolution in the metallicity. All of these results indicate that many of these metals must have been expelled into the IGM at higher redshifts.

Contents

List of Figures

List of Tables

Declaration

The work in this thesis is based on research that was carried out between October 2005 and May 2009 while the author was a research student under the supervision of Prof. Simon Morris and Dr. Tom Theuns. The work was completed with the Extragalactic Astronomy and Cosmology Research Group at the Department of Physics, Durham University. No part of this thesis has already been submitted for assessment at Durham University, or has been submitted at any other university for any degree or qualification. It is all the author's own work unless otherwise stated in the text.

All figures that appear in this thesis were produced by the author except the following.

- Figure 1.3 is from Rauch (1998).
- Figure 1.7 is from Janknecht et al. (2006).
- Figure 1.8 is from Davé et al. (1999).
- Figure 2.2 is from Ryan-Weber (2006).
- Figure 2.3 is from Pierleoni et al. (2008).
- Figures 2.4 and 2.5 are from Wilman et al. (2007).
- Figure 3.3 is from Bertin (2006).
- Figure 3.6 is from Jehin and O'Brien (2007).
- Figures 4.1 and 4.2 are from Churchill (2007).
- Figure 5.13 is from Hawkins et al. (2003).
- Figure 5.19 is from Guzzo et al. (2008).

The galaxy surveys described in Chapters 3 and 8, and the UVES spectra used in Chapters 4 and 6 used data based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. ESO program numbers 076.A-0312, 381.A-0874, 082.A-0812, 083.A-0940, 067.A-0567, 068.A-0570 and 069.A-0371.

The STIS and FOS data were based on observations made with the NASA/ESA Hubble Space Telescope, and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA) and the Canadian Astronomy Data Centre (CADC/NRC/CSA). These included the HST science proposals 3483, 5950, 6577, 9173 and 9885.

Most of the work in this thesis was completed by the author. Those parts not done by him, or sections which relied heavily on contributions from others include the following.

- The model of Lyα absorbers in the IGM used in Section 1.3.4 is derived from Schaye (2001).
- Richard Wilman was the primary observer for the FORS2 galaxy survey and was responsible for the design of the masks.
- All of the raw data that was taken at Paranal, that included both the imaging and spectroscopy, underwent initial data reduction at ESO Quality Control at Garching, Germany.
- The programme used to reduce FORS2 spectra was written by D. Kelson, The Observatories, Carnegie Institute of Washington, California 91101 USA (Kelson, 2003).
- The software that helped reduce STIS and UVES spectra, UVES Popler was written by M. Murphy, Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia (Murphy et al., 2007).
- The package that fit Voigt Profiles to the quasar spectra was written by R.F.Carswell, J.K. Webb, M.J. Irwin, A.J. Cooke, Institute of Astronomy, Cambridge, CB3 0HA,Uk. (Carswell et al., 1991).
- The UVES line-lists of quasars HE 1122-1648 and PKS 1127-145 had already been fit, line-lists for these are described in Kim et al. (2002) and Ryabinkov et al. (2003) respectively.

- Data for the galaxy auto-correlation function from the 2dFGRS (Hawkins et al., 2003) used to generate Figure 5.16 was supplied by Dr. P. Norberg, Institute for Astronomy, University of Edinburgh.

The GIMIC simulation used in Chapter 7 is a project of the Virgo Consortium that primarily includes; the Institute for Computational Cosmology (ICC) at Durham University, the Max Planck Institute for Astrophysics (MPA) in Garching, Germany, and the Leiden Observatory, The Netherlands. GIMIC used the GADGET simulation code that was written by Volker Springel (Springel, 2005a). Additional software, that included the program SPECWIZARD, was written by Joop Schaye, Claudio Dalla Veccia, Craig Booth and Rob Wiersma (Leiden), Tom Theuns (ICC), Klaus Dolag (MPA) and Luca Tornatore (SISSA, Italy). The galaxy catalogue and magnitudes from the GIMIC simulation were produced by Rob Crain, Swinburne University of Technology, Victoria 3122, Australia.

Several figures, even though produced by the author, used data that was from other sources not already acknowledged. These were:

- The plot in Figure 1.1 is from the CUBA simulation that is described by Haardt and Madau (2001).
- The evolution of the optical depth in Figure 1.6 uses an equation from Kirkman et al. (2005).
- Data for; the quantum efficiency of the FORS2 chips in Figure 3.1, for the transmission of the R and B Bessel filters (Figure 3.10), and for the 200I grism (Figure 3.5) is from the ESO web-pages, http://www.eso.org/ (Jehin and O'Brien, 2007).
- SEDs for the standard stars that are in Figure 3.7 were made by Hamuy et al. (1994), Oke (1990).
- The template for an absorption line galaxy in Figure 3.9 was made by Kurtz and Mink (1998).
- SEDs for the standard galaxy types in Figure 3.11 have been based on models made by Bruzual and Charlot (2003).
- The ion fractions of C IV, O VI and H I in Figures 6.1, 6.2 and 6.3 were made

using the CLOUDY simulation (version 07.02) (Ferland et al., 1998).

- Colour data from the DEEP2 Survey Data Release 3, that is plotted in Figure 8.1, was produced by the DEEP2 survey team (Coil et al., 2004, Davis et al., 2003a).

Acknowledgements

My sincerest thanks go to my supervisor Simon Morris, for his help over the last three and a half years, his guidance at our weekly meetings, and his limitless patience when proof-reading the chapters of this thesis. I would also like to thank my secondary supervisor Tom Theuns who provided instruction when I investigated the 2D 2-point correlation function using the GIMIC simulation.

Supervisors aside, nothing could have been done without the help of post-docs. Before his arrival I was unaware of the software VPFIT or UVES popler, and understood next to nothing about the finer details of quasar absorption spectra. So thank-you Neil Crighton for your invaluable help in these matters. When collecting data for my galaxy survey in Chile it was actually Richard Wilman who showed me the ropes. So my thanks to Richard for being the one responsible for the original ESO survey proposal, acting as primary observer on this trip, and for all his assistance during the initial year of my work.

On the more technical side, this investigation could never have been completed without computer support from Alan Lotts and Lydia Heck, without whom we are all sure this department would cease to function.

Despite my seven plus years of learning physics and astronomy, many of the intricacies of a computer still elude me, so I would also like to acknowledge my office peers for their continual support in all matters relating to IT and astronomy. Thanks go to: Rob Crain, who wrote and ran many of the GIMIC codes I used; David Murphy and John Helly who know their computing; and to Mark Swinbank, who can reduce spectra in his sleep and taught me the wonders of IDL.

Everyday at 11.15 and four o' clock I would be served tea in the Bransden room with maybe a piece of shortbread, so I feel a toast to Pat, Glenda and Lynne is in

order.

...and I suppose I should give a shout out for the people that pay the bills. Despite their shortcomings in accountancy STFC played their part: they did fund my two sojourns to Chile, and no matter what people may say about expenses, the pool and arboretum at the Paranal Residencia were essential.

A special mention also needs to be made for those people from beyond the department. Here's to the residents of 23 Geoffrey Avenue: Jon, for proof reading various parts of this thesis, including the very text you are currently reading, and for providing me with dinner when I was threatened with malnutrition; to Sophy, for her persistent nagging that I keep to a rigorous programme of regular food and sleep during the last few months of writing up when teetering on the periphery of insanity; to my team-mates at the Elm Tree Pub Quiz for providing release and sanctuary on a weekly basis; and, since he requested this, thanks go to Callum for being Callum as I couldn't find a clause to include him, other than his lemon meringue pie and lunchtime camaraderie in the Bransden Room.

My heartfelt gratitude goes to Kat for her enduring affection during these months, her support when I was feeling down, and for introducing me to the art of pole-dancing.

Above all I wish to thank my family for helping me all these years, and in more ways than just financially and as a means of transport -although the cash was definitely appreciated in enabling my career as a perpetual student, as were the countless lifts to and from Leeds. Their emotional support knew no end, and they could always be counted upon to lend an attentive ear when I tried to explain what exactly it was that I studied. So yes cheers everyone, I owe you all a drink.

Chapter 1

An Introduction to Absorbers in the Inter-Galactic Medium

"So basically you're studying the holes in Polos...."

Dr. P.A. Jaquin, 2008

1.1 Introduction

On a clear night the sky may seem full of stars, but for the information of Dr. Jaquin the holes in Polos are literally teeming with particles when compared to the near perfect vacuum of space, and particularly to the voids of the inter-galactic medium (IGM).

One of the first facts taught by my Physics Master was that the mean density of the universe was approximately 1 particle of hydrogen m^{-3}. That we exist is a sure sign that these particles are not evenly distributed throughout the universe. Despite the fact that we observe most visible matter in galaxies, approximately 90% of the baryons in the universe at present exist outside the galaxies in the IGM (Ryan-Weber, 2006). Under the influence of dark matter these have collapsed to form vast clouds that are part of a filamentary network.

Galaxies also follow the distribution of the dark matter and are embedded in the nodes and filaments of this cosmic web. The extent of association between the gas in these clouds and the galaxies will be the main subject of this investigation.

Chapter 1 will deal with the history of how the IGM was formed and how we use quasar absorption spectroscopy to study it. In Chapter 2 we ask how the relationship between gas and galaxies can be quantified, we define the 2D 2-point correlation function and we review some results already published from the literature. The reduction of optical spectra from the FORS2 instrument that produced a galaxy catalogue is described in Chapter 3. Details of the ultra-violet (UV) spectroscopy of the quasars that led to the creation of an absorber line-list are given in Chapter 4. Chapter 5 shows our result from the absorber-galaxy 2D 2-point correlation function, where a comparison has also been made to the galaxy auto-correlation function. Traces of carbon and oxygen were detected in the IGM. These were found by the transition lines of C IV and O VI in the quasar spectra. The abundance of these metals[1] is measured in Chapter 6. The GIMIC simulation (Crain et al., 2009) was then used to measure the 2D 2-point correlation function. The methods used and results that were obtained are contained in Chapter 7. Our final discussion regarding the results of the cross-correlation, along with conclusions and avenues of further work, has been written in Chapter 8.

1.2 The Cosmography

The following is a short and incomplete introduction to modern cosmology in order to understand the universe in which we live and explain the origin of the equations and terms that are used in this thesis.

1.2.1 Homogeneity and Isotropy in an Expanding Universe

The first step when investigating the origin, current state and fate of the universe is to accept two simple but far reaching statements that are based on the Cosmological principle. Firstly that the universe is homogeneous, on large scales it is observed to be the same at every point. Secondly it is isotropic, meaning the universe looks the same in all directions. Local observations of galaxies obviously refute this idea, but on a cosmic scale the universe has been shown to be remarkably evenly distributed. This is especially so since the advent of major galaxy surveys such as the Two-degree

---

[1] In astronomy a metal is any ion species that has a proton number greater than Helium.

Field Galaxy Redshift Survey (2dFGRS) (Colless, 1999).

Observations of local galaxies led Hubble (Hubble, 1929) to prove that we live in an expanding universe where the velocity of recession of a galaxy is proportional to its distance. This time dependent rate of expansion is the Hubble parameter denoted by $H(t)$. Common practice since the current value of $H(t = 0)$ is a matter of debate, is to use a dimensionless parameter h to scale the value. Therefore at the present epoch, denoted by a "0" subscript $H_0 = 100h\ \mathrm{km\ s^{-1} Mpc^{-1}}$. The latest observations made by the WMAP probe measure a value for h of $h = 0.719 \pm 0.026$ (Hinshaw et al., 2009).

1.2.2 Scale Factor and Redshift in an Expanding Universe

It is space itself that is expanding and this led astronomers to define the expansion factor $a(t)$, between a physical coordinate $\bar{r}$ and the co-moving coordinate $\bar{x}$.

$$\bar{r} = \frac{a(t)}{a_0}\bar{x} \tag{1.1}$$

a_0 is the expansion factor at the time of observation and is set to 1. The Hubble parameter is equal to $\dot{a}/a$ where the dot denotes the derivative with respect to time. A photon emitted from a source at a time t will therefore have a stretched wavelength by the time it is observed at a time t_0. Astronomers define this ratio of the observed to the emitted wavelength as the redshift z where,

$$1 + z \equiv \frac{\lambda_{obs}}{\lambda_{em}} = \frac{a_0}{a} = \frac{1}{a} \tag{1.2}$$

The redshift is also equal to the recessional velocity of a galaxy as a fraction of the speed of light (the relativistic limit has a slightly different Doppler formula).

1.2.3 The Friedman Equation and Parameterisation of the Universe

If Hubble's Law describes the rate of expansion then the Friedman equation describes how the universe expands. By adopting the Einstein field equations Friedman (Friedman, 1922) was able to equate (assuming the universe was isotropic) the rate of expansion of the universe to its total energy-matter density ρ, where both radiation and matter are considered to act as an ideal fluid (Liddle, 1999),

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

k is a constant and defines the curvature and geometry of the universe.

The density ρ can be split between "dust" which comprises all non-relativistic matter and radiation. The difference between the two is the pressure that they exert. For a fluid pressure P the equation of state that relates the pressure, density and expansion is,

$$\dot{\rho} + 3\frac{\dot{\mathrm{a}}}{\mathrm{a}}\left(\rho + \frac{\mathrm{P}}{\mathrm{c}^2}\right) = 0 \tag{1.4}$$

The pressure P is proportional to the density $\mathrm{P} = \mathrm{w}\rho\mathrm{c}^2$. For dust $\mathrm{w} = 0$ and so by equation (1.4) the matter density decreases as $\rho \propto \mathrm{a}^{-3}$. This makes sense because the particle density is diluted as the universe expands along 3 dimensions. For radiation $\mathrm{w} = 1/3$ and so radiation density decreases as $\rho \propto \mathrm{a}^{-4}$. The extra -1 in the exponent is because the energy of a photon decreases as the wavelength is stretched.

If $\mathrm{k} = 0$ in equation (1.3) then the matter-energy density has the critical value ρ_c.

$$\rho_\mathrm{c} = \frac{3\mathrm{H}^2}{8\pi \mathrm{G}} \tag{1.5}$$

The total matter-energy density present is expressed as a fraction of this critical density and is denoted by Ω.

$$\Omega(\mathrm{t}) = \frac{\rho(\mathrm{t})}{\rho_\mathrm{c}} \tag{1.6}$$

With Ω_m and Ω_γ each contributing towards the total Ω. The curvature and fate of the universe then depends on the values of Ω and k.

- $0 \leq \Omega < 1$, $\mathrm{k} < 0$. In this scenario the universe is underdense and has an open geometry.

- $\Omega > 1$, $\mathrm{k} > 0$. In this case the universe is overdense and will eventually contract. This has a closed geometry.

- $\Omega = 1, \mathrm{k} = 0$. This is the flat universe that is described by Euclidean geometry.

k is known to be so close to zero that this is taken to be the case with $\Omega = 1$. Nucleosynthesis, a theory we understand very well has predicted a value of $|\Omega - 1| \leq 10^{-28}$ (Liddle, 1999) at the time of the big-bang.

The result that we live in an Euclidean universe that is slowing down and will eventually stop at an infinite time would appear to be the case were it not for the evidence (such as distant type 1a supernovae that were used as standard candles (Perlmutter et al., 1999)) that the expansion of the universe is accelerating.

To account for this a third term $+\Lambda c^2/3$ was added on the right hand side of equation (1.3). This is the cosmological constant that arises from the negative energy density of a vacuum. So a third Ω component is required, Ω_Λ. Unlike dust and radiation Λ does not evolve with redshift.

The three Ωs

Therefore there are three components that contribute to Ω such that $\Sigma\Omega_\mathrm{i} = 1$.

- Ω_m includes those particles that posses mass and produce a gravitational field. Included in this are the baryons and in particular all the H I that pervades the IGM and is the main subject of this thesis. However all the baryons in the universe only make up about 16 – 24% of Ω_m and are denoted Ω_b.

  Even when the protons in the IGM are considered, the quantity of which was estimated using quasar absorption line studies (see Section 1.4 below), together with the visible baryons that make up the stars and galaxies, the mass is not sufficient to make up Ω_m. This was further evidence on top of the galaxy rotation curves (Sofue and Rubin, 2001) and the dynamics of galaxies in clusters (Zwicky, 1933) for the existence of dark matter. This is matter that we know exists because of its effect on the gravitational field. However it cannot be directly observed because of its weak/lack of interaction with electromagnetism.

  There are three main candidates for what this missing mass may be. The first are grouped under MACHOS (Massive Compact Halo Objects) and include black holes, brown-dwarfs and neutron stars. However the amount of baryons available in the universe is insufficient for the number of halo objects that would be required.

The second are the neutrinos that would have been formed during weak interactions in the early universe. These were a possibility because unlike baryonic matter they have a very low cross-section and this explained why there had been no direct detection for dark matter so far. However these too fall short of the mass required (Goobar et al., 2006).

A further problem with neutrinos is that when matter decoupled from radiation at the last scattering surface at $z \sim 1000$ it could either have a relativistic speed and be described as "hot", or travel much slower and be "cold". However in the early universe hot dark matter would never have collected in a manner that formed the large-scale structure we see today. That is why the most likely dark matter candidates are WIMPS (Weakly Interactive Massive Particles). This is cold dark matter, one that is non-baryonic, has a small cross-section and is susceptible to the weak force. The search is on for what dark matter is with investigations that include ZEPLIN-III at the Boulby mine (Lebedenko et al., 2008) and the Large-Hadron-Collider (LHC).

- Ω_γ. The energy density of radiation is proportional to $(1+z)^4$. This greater decline means that for the observational universe, even at high redshift ($z \leq 6$) Ω_γ can be ignored because the universe is matter and not radiation dominated. Ω_γ is now $2.47 \times 10^{-5}\,h^{-2}$ (Liddle, 1999).

- The cosmological constant that ensures ρ is of a critical density is made up of dark energy. Its contribution to Ω, Ω_Λ is by far the most significant and yet the least understood as to its nature.

With cold dark matter and dark energy dominating the value for Ω this model is called the ΛCDM universe and numerous observations have been carried out to constrain the values of these parameters. The latest results gathered using the WMAP probe (Hinshaw et al., 2009) are $\Omega_\Lambda = 0.742 \pm 0.030$ and $\Omega_m = 0.258 \pm 0.030$.

1.2.4 The Cosmic Microwave Background and Growth of Structure

Two problems in particular arise by insisting on homogeneity and isotropy on large scales. First of all if we measure the temperature towards two opposing points in the universe a black body temperature of $T = 2.728 \pm 0.004$ (Liddle, 1999) is observed. This is the cosmic microwave background radiation that is left over from the big-bang (Penzias and Wilson, 1965).

That this radiation appears isotropic, even in areas that could never have come under causal contact means that in the past all of the observable universe must have been in a state of thermal equilibrium. This is called the horizon problem. However if the universe is to be in equilibrium at an early point in time what was the initial seed that caused the growth of large scale structure?

Both of these problems are solved by Alan Guth's theory of inflation (Guth, 1981). Within the first 10^{-34} seconds after the big-bang the universe experienced a rapid period of growth during which the acceleration increased exponentially. This solves the horizon problem because small areas that were originally in thermal equilibrium drew apart at such a rate during inflation that by now light can still not traverse from one place to another within the Hubble time.

This background radiation was originally mapped using the COBE satellite (Smoot et al., 1992). Once the foreground sources had been removed the remaining image was a map of the universe at the last scattering surface at the epoch of re-combination. The measurement of the temperature and density of the photons at this point was shown to be incredibly uniform with perturbations of only 1 part in 10^5.

That these anisotropies exist is vital, as it is at these original overdensities that dark matter collected and collapsed to form the large scale structure of the universe. The original seed for these anomalies can also be explained by inflation. Quantum fluctuations in the original universe were captured at the moment of inflation before they could disappear back into the vacuum. These perturbations were then stretched to form irregularities in what was otherwise a homogeneous medium. These perturbations then evolved by gravitational instability to create the cosmic web of galaxies and gas that we observe today.

1.2.5 Observing in a ΛCDM Universe

It is within this universe that we must measure distances to galaxies and estimate the age of the universe when the light from these galaxies was originally emitted.

If we define the function E(z), where we have parameterised k in terms of the other contributors to Ω (Hogg, 1999), then the observables that we require are functions of this.

$$E(z) = \sqrt{\Omega_m (1+z)^3 + \Omega_\gamma (1+z)^4 + \Omega_\Lambda + (1 - \Omega_m - \Omega_\gamma - \Omega_\Lambda)(1+z)^2} \quad (1.7)$$

The Hubble parameter at a redshift z' is equal to:

$$H(z') = H_0 E(z') \quad (1.8)$$

The co-moving distance D_C to this source is equal to (Hogg, 1999),

$$D_C(z') = \frac{c}{H_0} \int_0^{z'} \frac{dz}{E(z)} \quad (1.9)$$

While the look back time to an object at this redshift is (Hogg, 1999)

$$t(z') = \frac{1}{H_0} \int_0^{z'} \frac{dz}{E(z)(1+z)} \quad (1.10)$$

Throughout this investigation we adopted a standard ΛCDM cosmology of $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 100\,h$ km s^{-1}Mpc^{-1} with $h = 0.7$.

1.3 Early History, Formation and Evolution of Absorbers in the Inter-Galactic Medium

1.3.1 Recombination and Reionisation

The most convenient way of studying the history of the IGM is by tracing its temperature since the Big Bang. This can be summarised by describing the two fundamental points at which a change in the temperature of the IGM changed its nature, the first is the epoch of recombination.

During the first $\sim$ 400000 years of the life of the universe the "primordial-soup" consisted of ionised hydrogen, helium-4 (with a baryonic mass-fraction of 0.24 (Schaye, 2001)) amidst a sea of electrons and photons. The universe cooled as it expanded and at a redshift of $z \sim 1000$ at a temperature $\sim$ 3000 K (Liddle, 1999) the electrons decoupled from the photon temperature and formed atoms with the nuclei. It was at this point that the photons, which up till now had a very short mean-free-path because of Thomson scattering, were able to propagate freely and the universe became transparent. This was the epoch of recombination that has been captured in the map of the Cosmic-Microwave-Background (CMB).

After recombination the universe continued to cool adiabatically for up to $\sim 10^6$ years until effectively all the nuclei had formed atoms. By $z \sim 200$ (Barkana and Loeb, 2007) the now cool gas condensed and collected under gravitational collapse into the "haloes" formed by dark matter.

While it collapsed the gas could continue to radiate heat to the surrounding medium and so collapsed further. At this point the density became sufficient to form the first stars. This second fundamental change was the epoch of reionisation at $z \sim 11$ (Hinshaw et al., 2009). These first ever population III type stars were large with a mass $\gtrsim 100M_{\odot}$. These quickly spent their fuel which only consisted of hydrogen within 2-3 million years expelling large amounts of UV radiation in the process (Bromm and Larson, 2004, Yoshida et al., 2006). Up till now these stars have only been realised in simulations, so identifying these sources that first reionised the universe is one of the key investigations of the future James Webb Space Telescope.

These stars originated in the regions with the highest density. Therefore it was these denser regions that were the first to be reionised from the large expanding bubbles of UV radiation. The bubbles expanded until the pockets were overlapping with the result that by a redshift of z ~ 5 the IGM had become an ionised plasma.

1.3.2 The Photo-Ionisation Rate - Γ

The IGM continues to cool but the background UV radiation from sources such as QSOs, galaxies and AGN ensure a minimum plasma temperature of $T \sim 10^4$ K (Haardt and Madau, 1996). This UV background ($J(\nu, z)$) heats the IGM by photo-ionisation at a rate of $\Gamma(z)$ where the photons have an energy of $h\nu \geq 13.6$eV (the energy required to ionise H I from the ground state). The photons that are of a frequency below this ($\leq \nu_{th}$) contribute to the thermal energy of the IGM. With a cross-section of $\sigma(\nu)$, Γ is (Schaye, 2001),

$$\Gamma(z) = \int_{\nu_{th}}^{\infty} \frac{4\pi J(\nu, z)\sigma(\nu)}{h\nu} d\nu \tag{1.11}$$

The simulation program CUBA (Haardt and Madau, 2001) modelled the background UV intensity, taking into account the most significant contribution from both quasars and galaxies. Using this value for $J(\nu, z)$ the rate of photo-ionisation $\Gamma(z)$ as a function of redshift was found. This was parameterised as:

$$\Gamma(z) = \frac{61.6 \times 10^{-13} \times e^{2.39\,z}}{45.6 + e^{2.64\,z}} \times (1+z)^{-0.647} \tag{1.12}$$

This function is plotted in Figure 1.1.

Therefore the neutral fraction of the IGM has changed with time so that at present the diffuse plasma in the IGM is highly ionised with approximately only 1 in 10^4 neutral particles. However this rate of ionisation has decreased significantly since z ~ 2 with a fall caused by the drop in the quasar population. The density of star forming galaxies which also contribute to the level of background UV radiation also falls at this epoch (Hopkins, 2004).

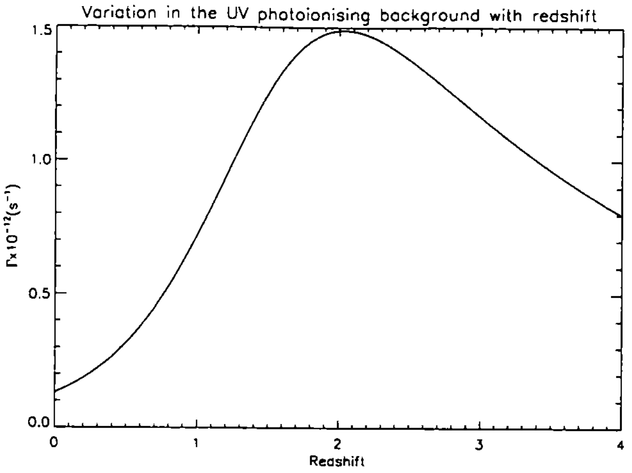


Figure 1.1 The variation in the value for the photoionisation rate Γ(z) with redshift. This breaks at $z \sim 1.7$ to local redshift as the UV background decreases with the lower quasar population and lower rate of star formation. The values for Γ(z) was calculated by the CUBA simulation that is described in Haardt and Madau (2001).

1.3.3 Absorbers in the Inter-Galactic Medium

The four phases of the IGM

At redshifts $z \geq 2$ the vast majority of baryons in the IGM existed as a diffuse plasma at a temperature of $T \sim 10^4$ K (Davé et al., 1999). Since that time structural collapse, Hubble flow, shock-heating and feedback from galaxies have perturbed this gas into a complex multi-phase IGM.

At the current epoch only 6% of Ω_b make up the stars, galaxies and other stellar remnants. Within the galaxies, cold ($T \lesssim 10^4$ K), condensed gas also contributes a further $\sim$ 1.7% (Bregman, 2007). This gas was observed in the H I Parkes All-Sky-Survey (HIPASS) that detected 21 cm radiation in the near universe (Zwaan et al., 2003).

The hot gas component which is at a temperature of $T \geq 10^7$ K is bound in galaxy clusters. The gas is shock heated as it enters the cluster and the temperature is increased to the virial temperature. These baryons are observed using x-ray emission lines, that are created in the hot gas that surround elliptical galaxies using the Chandra and XMM Observatories (Mathews and Brighenti, 2003).

Thus only 29 ± 4% of the baryons are now located in the diffuse IGM at temperatures T $\sim 10^4 - 10^5$ K (Sembach et al., 2004). This can be seen in QSO absorption spectra and will be the gas that is studied in this investigation. This gas is thought to lie along the filaments and in pancakes that trace large scale structure and in the voids of the IGM. So this leaves at least 60% of the baryons still unaccounted for.

Gas which is only $10 - 100$ times the mean cosmic density is more susceptible to shock heating. Shock waves that are caused by gravitational collapse travel along the filamentary structure heating the gas to T $= 10^5 - 10^7$ K. These baryons form part of the diffuse warm-hot IGM (WHIM) (Cen and Ostriker, 1999). However this gas remains elusive to observation as it is not significant in either emission or absorption. This field of research has sparked major recent interest as it is believed that up to half of the baryon deficit may be located here. The main observable signatures of the WHIM are the O VI doublet that is thought to trace this gas (Richter et al., 2004), thermally broad Lyα lines in UV and optical spectra (Prause et al., 2007), and the soft X-ray lines of O VII and O VIII at 0.57 and 0.65 keV respectively (Fang et al., 2003). The detection and study of the WHIM using these oxygen lines is a major science case for the future space instruments XEUS (the X-Ray Evolving Universe Spectrometer) and the new COS (Cosmic Origins Spectrograph) on the Hubble-Space Telescope (HST).

What is it that binds diffuse gas in the IGM?

Returning to the diffuse gas, one of the foremost questions when it was detected in the absorption lines of quasar spectra by Gunn and Peterson (1965) was what was its distribution? In the Gunn-Peterson model there is a smooth ionised IGM with pockets of gas, that even though still highly ionised contain about 1 in 10000 neutral particles that give rise to absorption lines. Questions arose as to what confined these clouds. Suggestions varied between pressure in a hot IGM (Sargent et al., 1980), gravity (Melott, 1980) and collapse into dark matter haloes (Rees, 1986) or a mixture of all three. The bound model developed with the idea that Lyα absorbers that are visible in QSO absorption spectra are formed in galaxy haloes. At redshifts lower than z ≤ 0.8 Lanzetta et al. (1995) found that Lyα absorbers were to be found out to 160 h^{-1}kpc from galaxies. This was confirmed by Chen et al. (1998) who also added that this correlation with galaxies grew with increasing absorption strength and decreasing impact parameter.

However this model contradicted the evidence where isolated Lyα absorbers were detected with no galaxy located close by (Morris et al., 1993, van Gorkom et al., 1996). Improved simulations and data from spectroscopy of a higher resolution supported this idea, and a theory was formed that Lyα absorbers arise because the neutral gas traces large scale structure. The correlation that was observed in Chen et al. (1998) between strong absorption lines and galaxies was just a consequence of galaxies lying in the more dense regions of the cosmic web.

It would be naive to accept just one side of these two contradicting models. Even though opinion has sided with the latter there is no definitive answer to this question. At $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 21$ the density of gas in an absorber would be comparable to that in the outer regions of a galaxy and this gas then forms tidal streams within the halo (Murali and Dubinski, 1999). There is then a continuous exchange of matter in the halo and the IGM as gas is accreted and then expelled from the galaxy by winds (Bertone and White, 2006, Kereš et al., 2005). Therefore at an impact parameter of $\sim$ 10 kpc it would be pointless to draw a line where a galaxy terminates and an absorber begins. Meanwhile those clouds with a column density that is several decades lower at $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\leq 14$ are more likely to make up the filaments and lie in the voids of the IGM. The location, column density and number density of these absorbers must also be a function of redshift because over time the gas drains along filaments and into haloes. The question as to the extent of the association between Lyα absorbers and galaxies which is the main topic of this thesis is discussed further in Chapter 2.

The dynamic absorbers

The gas in the IGM is far from static and is continuously acted upon by several processes.

- Hubble expansion. The expansion of the universe effects the absorbers directly by lowering the mean cosmic density, $\overline{\mathrm{n_H}} \propto (1+\mathrm{z})^3$. A lower density means that the characteristic size and number of absorbers changes since this

"Jeans size" (see Section 1.3.4 and equation (1.13) below) is set by an equilibrium between the mutual gravitational collapse and outward pressure from the temperature of the gas. The recombination rate of the ions is also proportional to the density. So a lower density will change the neutral fraction n_{HI}/n_{H}.

- Structural evolution. The gas in absorbers is undergoing a process of continuous gravitational collapse. After the epoch of recombination the gas started to fall into the dark matter haloes that define the large scale structure. The low density gas collects to form filaments and sheets and over time the gas drains along these filaments from regions of low to high density. Eventually these sheets may collapse to form smaller bodies, or the gas drains into the high density nodes of the cosmic web and gets trapped within a halo. This flow of gas may work in addition to or against the Hubble flow depending on the orientation of the filament to the observers line of sight.

- The photo-ionisation rate. The rate of recombination of hydrogen (labelled β_{HII} see Section 1.3.4) is proportional to the density and so is effected by Hubble expansion and structural evolution. As expansion is dominant β_{HII} is decreasing. However it is offset by the decrease in the rate of photo-ionisation Γ that has been observed since $z \sim 2$. Therefore the hydrogen neutral fraction is determined by the density and the UV background. This is relevant because it is only the neutral hydrogen gas that we can observe in quasar absorption lines.

- Feedback and metal enrichment. The feedback into the IGM not only heats the baryons to create the WHIM but it also enriches the universe with metals. The stars create metals such as carbon, silicon and oxygen that even at high redshift pervade the IGM and are visible in QSO absorption spectra. These metals were distributed in the early universe by powerful winds at $z \gtrsim 2$ (Aguirre et al., 2005, Porciani and Madau, 2005). For example Adelberger et al. (2003) showed that C IV systems at a redshift of $z \sim 3$ originated from super-winds travelling at 600 km s^{-1} that emanated from Lyman-Break-Galaxies (LBGs). Further details about the metallicity of the IGM and what this says about its connection with galaxies will be discussed in Section 2.2.2.

1.3.4 A Theoretical Model for Lyα Absorbers

When under the influence of so many competing forces it is not surprising that the nature of an absorber, even when stationary with no net gain or loss in mass will change with redshift.

What follows is a simple model in order to determine the physical properties of absorbers that exist in the Lyα forest between the redshifts of $z = 0 \rightarrow 4$. This model and the assumptions made are from Schaye (2001). What we would like to know is what is a typical size for these clouds and what is their density when compared to the surrounding mean density of the cosmos?

In order to use this model we have to assume that the gas is in hydrostatic equilibrium with negligible collisional excitation, as we only consider photo-ionised optically thin clouds. This is where a photon of wavelength $912.0 \leq \lambda \leq 1215.6701$ Å (the wavelength range of the Ly absorption series) can traverse the medium unimpeded except at the resonance absorption frequencies. The calculation for the neutral fraction of clouds with densities greater than this does not have an analytic solution because additional complexities, such as self-shielding, mean numerical simulations are required. The gas at all points in the derivation is assumed to be part of the diffuse IGM and not part of the WHIM.

By assuming hydrostatic equilibrium, the dynamical (free-fall) time ($t_{dyn} \equiv 1/\sqrt{G\rho}$) of the cloud can be approximated to the sound crossing time ($t_{sc} \equiv L/c_s$), for the speed of sound (c_s) through this medium that has a length L. When this occurs the characteristic size of the cloud is the Jeans length L_{Jeans}. Equating these time scales.

$$\begin{aligned} t_{sc} &= t_{dyn} \\ L_{Jeans}\left(\frac{\mu m_H}{\gamma k T}\right)^{\frac{1}{2}} &= \left(\frac{(1-Y)\,f_g}{G m_H n_H}\right)^{\frac{1}{2}} \qquad (1.13) \\ L_{Jeans} &= \left(\frac{\gamma K}{\mu m_H^2 G}\right)^{\frac{1}{2}} (1-Y)^{\frac{1}{2}}\, f_g^{\frac{1}{2}}\, T^{\frac{1}{2}}\, n_H^{-\frac{1}{2}} \end{aligned}$$

Where μ is the mean molecular weight of a fully ionised plasma, γ the specific heat capacity for a monoatomic gas, Y is the Helium baryonic mass fraction equal to 0.24 and f_g is the gaseous fraction of the mass $\sim \frac{\Omega_b}{\Omega_m}$. T is the temperature of the gas and has a value of 10000 K. The temperature of the Lyα absorbers are assumed to be of this order at the redshifts considered because if they were cooler then they would be heated by photo-ionisation. Temperatures derived from the widths of absorption lines also measure this value (Carswell et al., 1984).

The assumption that the size of an absorber is of the order of the Jeans length is not too crude for this model because if $t_{dyn} \gg t_{sc}$ then the cloud would expand and disperse before it could continue to collapse. Likewise if $t_{dyn} \ll t_{sc}$ then the cloud would not be stable. The pockets of gas with a greater density would cause the cloud to fragment. Schaye (2001) then defines a "Jeans density" equal to the particle density n_H times equation (1.13).

$$N_{H,J} = \left(\frac{\gamma k}{\mu m_H{}^2 G}\right)^{\frac{1}{2}} (1-Y)^{1/2} f_g{}^{\frac{1}{2}} T^{\frac{1}{2}} n_H{}^{\frac{1}{2}} \tag{1.14}$$

If we assume that the gas is optically thin and has a constant density then the fraction of neutral hydrogen is,

$$\frac{n_{HI}}{n_H} = n_e \frac{\beta_{HII}}{\Gamma(z)} = n_H \frac{(1-Y/2)}{1-Y} \frac{\beta_{HII}}{\Gamma(z)} \tag{1.15}$$

Where n_e is the number density of free electrons. The hydrogen recombination rate, which is proportional to the temperature has a value $\beta_{HII} \approx 4 \times 10^{-13} (T/10000\ K)^{-0.76}$. Our value for the redshift dependent photo-ionisation rate $\Gamma(z)$ was that modelled on results from the CUBA simulation (Haardt and Madau, 2001) and plotted in Figure 1.1.

Using the neutral fraction of equation (1.15) in equation (1.14) we can now compute the column density N_{HI} over the Jeans Length.

$$N_{HI} = \left(\frac{\gamma k}{\mu m_H{}^2 G}\right)^{\frac{1}{2}} \frac{1-Y/2}{(1-Y)^{1/2}} f_g{}^{\frac{1}{2}} T^{\frac{1}{2}} n_H{}^{\frac{3}{2}} \beta \Gamma(z)^{-1} \tag{1.16}$$

Rather than deal with the local density of the gas it is better to re-write equation

(1.16) using the mean baryon density of the universe $\overline{n_H}$ and the density contrast δ where $\delta = \left(n_H - \overline{n_H}\right)/\overline{n_H}$.

The particle density n_H as a function of the density contrast and redshift is:

$$n_H = \frac{3\Omega_b H_0^2}{8\pi G m_H}\left(1 - Y\right)\left(1 + z\right)^3\left(1 + \delta\right) \tag{1.17}$$

Using equation (1.17) in equation (1.16) produces the analytic solution to the column density of an optical thin Lyα absorber with density contrast δ.

$$N_{HI} = \left(\frac{3}{8\pi}\right)^{\frac{3}{2}}\left(\frac{\gamma k}{\mu m_H^5 G^4}\right)^{\frac{1}{2}}\left(1 - Y\right)\left(1 - \frac{Y}{2}\right) f_g^{\frac{1}{2}} T^{\frac{1}{2}} \beta \Gamma(z)^{-1}\left(1 + \delta\right)^{\frac{3}{2}}\left(1 + z\right)^{\frac{9}{2}}\left(\Omega_b H_0^2\right)^{\frac{3}{2}} \tag{1.18}$$

Therefore the column density depends heavily on the redshift with $(1 + z)^{4.5}$. This is not surprising given the statements made in Section 1.3.3 where the neutral fraction (hence observable column density) is set by the density and Γ, all of which are dependent on the redshift.

Using the stated cosmology from Section 1.2.5 with a value for $\Omega_b = 0.048$ we found the column density of Lyα absorbers for systems of a constant density contrast. These contours are plotted in Figure 1.2.

Points in Figure 1.2 where the column density is greater than $\log_{10}\left(N_{HI}\left(cm^{-2}\right)\right) \geq 17.3$ should be ignored as these absorbers have exceeded the limit for them to be considered optically thin and the above assumptions for the neutral fraction no longer apply. However it is interesting to note that systems where $\delta \geq 1000$ enter the optically thin case at more recent epochs.

Figure 1.2 shows how we must be careful when equating a column density with a particular physical density. For example at a redshift of $z = 0$ an absorber that creates a line that has a column density $\log_{10}\left(N_{HI}\left(cm^{-2}\right)\right) \sim 13$ has the same over-density as an absorber that creates a line of $\log_{10}\left(N_{HI}\left(cm^{-2}\right)\right) \sim 15$ at a redshift of 3.

The dependency of N_{HI} on Γ is also visible, for example the slow-down in evolution at $z \leq 2$ that corresponds to the sudden decrease in Γ that is observed in Figure 1.1. A lower value of Γ means a higher neutral fraction, hence greater neutral col-

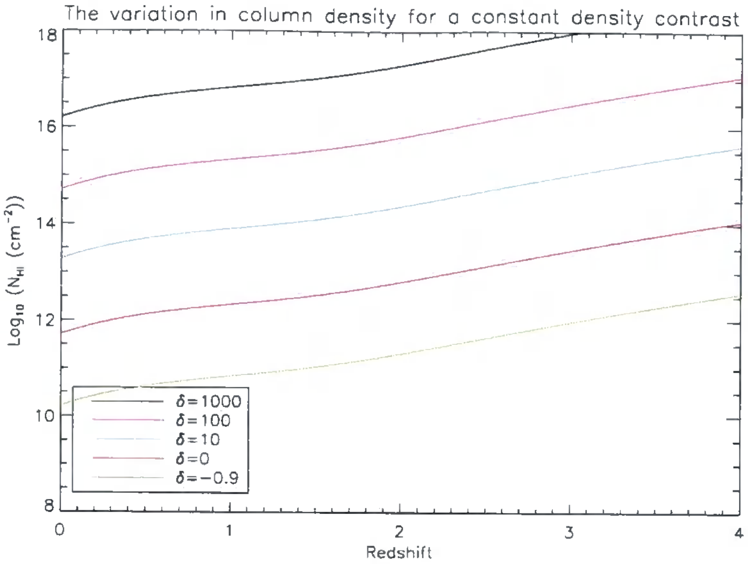


Figure 1.2 These plots show how the column density changes with redshift, the contours are of a constant density contrast δ ($\delta \equiv \frac{n_H - \overline{n}_H}{\overline{n}_H}$). Lines have been drawn at $\delta = 1000, 100, 10, 0$ and at $\delta = -0.9$ at 1/10 the mean density of the cosmos.

umn density, however this is balanced by the rarefaction of baryons as the universe expands.

Figure 1.2 also gives an indication as to the typical density of neutral hydrogen that produce the column densities of the absorbers observed in quasar spectra. The current detection limit of modern echelle spectrographs of $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 12$ corresponds to densities of the order of $\overline{n}_H$. However this is not always the case because when observing at a redshift of 3 an absorber with an under-density when compared to the cosmic mean is at the detection limit of an optical spectrograph.

We can also estimate the typical Jeans length of our model absorber based on the column density by rearranging equation 1.16 and using this in equation 1.13.

$$\mathrm{L_{Jeans}} = \left(\frac{\gamma \mathrm{K}}{\mu \mathrm{m_H^2 G}}\right)^{\frac{1}{2}} ((1-\mathrm{Y})(1-\mathrm{Y}/2))^{\frac{1}{3}} \mathrm{f_g^{\frac{2}{3}}} \mathrm{T^{\frac{2}{3}}} \left(\frac{\beta}{\mathrm{N_{HI}}\Gamma(\mathrm{z})}\right)^{\frac{1}{3}} \tag{1.19}$$

Which for an absorber with $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 14$ at a redshift of z = 3 is approximately 100 $\mathrm{h^{-1}kpc}$.

Therefore because of the dependency of $(\Gamma(\mathrm{z})\mathrm{N_{HI}})^{-1/3}$ both of which decrease with redshift, since a redshift of z = 4, the length for a Jeans absorber of a constant over-density has increased. This increase in length is necessary in order that the cloud remains in a stable equilibrium.

1.4 Observing QSO Absorption Lines

1.4.1 The Cause of Lyman Absorption Series in Quasar Spectra

The baryons in the IGM that we shall concentrate on in this investigation are those located in the diffuse phase at temperatures of $T \sim 10^4$ K. The neutral hydrogen, H I, that lies within clouds that are intersected by the line of sight towards a background quasar, can be studied by analysing the absorption lines found in the quasar spectrum. This phenomenon was predicted by and observed in Gunn and Peterson (1965). If this line of sight happens to pass through gas that is near to or is part of a galaxy then this too appears in the spectrum as an absorption line with a much higher equivalent width (Bachall and Salpeter, 1965).

It takes a photon with energy $E = h\nu \geq 13.6$ eV in order to photo-ionise a hydrogen atom from the ground state.

$$h\nu \; + \; H\,(1s) \rightarrow H^{+}$$

Therefore below 912 Å ($\nu \geq 3.3 \times 10^{15}$ s^{-1}) there is continuous absorption of the quasar flux. The most prominent absorption line is caused by the Lyα transition, this is where the hydrogen atom is excited from the 1s to 2p state.

$$h\nu \; + \; H\,(1s) \rightarrow H\,(2p)$$

When at rest this resonance line is found at $\lambda_{0\mathrm{Ly}\alpha}$ = 1215.6701 Å. When observed in a quasar spectrum what we see is a series of discrete absorption lines along the continuum at wavelengths that correspond to the redshift of the absorbers. This phenomenon is illustrated by the spectrum in Figure 1.3. This is the spectrum of QSO 1422+23 (z_{QSO} = 3.62) taken with the Keck HIRES (HIgh Resolution Echelle

Spectrometer) spectrograph (Rauch, 1998).

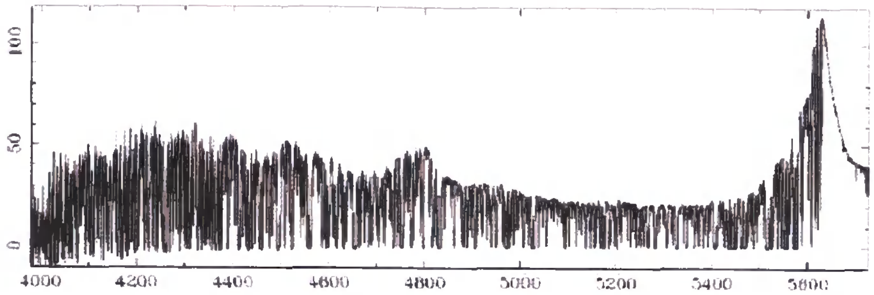


Figure 1.3 The spectrum of QSO1422+23 showing the Lyα Forest taken with the Keck HIRES with data from Womble et al. (1996). This figure is from Rauch (1998).

At the far right of Figure 1.3 is the Lyα emission line of the quasar and then at the preceding wavelengths absorption from the IGM. However not all of these lines are caused by the Lyα transition. The Lyβ emission line of the quasar is at 4736 Å corresponding to the H (1s) → H(3p) excitation. Consequently each Lyα absorption also has an associated Lyβ absorption line. This pattern repeats itself all the way through the Ly series. Therefore below 5616 Å is the Lyα forest, this becomes more convoluted below 4736 Å with the addition of the Lyβ forest and so on.

Mixed in with these lines there may also be absorption caused by metals, such as the various states of C, O, Mg, Si, Fe and Al. (For example there are strong Mg II absorbers that maybe found in the vicinity of galaxies (Churchill et al., 2005), or O VI that is believed to trace the warm-hot inter-galactic medium (WHIM) (Sembach et al., 2003).) Metals such as O VI and C IV are easy to identify because transitions from the ground state are doublets. This is because the single valence electron has two different energy levels caused by fine-structure splitting. So the metal absorption lines are of a known distance apart and will have predictable relative strengths.

At a sufficiently high column density the neutral gas in an absorber will become optically thick. This is where a photon with wavelength $912.0 \geq \lambda \geq 1215.6701$ Å is more likely to be absorbed at the Lyman resonance frequencies and this will produce an absorption line with a greater optical depth (see equation (1.27) below). Eventually the absorption line will become saturated as the probability of a photon in this wavelength range traversing the absorber tends to zero.

At a high enough column density ($\log_{10}(\mathrm{N_{HI}}(\mathrm{cm}^{-2})) \geq 17.3$ see section 1.4.3 below) it is almost certain that any photon that had a wavelength < 912 Å will have been absorbed. Therefore at wavelengths below this there will be a Lyman limit. This is where the quasar continuum drops to a negligible amount and the forest ends. For example consider a saturated Lyα absorption line that is centred on $\lambda = 2431.3402$ Å at a redshift $z = 1$. The Lyman limit in this quasar spectrum will occur at $\lambda = (z + 1) \times 912.0 = 1824$ Å.

1.4.2 The Shape of an Absorption Line

By studying the location, height, full-width-half-max (FWHM) and shape of a Lyα absorption line we can ascertain the redshift, neutral H I column density and an approximate temperature of the cloud.

The Heisenberg uncertainty relation dictates that the amount of energy required for excitation is not exact so an absorption line is of a finite width. The intensity profile I_ν is described by a Lorentz profile centred on ν_0. γ is the rate of spontaneous transition between energy levels and defines the natural line width. (Rybicki and Lightman, 1979).

$$\mathrm{I}(\nu) = \frac{\frac{\gamma}{4\pi^2}}{(\nu - \nu_0)^2 + \left(\frac{\gamma}{4\pi}\right)^2} \tag{1.20}$$

This predicted line-shape is different to what is observed because of several broadening processes. The most significant of these is the Doppler broadening that occurs because of the random motion of atoms within the cloud.

Consider an atom i which has a component velocity along the line of sight $\mathrm{v_z^i}$. This will cause a frequency shift in the absorption of $|\nu - \nu_0| = \mathrm{v_z^i}\nu_0/\mathrm{c}$. The velocity of these particles is proportional to their energy and so is described by a Boltzmann distribution.

When summed over the contributions from all $\mathrm{v_z^i}$ this causes a Doppler broadening of width $\Delta\nu_\mathrm{D}$ (Rybicki and Lightman, 1979).

$$\Delta\nu_\mathrm{D} = \frac{\nu_0}{\mathrm{c}}\sqrt{\frac{2\mathrm{kT}}{\mathrm{m}}} \tag{1.21}$$

Where the intensity profile is now described by,

$$\mathrm{I}(\nu) = \frac{1}{\Delta\nu_{\mathrm{D}}\sqrt{\pi}} \mathrm{e}^{-\left(\frac{\nu-\nu_0}{\Delta\nu_{\mathrm{D}}}\right)^2} \tag{1.22}$$

This curve is described by a Gaussian function with a profile that has a wider FWHM, a more pronounced peak but smaller wings than a Lorentzian. Additional contributions to this Gaussian will include turbulence within the cloud and bulk motion along the line of sight that will move the central frequency ν_0.

When this Gaussian is convolved with the Lorentz function in equation (1.20) the resultant curve can be described by a Voigt profile (Rybicki and Lightman, 1979).

$$\mathrm{I}(\nu) = \frac{\gamma}{4\sqrt{2}\pi^{\frac{5}{2}}} \sqrt{\frac{\mathrm{m}}{\mathrm{kT}}} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{\frac{-\mathrm{m}\mathrm{v}_{\mathrm{z}}^2}{2\mathrm{kT}}}}{\left(\nu - \nu_0 - \frac{\nu_0 \mathrm{v}_{\mathrm{z}}}{\mathrm{c}}\right)^2 + \left(\frac{\gamma}{4\pi}\right)^2} \, \mathrm{dv}_{\mathrm{z}} \tag{1.23}$$

A general Voigt function is written as (Rybicki and Lightman, 1979),

$$\mathrm{H}(\mathrm{a}, \mathrm{u}) \equiv \frac{\mathrm{a}}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{y}^2}}{\mathrm{a}^2 + (\mathrm{u} - \mathrm{y})^2} \, \mathrm{dy} \tag{1.24}$$

Therefore the intensity profile can be described as,

$$\mathrm{I}(\nu) = \frac{\mathrm{H}(\mathrm{a}, \mathrm{u})}{\Delta\nu_{\mathrm{D}}\pi^{\frac{1}{2}}} \tag{1.25}$$

With

$$\mathrm{a} \equiv \frac{\gamma}{4\pi\Delta\nu_{\mathrm{D}}} \qquad \mathrm{u} \equiv \frac{\nu - \nu_0}{\Delta\nu_{\mathrm{D}}}$$

Any collisional broadening has been ignored. Collisional broadening is proportional to the time between consecutive collisions. The low density of the IGM means that this is on a timescale of thousands of years and is negligible when compared to the frequency γ.

With so many factors that contribute to the Voigt profile there is no analytical solution to equation (1.25) that can derive the properties of an absorber. Instead the quasar spectrum is normalised to unity by dividing by the predicted quasar continuum and the amount of flux (W) absorbed as a function of frequency is measured.

As Angstroms have been used in this investigation the following equations will be a function of λ rather than ν.

$$\mathrm{W} = \int \left(1 - \frac{\mathrm{I}(\lambda)}{\mathrm{I}_0(\lambda)}\right) \mathrm{d}\lambda \tag{1.26}$$

Where $\mathrm{I}_0(\lambda)$ describes the SED of a quasar.

If the absorption line were to be transformed into a rectangle such that the area under the normalised continuum was the same as that integrated over the Voigt profile, then the width of this rectangle is the equivalent width (EW) measured in Å.

The amount of radiation absorbed when passing through a homogeneous medium of unit length S is proportional to the amount of radiation present. So for opacity coefficient $\kappa(\lambda)$ (Dyson and Williams, 1997).

$$\begin{aligned} \frac{d\mathrm{I}(\lambda)}{dS} &= -\kappa(\lambda)\mathrm{I}(\lambda) \\ \therefore \mathrm{I}(\lambda) &= \mathrm{I}_0(\lambda)\, \mathrm{e}^{-\tau(\lambda)} \end{aligned} \tag{1.27}$$

Where $\tau(\lambda) = \int \kappa(\lambda)\, \mathrm{dS}$ is the optical depth. Hence equation (1.26) now becomes,

$$\mathrm{EW} = \int \left(1 - \mathrm{e}^{-\tau(\lambda)}\right) \mathrm{d}\lambda \tag{1.28}$$

Rather than use the optical depth, it is customary to measure the neutral column density of the gas. By modelling the Thomson scattering within an absorber we get a value for the cross-section $\sigma(\lambda)$. The column density measured in cm^{-2} is then defined as $\mathrm{N} = \tau(\lambda)/\sigma(\lambda)$

As well as the column density of an absorber we also measure the FWHM of the profile. Both the natural and Doppler broadening terms are incorporated into a single measurement of the FWHM called the b-parameter.

$$\mathrm{b} = \frac{\mathrm{FWHM}}{2\sqrt{\ln 2}} \tag{1.29}$$

The value for the b-parameter therefore depends on whether the absorber is dominated by natural line broadening set by γ, or Doppler broadening set by $\Delta\nu_{\mathrm{D}}$. When Doppler broadening is dominant the value of b will be dependent on the temperature and any turbulence in the cloud and these terms are added in quadrature (Rauch, 1998).

$$\mathrm{b} = \sqrt{\frac{2\mathrm{kT}}{\mathrm{m}} + \mathrm{b}_{\mathrm{turb}}^2} \tag{1.30}$$

The b-parameter is related to the σ width of the Gaussian profile by $\mathrm{b} = \sqrt{2}\sigma$.

It is only when the column density exceeds $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 18$ that γ will dominate and the profile will have large damping wings.

1.4.3 The Curve of Growth

How we determine the column density of an absorber from the equivalent width depends on N because the relation is not linear. EW(N) is given by the curve of growth that is plotted in Figure 1.4 for values of column density that range from $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12 - 21$.

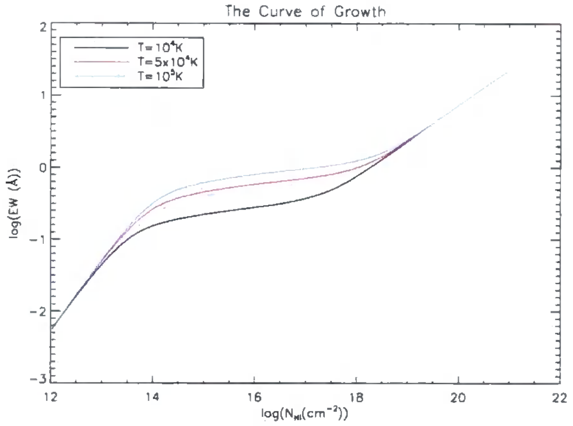


Figure 1.4 The EW dependency on the column density is not linear and can be divided into 3 regions. For unsaturated lines ($\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\lesssim 14.5$) the EW is directly proportional to N. When $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 14.5 - 18$ the EW is only weakly dependent on N. These profiles are still dominated by Gaussian broadening so the EW becomes dependent on the temperature of the absorber as the b-parameter is proportional to $\Delta\nu_{\mathrm{D}}$. Fiducial temperatures of $\mathrm{T} = 10^4$, 5×10^4 and 10^5 K have been plotted. Above $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 18$ the natural broadening now dominates and the width depends on the column density as $\mathrm{EW} \propto \mathrm{N}_{\mathrm{HI}}^{\frac{1}{2}}$.

Figure 1.5 shows a plot of Gaussian and Lorentzian dominated Voigt profiles. The

Voigt profiles were plotted assuming column densities of $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12-19$ and a temperature of 10^4 K. These illustrate the different regimes outlined in the curve of growth.

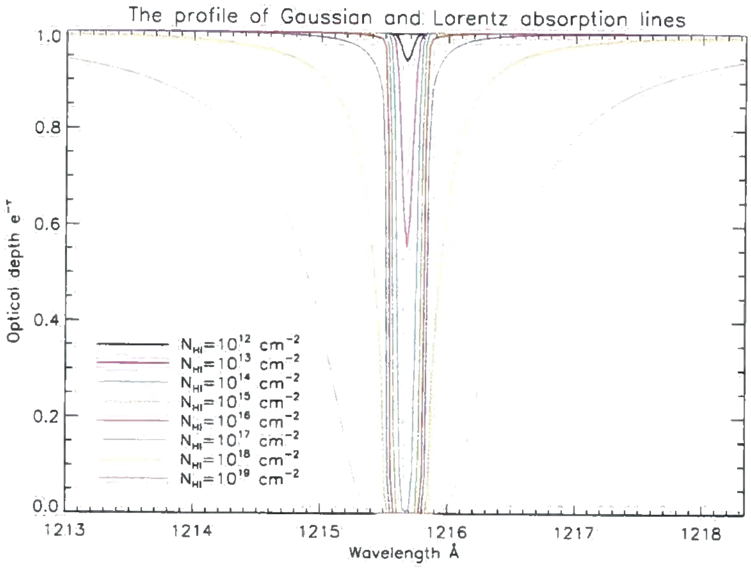


Figure 1.5 These 8 absorbers were all simulated using the same temperature of 10^4 K. They show the change in shape of a Voigt profile from a Gaussian to a Lorentzian as the column density increases from $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12$ (black line) to 19 (peach line).

Up to a column density of $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\sim 14.5$ the absorption line remains unsaturated. We now express $\tau(\lambda)$ as $\tau(\lambda) = \mathrm{N}\sigma_{\mathrm{T}}\Delta\lambda$ with Thomson scattering σ_{T} over a wavelength window of $\Delta\lambda$. For optically thin values of τ equation (1.28) is simplified to

$$\mathrm{EW} = \int \tau(\lambda)\, \mathrm{d}\lambda = \mathrm{N}\sigma_{\mathrm{T}}\Delta\lambda \tag{1.31}$$

Therefore EW is directly proportional to N and this gives the linear portion of the curve of growth in Figure 1.4. In this instance the profile is dominated by Doppler broadening and parameter ‘a’ from equation (1.25) is small. This case is illustrated by the black, pink and light blue profiles in Figure 1.5. Absorbers with a column density of $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12$ (black line) are very shallow and a spectrum would need to be of very high signal-to-noise in order to detect a line this weak. However with an increase of just 2 orders of magnitude (light blue line) the column density has reached the maximum before the absorption line becomes saturated. This shows that Lyα absorption is very sensitive to small amounts of neutral gas, as even a column density of $\log_{10}(N_{HI}(cm^{-2}))= 14.5 - 15$ can produce an absorber that will saturate an absorption line.

When the column density is $\log_{10}(N_{HI}(cm^{-2}))\sim 14.5 - 17$ the Gaussian profile becomes saturated and the equivalent width is weakly dependent on N, $W \propto (\log N)^{\frac{1}{2}}$ (Dyson and Williams, 1997). This is the flat region of the curve of growth and is why the green, red and dark blue profiles in Figure 1.5 are so close. By $\log_{10}(N_{HI}(cm^{-2}))= 17$ we begin to see the Lorentzian damping wings.

Of much more relevance in this regime is the FWHM. The Doppler broadening is still greater than the natural broadening, so by equation (1.21) the FWHM is dependent on the temperature of the absorber. This is apparent in Figure 1.4 where 3 different plots have been drawn assuming a temperature of $T = 10^4$, 5×10^4 and 10^5 K. This is why determining the column density of absorbers is so hard in this region. A minor error in the equivalent width or b-parameter produces an error that can stretch over an order of magnitude in the column density.

In Lyα systems where $\log_{10}(N_{HI}(cm^{-2}))\geq 18$, the natural broadening becomes the most significant as γ increases. A large value for parameter 'a' in equation (1.25) implies large Lorentzian wings and once again the width becomes predominantly dependent on the column density, $EW \propto N^{\frac{1}{2}}$. The orange and peach Voigt profiles in Figure 1.5 which have a column density of 10^{18} and 10^{19} cm^{-2} show this. The square-root dependency is observed in Figure 1.4 at the top end of the plot.

1.4.4 Extracting Information about the IGM from Quasar Spectra

Ever since the discovery of absorption lines in quasar spectra that are caused by gas in the IGM, spectrographs have been trained on these targets. During the intervening years there has been considerable progress in the signal-to-noise and spectral resolution of this data that has lead to major breakthroughs in the understanding of the IGM. Depending on the epoch under investigation and the resolution of the spectra, different techniques have been employed to extract the information available.

Low resolution spectroscopy

The first spectra were hampered by low resolution and a minimum detectable EW of $\sim$ 15Å (Rauch, 1998) so that all the lines were blended and the best that could be achieved (as was the original plan in Gunn and Peterson (1965)) was an estimate of the mean flux absorbed with redshift, the flux decrement $D_A(z)$ (Kirkman et al., 2007, Oke and Korycansky, 1982).

$$\begin{aligned} D_A(z) &= \langle 1 - \frac{f_{obs}}{f_{cont}} \rangle \\ &= \langle 1 - e^{-\tau} \rangle \\ &= 1 - e^{-\bar{\tau}(z)} \end{aligned} \tag{1.32}$$

The standard procedure is to find the mean flux absorbed between the Lyα and Lyβ emission lines which is then averaged over many sight lines.

This technique can of course be used with data of higher resolution as was done by Kirkman et al. (2007). By using 74 quasars observed using the Faint-Object-Spectrograph (FOS) during the QSO Absorption Line Survey Hubble Key project (Bechtold et al., 2002), $D_A(z)$ was measured between $0 \leq z \leq 1.6$. Additional surveys at higher redshifts using HIRES and the KAST spectrograph (Kirkman et al., 2005, Tytler et al., 2004) pushed this measurement up to a redshift of 3 with the following result. $D_A(z)$ was fit with two power laws with a break at $z_c = 1.10$.

$$D_A(z) = \begin{cases} 0.013\,(1+z)^{1.54} & z \leq z_c \\ \frac{0.013(1+z_c)^{1.54}}{(1+z_c)^{2.98}}\,(1+z)^{2.98} & z \geq z_c \end{cases} \tag{1.33}$$

The largest source of error when measuring the level of absorption is fitting the quasar continuum.

By assuming that the absorption properties of these clouds do not change over cosmic time we can now use $D_A(z)$ to estimate the evolution in the number density of absorbers with redshift. This follows a power law that measures the evolution in the line-density of absorbers above a minimum EW limit W_{min}.

$$\left.\frac{dN}{dz}\right|_{W_{min}} = \left(\frac{dN}{dz}\right)_0 (1+z)^{\gamma} \tag{1.34}$$

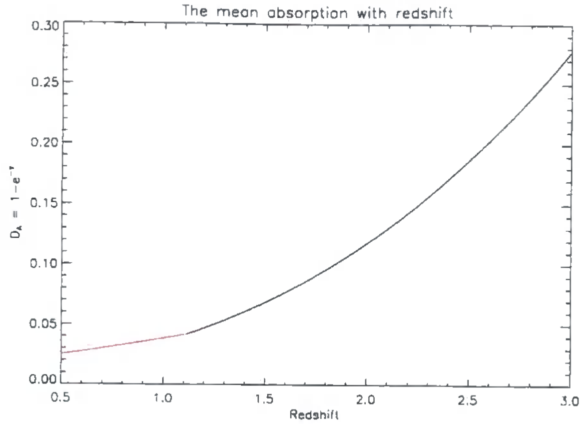


Figure 1.6 Using FOS spectra of quasars Kirkman et al. (2007) estimated the amount of absorption from the IGM up to a redshift of 1.6. It was found that 79% of the absorption is caused by the diffuse IGM, 12% from metals and 9% from strong H I lines. The redshift limit was increased using data taken from HIRES and KAST spectra (Kirkman et al., 2005, Tytler et al., 2004).

However the error associated with γ is large as a constant b-parameter and column density must be assumed.

Medium resolution spectroscopy

The situation is improved when the lines are of a sufficient resolution that they can be individually identified. It is now possible to do a line count to establish the number density of absorbers per range in equivalent width per unit redshift (Sargent et al., 1980, Young et al., 1982).

$$\frac{d^2N}{dz\,dW} = \frac{W_r}{W_*}e^{-\frac{W_r}{W_*}}(1+z)^{\gamma} \tag{1.35}$$

With a more reliable estimate for the number of absorbers per unit redshift a more accurate exponent for the line-density evolution can be found. The coefficient has been normalised by the parameter W_*. This could be the minimum detectable EW W_{min}. The minimum value depends on the resolution of the spectra, for example the Key Project FOS data of medium resolution that had $\Delta v \sim 230$km s^{-1} used a value of $W_* = 0.283 \pm 0.012$ (Davé et al., 1999).

Unfortunately, although the individual peaks of absorption lines may be visible the wings of the profiles will still be blended. Algorithms do exist that attempt to deconvolve the spectrum and predict the most likely shape for the absorption lines. However a standard value for the b-parameter needs to be chosen in order to constrain the column density.

High resolution spectroscopy

The only way to accurately measure attributes such as the b-parameter, line broadening, the EW and derive a column density is by Voigt profile fitting and this needs high resolution spectroscopy. Optical spectrographs that are capable of this include the Ultra Violet Echelle Spectrograph (UVES) that is based at the ESO Observatory and HIRES at the Keck. These are capable of a resolution of up to R $\sim$ 65000. The UV echelle spectrograph used in this investigation was the Space Telescope Imaging Spectrograph (STIS) that had a resolution of R $\sim$ 30000.

1.4.5 Line-Density Evolution

Numerous quasars have been studied in order to obtain a value for γ from equation (1.35). One of the first major studies at low redshift was completed by Weymann et al. (1998) that studied 63 FOS quasars up to a redshift of $z = 1.5$ from the HST Key Project. The evolution was found to be flat with $\gamma = 0.26 \pm 0.22$ with a minimum EW of 0.24 Å. This is in agreement with later works that studied larger FOS quasar samples from the HST archive. Dobrzycki et al. (2002) found $\gamma = 0.54 \pm 0.21$ that extended the redshift up to $z \sim 1.7$.

The results for γ that were taken from spectra at optical wavelengths (Bechtold, 1994, Lu et al., 1991) show a very different value. When observed at optical wavelengths γ was found to be significantly higher. For example Lu et al. (1991) found $\gamma = 2.75 \pm 0.29$, while Bechtold (1994) had a value of $\gamma = 1.85 \pm 0.27$.

Early extrapolations of the evolution based on optical spectra had predicted that no lines would be visible at all when observed in the ultra-violet. With the advent of UV spectroscopy from space telescopes the only way to fit these regimes was by two separate power-laws with a break at $\log(1 + z) = 0.4$ ($z \sim 1.7$). Unfortunately the change in the relationship of dN/dz occurred at the change over between UV and optical wavelengths.

This sharp break has been smoothed with the advent of echelle spectroscopy at wavelengths in the region of ~ 3000 Å. These spectrographs are able to produce a higher number of lines with a lower EW limit. Using 9 quasars that have been studied with STIS, UVES and Keck, Janknecht et al. (2006) found evidence for a smoother transition between the two redshift regimes.

Finding an exponent that will fit the evolution is made harder still because the decrease in line density may also be a function of column density. Janknecht et al. (2006) found tentative evidence for this when they measured $\gamma = 0.74 \pm 0.31$ for $\log_{10}(N_{\rm HI}\,(\rm cm^{-2})) = 12.9 - 14$ between a redshift of $0.7 \leq z \leq 1.9$. Increasing the minimum column density to $\log_{10}(N_{\rm HI}\,(\rm cm^{-2})) \geq 13.64$ caused γ to increase to $\gamma = 1.50 \pm 0.45$. Figure 1.7 which is from Janknecht et al. (2006) compares a collection of quasar absorption line surveys. The results from Janknecht et al. (2006) have been plotted with the solid line.

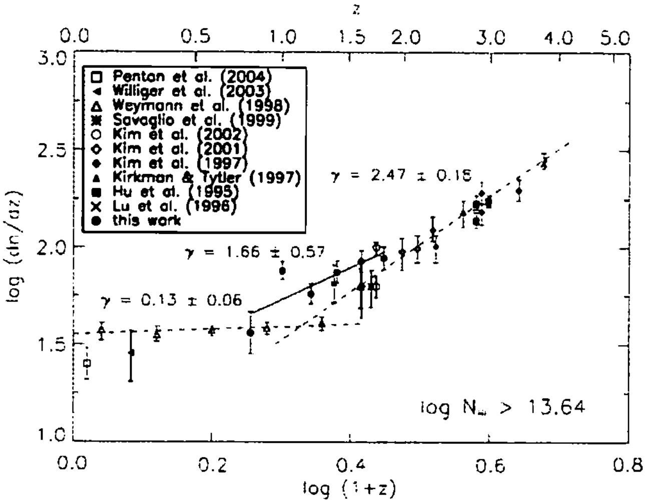


Figure 1.7 This figure from Janknecht et al. (2006) shows the line-density evolution from optical and ultra-violet surveys and shows the need for two separate power laws for γ. The echelle spectra used in Janknecht et al. (2006) were able to fit a smoother drop in the line evolution in the critical region at ~ 3000 Å.

There are 3 competing effects that cause the power-law in the line-density evolution. The most predominant effect which causes the decrease in density from high redshift to low is Hubble expansion (Davé et al., 1999).

As the universe expands the mean physical density falls. As the recombination rate is proportional to the density this causes the neutral fraction to decrease. Fewer H I atoms means there are fewer absorbers above a given equivalent width limit.

dN/dz would continue to fall were it not for the sudden drop in Γ that occurs at $z \lesssim 2$ (see Figure 1.1). By decreasing the rate of photo-ionisation this compensates for the drop in density and the evolution is now flat.

A third process is structural evolution (Davé et al., 1999). As gas drains from regions of low to high density the diffuse IGM, which is the most common means of absorption, is depleted. Meanwhile high density systems collapse under gravity but will also have feedback from galaxies causing the clouds to disperse. So the structure of absorbers at these two different epochs is different and this causes the line-density to decrease.

The same evolution in line density has also been observed in simulations (Davé et al., 1999, Theuns et al., 1998). In Davé et al. (1999) 400 artificial spectra were generated by firing random sightlines through the simulation at intervals of $\Delta z = 0.25$ between the redshifts of $0 \leq z \leq 3$. The UV background was modelled on the latest predictions from Haardt and Madau (1996).

Figure 1.8 which is from Davé et al. (1999) shows the line-density evolution taken from the simulation of the ΛCDM universe. The solid line shows the evolution in

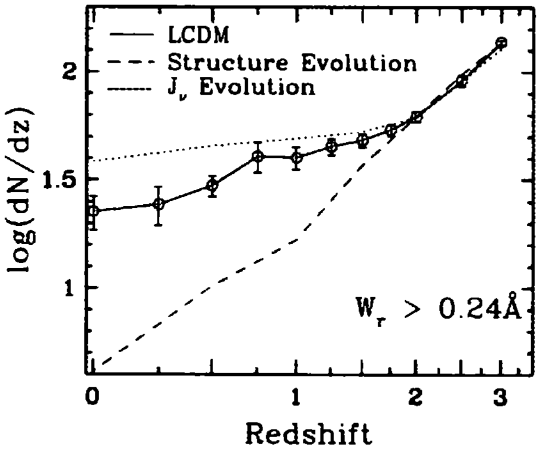


Figure 1.8 The solid line is of the line-density evolution from the simulation with the change in γ at $z \lesssim 1.7$. The finely dotted line is of the evolution with the structural growth that causes γ to increase switched off. The dashed line has a constant value for the photo-ionising background. Clearly the decrease in the photo-ionising rate $\Gamma(z)$ at $z \lesssim 2$ is the cause for the flat evolution. Figure is from Davé et al. (1999).

line-density from the simulation. The finely dotted line is for a universe with only evolution in $\Gamma(z)$ and Hubble flow. The dashed line maintains the same value for $\Gamma(z)$ from high redshift and includes structural growth. Figure 1.8 shows that the drop in the rate of photo-ionisation at $z \lesssim 2$ is the dominant process and the cause for the decrease in γ. The solid line from the simulation lies below this proving that structural growth does have a minor effect.

Now that we know about the diffuse gas in the IGM and how to detect it by the absorption lines, the next question is to what extent is this gas associated with galaxies. The principles behind how the IGM interacts with galaxies and how we can quantify this are discussed in Chapter 2.

Chapter 2

The Interaction of Galaxies with the Inter-Galactic Medium

2.1 The Question we Should be Asking

We now return to the two opposing straw-man models that were described in Section 1.3.3. Do quasar absorption lines arise because of the cosmic web of gas in the IGM, or are they created as the line of sight towards a quasar passes through the extended halo of an individual galaxy? This is our starting question because any investigation that seeks to address the IGM-galaxy relationship using absorption lines must know their origin.

By surveying galaxies within the same field of view as the sightline to a background quasar we can judge whether galaxy haloes and absorption lines occupy the same space. When doing this we determine two distances, the distance between the galaxy and the sightline, the projected separation (or impact parameter) σ, and the separation along the line of sight π. This is illustrated in Figure 2.1.

The question of haloes or cosmic web was investigated at low redshift by Morris et al. (1993). This was done by comparing Lyα absorbers that were located in the sightline towards quasar 3C 273 with galaxies found in a 17.2′ diameter field down to a limiting apparent magnitude of B $\approx$ 19. The results showed that there was no direct 1:1 association between any of the absorption lines and individual galaxy haloes, with the closest galaxy at an impact parameter of 160 kpc from the quasar sightline. There was even the case of a lone absorber at an impact parameter of more than 4.8 Mpc from the nearest galaxy with $M_B \leq -18$. However the distribution of

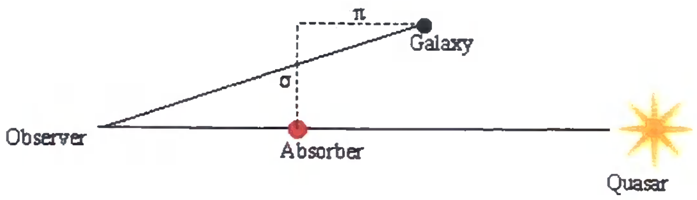


Figure 2.1 This diagram (not to scale) shows a single absorber-galaxy pair in a quasar field. The π and σ distances along the line of sight and projected separation respectively are what we measured to determine the absorber-galaxy 2D 2-point correlation function.

galaxies and absorbers was not found to be random, with both following the same distribution.

A follow-up paper that used the same line of sight (Mo and Morris, 1994) that considered a model of galaxy mini-haloes considered a compromise whereby only 20% of the Lyα absorbers could be attributed to galaxy haloes.

The opposite conclusion was reached by Lanzetta et al. (1995). They suggested that at redshifts z $\leq$ 1 luminous galaxies each have a gaseous halo that extend out to 160 h^{-1}kpc and it is within these that at least $32 \pm 10\%$ or as many as $60 \pm 19\%$ of the absorbers lie.

After many more studies and much conflicting evidence a compromise between these two models has been adopted. The advent of echelle spectroscopy and large simulations (Davé et al., 1999, Theuns et al., 1998) that can model the creation of absorption lines has shown that the contribution from the absorbers in the cosmic web should dominate for column densities that are typical of the diffuse IGM ($\log_{10}(N_{HI}(cm^{-2}))\sim 14$). However just because galaxies cannot be found within the region of an absorber does not mean that none exist. At best the galaxy surveys are complete down to a limiting magnitude and so the original question needs to be amended. What is the association between absorbers and galaxies down to a minimum galaxy luminosity?

The question should also include a statement as to the column density of the absorption line. There is strong evidence that those saturated lines, the damped Lyman-α systems (DLAs) that have $\log_{10}(N_{HI}(cm^{-2}))\geq 20.3$, are directly connected with a single galaxy or group of galaxies. For example the DLA system at z =

0.313 in the spectrum of PKS 1127-145 (Rao et al., 2003a) lies 6.5 kpc from the galaxy responsible. On the other hand these galaxies are embedded in a cosmic web structure. These filaments can be Mpc across and contain low column density systems that have a column density that are several orders of magnitude lower at $\log_{10}(N_{HI}\,(cm^{-2})) \sim 13$.

The final question cannot stop there, any association must also be dependent on the redshift for reasons stated in Section 1.3.3 regarding the evolution of absorbers. The galaxy population is evolving as well. The dwarf galaxies at a redshift $z \sim 3$ will interact in a different way than the evolved spirals in the present universe, not least because of the different methods of accretion and galactic winds that are discussed below.

The environment surrounding a galaxy could also have an impact on the measured association. The high temperature gas at $T \sim 10^7$ K in the intra-cluster medium (ICM) identified by the Chandra X-ray satellite (Mathews and Brighenti, 2003) means that the neutral hydrogen fraction will be very low and so the observable gas will be depleted. This would lead to an anti-correlation between absorption lines and galaxy clusters at small separations along σ and π.

Further investigations have suggested (Chen et al., 2005) that the correlation between absorbers and galaxies may also depend on whether the galaxies are dominated by emission or absorption spectral features. A luminous star-forming galaxy that is identified by strong [O II] emission lines requires a supply of gas that could be fed by an absorber. Therefore this would increase the correlation of emission type galaxies with absorbers. However this star formation would also expel winds and heat the surrounding IGM, thereby decreasing the presence of neutral gas and so lead to an anti-correlation.

So rather than just ask what is the association between the gas in the IGM and galaxies we should be asking;

> *"What is the correlation between the neutral gas that is observed in quasar absorption lines and galaxies brighter than a minimum luminosity as a function of redshift, column density, the distances σ and π and galaxy environment?"*

If the relationship between neutral gas and galaxies is better described by the theory of Lanzetta et al. (1995), then we would expect the correlation to be high at short separations of π and σ, and that this correlation would decay beyond the typical size of a galaxy halo ($\sim$ 100 kpc).

A theory which describes the neutral gas as being an entity of the large-scale-structure in the IGM (Morris et al., 1993) would show a low correlation at small distances, and this level would remain flat along both the π and σ separations.

2.2 The Inflow and Outflow of Matter from a Galaxy

The method by which matter from the IGM is accreted by galaxies and then expelled by galaxy winds, super novae and active-galactic nuclei (AGN) is of course a huge area of active research, the many details of which are too numerous to explain in detail here. This investigation will contribute to this topic, as it is hoped that by knowing how and to what extent the gas from the IGM interacts with galaxies we can form a better understanding of gas accretion, galaxy evolution and feedback processes that distribute metals back into the IGM.

What follows is a short summary of the accretion of gas by galaxies and their expulsion of gas and metals in feedback mechanisms. This is in order to understand how we arrive at a situation where we are able to detect material in the IGM that originated from galaxies.

2.2.1 Gas Accretion

Traditional hot accretion with models that appear in Fall and Efstathiou (1980), Rees and Ostriker (1977), White and Rees (1978) includes cold gas that falls through the potential of a dark halo. As this happens it is shock heated to the virial temperature, which for a Milky Way type galaxy is T $\sim 10^6$ K (Kereš et al., 2005). This gas then cools and condenses, and when it is no longer supported by pressure it falls onto the galactic disk and forms stars.

Once these stars have spent their fuel, their gas is then expelled into the interstellar medium (ISM) in a hot phase where a fraction of this mass is returned as metals. This gas with the metallic fraction then either cools and returns to form new stars, or is ejected by a galactic wind. Once again in the outer galactic halo and in a cool state, the gas then rains down back onto the galactic disk, or if the wind is of sufficient energy returns to the IGM. All of these processes have been well documented and are investigated using smoothed particle hydrodynamic simulations (SPH) and semi-analytic models such as the galaxy formation code GALFORM (Cole et al., 2000).

Using SPH simulations Kereš et al. (2005) have modelled a cold method for accretion. This is where gas of a lower temperature T $\sim 10^5$ K is drawn into a galaxy along filamentary structures, for example when the Milky Way feeds from the Magellanic Stream, this enables the galaxy to draw gas from a further distance than that achieved via hot accretion. Kereš et al. (2005) predict that this is the dominant method of accretion for galaxies of a lower mass, $M \leq 10^{10.3} M_{\odot}$.

2.2.2 Winds and Feedback

Observations of star-forming, gas rich galaxies in the local universe, through filters that target species such as [O II] and Hα as well as ranging from the radio to x-ray regimes of the electromagnetic spectrum have shown material being expelled to large distances from the galactic plane. This energy and matter is being carried by powerful galactic winds and this is how metals leave the ISM and contaminate the galactic halo and IGM.

The main questions astronomers ask are, what powers these winds, how energetic are they, and what effect do they have on the IGM and on the evolution of galaxies?

The two sources of energy that drive these winds are the supernovae explosions of Population II stars, that can be observed by tracing [O II] (Burke, 1968), and if it is active, the galactic nucleus. For example one of the first central engines that was detected was the one that powers the outflows of M82 (Lynds and Sandage, 1963).

Having to observe over many orders of magnitude in wavelength in intricate detail, and having two energy sources means that wind models are very hard to constrain using observations. Therefore the processes that govern winds are usually investigated using hydrodynamic simulations and semi-analytic models.

In the supernovae model energy is radiated away from the starburst region. This mixes with the ISM to produce a bubble of super-heated material that forms a momentum driven wind (Oppenheimer and Davé, 2006).

Winds that are powered by AGN get their energy from accretion. Such winds can be very powerful. For example winds with speeds of $\sim$ 0.03c have been observed emanating from broad absorption line quasars that are accreting matter at rates of up to $\sim$ 100 $M_{\odot}year^{-1}$ (Veilleux et al., 2005).

As the wind escapes it picks up contaminants from the ISM. These supply a drag factor that slows the wind down causing it to drop some of this load. Therefore we expect the heavy metals, such as Fe II and cold molecular gas to inhabit the immediate galactic halo. For example the spectral features of extended components from M82 that have been observed escaping at a velocity v $\sim$ 230 km s^{-1}. Meanwhile the lighter species, O VI, was observed in the galaxy corona at an escaping velocity of v $\sim$ 525 – 655 km s^{-1} (Veilleux et al., 2005). It is determining this drag factor, the efficiency of the radiation and size of the energy supply that makes it hard to predict what distance these winds can reach, and what fraction of the metals enter the IGM at distances $\sim$ 100 kpc and beyond.

The winds that are observed at low redshifts are expected to differ to those predicted and detected at high redshift. Adelberger et al. (2003) discovered a lower than average concentration of H I within 0.5 h^{-1}Mpc of a sample of Lyman-Break galaxies (LBGs). This was evidence for super-winds (v $\sim$ 600 km s^{-1}) that escape from LBGs (Pettini et al., 2001) that have high rates of star formation. When using an extended dataset Adelberger et al. (2005) found that this H I deficit disappeared and so ruled out there being a strong wind from every LBG. However in 1/3 of the galaxies evidence remained that winds were expelling a fraction of the neutral gas.

Adelberger et al. (2005) were also able to detect C IV and O VI absorption systems at distances of $\sim$ 80 and $\sim$ 400 h^{-1}kpc from the nearest known LBG. Either these metals were products of these strong winds, or they are a relic from an earlier epoch during the initial distribution of primordial metals. The best tool available to answer this question is by studying metal lines in quasar absorption spectra and finding out to what extent these metal systems are a feature of galactic haloes.

2.3 The Metal lines in Quasar Spectra

2.3.1 Measuring the Metallicity

To adapt to the limitations of spectrographs there have been various methods to compute the metallicity of absorbers. The problem was that for a particular signal-to-noise and minimum detectable equivalent width most metal lines were too insignificant. As they have a much lower density, the optical depth of metal absorptions is significantly lower than that of Lyα lines. The b-parameter for these Gaussian dominated profiles is also lower because of the $1/\sqrt{m}$ thermal dependency in equation (1.30). This reduces the line FWHM. To overcome this a "shift and stack" method was employed. This is where both the Lyα absorber and associated metal line, whether or not it could be seen, were transformed to a rest equivalent width. Each transformed Lyα and metal line was then superimposed with the rest to boost the equivalent width significance. The ratio of the column densities of these two overall absorbers was then used to determine the mean metallicity for the whole spectrum.

Improvements in the resolution and the sensitivity of spectrographs has meant that significant individual metal absorbers can now be detected and the metallicity of a Lyα absorber with contaminant metal X is calculated to be

$$Z_X = \left[\frac{X}{H}\right] = \log\left(\frac{N_X}{N_{HI}}\frac{f_{HI}}{f_X}\right) - Z_\odot \tag{2.1}$$

Where N_X/N_{HI} is the ratio of column densities and f_{HI}/f_X the ratio of the ion fraction of each species. It is common practice to quote the metallicity as a fraction of the solar metallicity $Z_\odot$. Metallicities quoted in this investigation were compared with the solar values given by Asplund et al. (2005).

Even if the signal-to-noise of the spectrum is large the error in metallicity that is derived from the ratio of the column densities will generally be small in comparison with the assumption made about the ion fractions. The neutral fraction of H I and the ion fraction of species X will depend to a great extent on the temperature, density, photoionisation rate $\Gamma(z)$ and whether the species is collisionally or photoionised (see Chapter 6 to see how these vary). This is why values for metallicity must be treated with caution because there is increasing evidence that we must acknowledge a multi-phase IGM.

2.3.2 The Halo Metal Magnesium II

Whether or not metals can escape to the IGM using momentum driven winds is important because if not it would be further evidence that the metal lines that are observed in quasar spectra arise only because the sightline has passed through or near to a halo.

For some metals this is almost certainly the case, such as the Mg II doublet (2796.352,2803.531 Å (Verner et al., 1994)). Mg II absorbers have been identified with Lyα absorbers that range in column density from the Lyman-limit to DLA systems, $\log_{10}(N_{HI}\,(cm^{-2})) \geq 17.3$ (Churchill et al., 2000). In fact many DLA systems at UV wavelengths have been identified by their Mg II absorption. For example the DLA system (z=0.313) that was observed in the FOS G160L spectrum of PKS 1127-145 has a corresponding Mg II system that was used to locate the redshift of the DLA first (Rao et al., 1995).

Mg II is used to probe galactic haloes because it is known to be closely associated with high column density systems and at relatively close impact parameters with galaxies. In these high column density systems the Lyα line is saturated, however the corresponding Mg II lines have a smaller optical depth. Therefore the separate components of a blended Mg II absorption system are more easily resolved and remain unsaturated.

The heavier metal species that also would be deposited by an escaping wind, such as the rare Ca II line and Fe II, are also closely associated with DLA systems and galaxy haloes. For example in SDSS spectra at redshifts $z \leq 0.5$ Ca II absorption lines were found at an impact parameter no greater than $\sim$ 24 kpc from the nearest galaxies (Zych et al., 2007).

2.3.3 Metals in the IGM

What of those metals lines that are also associated the with weak H I absorbers that pollute the IGM beyond a galaxy halo? These include the C IV (1548.1950,1550.7700 Å) and O VI (1031.9261,1037.6167 Å) species that were investigated with results that are summarised in Chapter 6.

Carbon IV

C IV has been found to be the most ubiquitous detected ion in the IGM and studies have been taken which measure the carbon metallicity of both DLAs and sub-DLAs. Fox et al. (2007) studied 74 DLA systems in UVES spectra which had $\log_{10}(N_{HI}\,(cm^{-2}))\geq$ 19 between the redshifts of $z = 1.75 - 3.61$. Every single line had an associated C IV doublet and these had a median column density of $N_{CIV} = 10^{14.2}$ cm^{-2}.

However since the advent of high resolution optical spectrographs, such as HIRES and UVES, C IV has also been found at redshifts $z \sim 3$ linked to Lyα absorbers that have $\log_{10}(N_{HI}\,(cm^{-2}))\sim$ 14.5. These are the C IV absorbers which may be the product of super-winds from Lyman-break galaxies (LBGs) (Adelberger et al., 2003).

On top of the contaminants in the IGM from these super winds we also expect a homogeneous minimum in the metallicity that was created shortly after the onset of reionisation. Early star formation would have polluted the IGM, and simulations suggest a minimum floor metallicity of $Z \sim 10^{-3.5}\ Z_{\odot}$ at a redshift $z \sim 10$ (Stocke et al., 2007). Therefore we should also ask whether there is a minimum Lyα column density at which C IV (or indeed any metal) can be seen?

Stocke et al. (2007) investigated this at redshifts of $z \leq 0.1$ by looking for evidence of C IV, O VI or Si III in high signal-to-noise STIS and FUSE spectra. 61 Lyα absorbers were identified as "void" absorbers as they were at a distance of more than 1.4 h_{70}^{-1}Mpc from the nearest L* galaxy. No significant metal absorption lines were detected, with an upperlimit in the metallicity $10^{-1.8\pm0.2}Z_{\odot}$. This could mean that these voids have never been contaminated, hence winds never reached this distance at these redshifts. However this minimum is not enough to rule out a primordial metallicity.

Even with all the metals observed at redshifts $z \gtrsim 2$ there is still a very large deficit, we can currently only see a fifth of the quantity of metals that are expected (Péroux et al., 2007). This has become known as the "missing metals problem". One explanation is that these metals exist in a different ionisation state as an undetectable species, such as the x-ray lines of O VII and O VIII. These "missing-metals" may also reside in the WHIM or the hot ICM.

Oxygen VI

A species that is believed to trace the WHIM at low redshift is O VI. Richter et al. (2004) came to this conclusion when comparing the positions of O VI absorbers in FUSE and STIS spectra with thermally broad Lyα lines. These absorbers had velocity widths that implied a temperature of $T \sim 10^5 - 3 \times 10^6$ K.

What causes these high temperatures and high metallicity is thought to be the O VI proximity to galaxies that exhibit winds and feedback.

At a redshift of $z \leq 0.15$ Stocke et al. (2006) observed 37 O VI absorbers in FUSE spectra with metallicities that were $0.10 \pm 0.05\ Z_\odot$. All of these lines were within 350-500 h_{70}^{-1}kpc from the nearest L^* galaxy and 200-270 h_{70}^{-1}kpc from the nearest $0.1L^*$ galaxy. This would place the O VI absorptions beyond the furthest extent of a wind and in a region level with or beyond the halo-IGM boundary. Also at these large distances O VI cannot be associated with an individual galaxy. Therefore rather then have O VI as part of the halo Stocke et al. (2006) placed them in the cosmic web.

Ten of the prominent Lyα lines were found at a distance of more than 3 h_{70}^{-1}Mpc away from the nearest galaxy and were assumed to lie in voids. As with C IV (Stocke et al., 2007) no O VI absorbers were associated with these Lyα lines.

So, if not within a galaxy halo, O VI is definitely observed to trace those regions that are above the mean density. This was confirmed by Tripp et al. (2006) while studying two strong O VI absorption systems in the sightline of PG 0953+415 at redshifts $z = 0.06807$ and 0.14232. Neither O VI line was found to be within 195 h_{70}^{-1}kpc of a $L \geq 0.04L^*$ galaxy. Instead Tripp et al. (2006) argue that O VI absorbers are located in an enriched cosmic web. They hypothesise that the dominant process of IGM O VI enrichment occurred at higher redshifts ($z \sim 2$). Those galaxies responsible have since moved on a significant distance and this explains why there are no galaxies within the immediate vicinity of O VI.

Guzzo et al. (2008) found a mean value for the galaxy pair dispersion velocity at a redshift of $z \sim 0.8$ of 412 km s^{-1}. Using this velocity and equation (1.10) to calculate the look-back time to a redshift of $z = 2$ we find that galaxies will have moved 4.2 h_{70}^{-1}Mpc, so this explanation suggested by Tripp et al. (2006) is highly plausible.

Together with results from simulations (Oppenheimer and Davé, 2008), Tripp et al. (2006) found evidence that the Lyα absorbers that are associated with the O VI lines favoured a model where they are predominantly photoionised at only $T \sim 10^4$ K. This means that only a minor fraction would trace the WHIM. These results conflict with the broad Lyα lines detected by Richter et al. (2004) but can be explained if we consider a multi-phase IGM.

2.4 Quantifying the Association between Galaxies and the IGM

2.4.1 The 2D 2-Point Correlation Function

Studies that attempt to quantify the association of Lyα and metal absorbers in QSO spectra with galaxies have used two methods. Early work, for example Lanzetta et al. (1995), Morris et al. (1993), because of the fewer number of absorbers available in low resolution spectra, tended to focus on individual systems of galaxies and absorbers. Now that a greater number of galaxies near the line of sight are available at lower limits in luminosity, and there are more absorbers from high resolution spectroscopy across more lines of sight, the analysis has become more statistical.

One approach is the 2D 2-point correlation function (ξ_{AG}) between absorbers and galaxies. The calculation is based on the estimator devised by Davis and Peebles (1983). The correlation function in equations (2.2a) and (2.2b) gives the measure of association between the Lyα absorbers and galaxies, across both the projected (σ) and line of sight (π) separations, by comparing the distribution of absorber-galaxy pairs with pairs drawn from random catalogues. Three versions of the estimator were used in this investigation. One where the number of real pairs was compared with the number of pairs between each real absorber and each random galaxy, as in (2.2a). This was the formula used by Ryan-Weber (2006) and Wilman et al. (2007). The second used equation (2.2b) and divided by the number of pairs between each random absorber and random galaxy.

$$\xi_{\mathrm{AG}}\,(\sigma,\pi) = \mathrm{N_{dr}}\frac{\mathrm{DD}}{\mathrm{DR}} - 1 \tag{2.2a}$$

$$\xi_{\mathrm{AG}}\,(\sigma,\pi) = \mathrm{N_{rr}}\frac{\mathrm{DD}}{\mathrm{RR}} - 1 \tag{2.2b}$$

DD, DR and RR are the number of real-real, real-random and random-random pairs respectively in each bin along the π and σ separations.

The normalisation constants $\mathrm{N_{dr}}$ and $\mathrm{N_{rr}}$ fix the total number of random pairs to the same value as the number of real pairs across all σ and π. This way $\xi_{\mathrm{AG}}\,(\sigma,\pi)$ gives a value for the extent of association, with no correlation in a particular bin producing 0.

If $\mathrm{n^i_a}$, $\mathrm{n^i_g}$, $\mathrm{n^i_p}$ and $\mathrm{n^i_q}$ are the number of real absorbers, real galaxies, random absorbers and random galaxies for each sightline i respectively, then the normalisation constants from equations (2.2a) and (2.2b) are:

$$\mathrm{N^i_{dr}} = \frac{\mathrm{n^i_a n^i_q}}{\mathrm{n^i_a n^i_g}} = \frac{\mathrm{n^i_q}}{\mathrm{n^i_g}} \tag{2.3a}$$

$$\mathrm{N^i_{rr}} = \frac{\mathrm{n^i_p n^i_q}}{\mathrm{n^i_a n^i_g}} \tag{2.3b}$$

We want to measure to what extent the gas detected in the IGM is part of a galaxy and how the nature of this association changes with redshift. One way to show this is by comparing ξ_{AG} with the galaxy auto-correlation function, ξ_{GG}. This is well understood and has been studied many times at various redshifts (Coil et al., 2007, Hawkins et al., 2003, Pollo et al., 2005). Due to structural collapse ξ_{GG} increases at lower redshift. By comparing ξ_{AG} with ξ_{GG} we will then know whether or not the gas can be thought to be part of these collapsed structures.

As the auto-correlation studies only one population the calculations for the normalisation constants N_{dr} and N_{rr} are different. The way we generated the random galaxies was to have X (a value of 200 was used) random galaxies for each real one. When producing the DR and RR pairs no pairs were made that contained a real galaxy and/or randoms generated from the same real galaxy. The normalisation constants were calculated to be:

$$N^i_{dr} = \frac{n^i_q n^i_g \left(n^i_g - 1\right)}{\left(n^i_g\right)^2 \left(n^i_g - 1\right)} = \frac{n^i_q}{n^i_g} = X \tag{2.4a}$$

$$N^i_{rr} = \frac{\left(n^i_q\right)^2 n^i_g \left(n^i_g - 1\right)}{\left(n^i_g\right)^3 \left(n^i_g - 1\right)} = \left(\frac{n^i_q}{n^i_g}\right)^2 = X^2 \tag{2.4b}$$

Therefore the normalisation constants depended only on the ratio X.

A third estimator devised by Landy and Szalay (1993) and used in Hawkins et al. (2003), Pollo et al. (2005) and Coil et al. (2007) when investigating the galaxy auto-correlation function used the equation.

$$\xi_{GG}\left(\sigma,\pi\right) = \frac{N_{rr}DD - 2N_{dr}DR + RR}{RR} \tag{2.5}$$

The variance associated with this formula is closer to a Poisson error and Landy and Szalay (1993) also states that equation (2.2a) introduces a small bias. This third estimator was also used in this investigation for both the cross and auto-correlation function, where the results which were of mixed success are summarised in Chapter 5.

2.4.2 Absorber-Galaxy Correlation in the Literature

Previous investigations that have studied the 2-point correlation function can be divided into four categories. First of all there is work that covers the local universe like in Rosenberg (2005) and Ryan-Weber (2006).

To compare with the results of Ryan-Weber (2006), Pierleoni et al. (2008) investigated ξ_{AG} by firing imaginary quasar sightlines through a hydrodynamic simulation at $z = 0$.

Then there are those studies at intermediate redshifts which utilised the quasar surveys that detected Lyα absorption in the near UV. These works include Chen et al. (2005), Morris and Jannuzi (2006), Wilman et al. (2007) and Williger et al. (2006).

At higher redshift Lyα passes into the optical regime and Adelberger et al. (2005, 2003) correlated these absorption lines with LBGs.

Observations at low redshift

The studies made by Rosenberg (2005) and Ryan-Weber (2006) at low redshift both correlated Lyα absorption with gas rich galaxies that had been picked out using 21 cm radiation. The expectation was that a strong correlation would be found.

Rosenberg (2005) comprised of two investigations. The first studied the H I environment of the closest DLA outside the local group, SBS 1543+593. Three gas rich galaxies were located within 183 kpc each with a mass $M \sim 10^8$ $M_\odot$. The evidence from this system suggested a close connection. The second data set studied 16 Lyα systems along 4 quasar sightlines, each which had a column density within the range $\log_{10}(N_{HI}\,(\text{cm}^{-2})) = 14.09 - 21.81$. In the fields of these absorbers were 8 gas rich galaxies and 3 spirals. However the closest galaxy-absorber pair was 95 kpc apart and only half of the absorbers were located within 500 kpc of a galaxy. Two absorbers were located in voids at 2 Mpc from any of the galaxies.

This lack of correlation at low redshift was contradicted by the work from Ryan-Weber (2006). 5317 gas rich galaxies that had been detected in the HIPASS galaxy survey (Zwaan et al., 2003) were correlated with 129 absorbers from 27 lines of sight. These spectra were gathered from STIS and GHRS data and all absorbers had $\log_{10}(N_{HI}\,(\text{cm}^{-2})) = 12.41 - 14.81$. These real absorber-galaxy pairs were compared with a random catalogue of 20000 galaxies and 5000 absorbers. The galaxy auto-correlation was also calculated and the results to both $\xi_{AG}\,(\sigma,\pi)$ and $\xi_{GG}\,(\sigma,\pi)$ are shown in Figure 2.2.

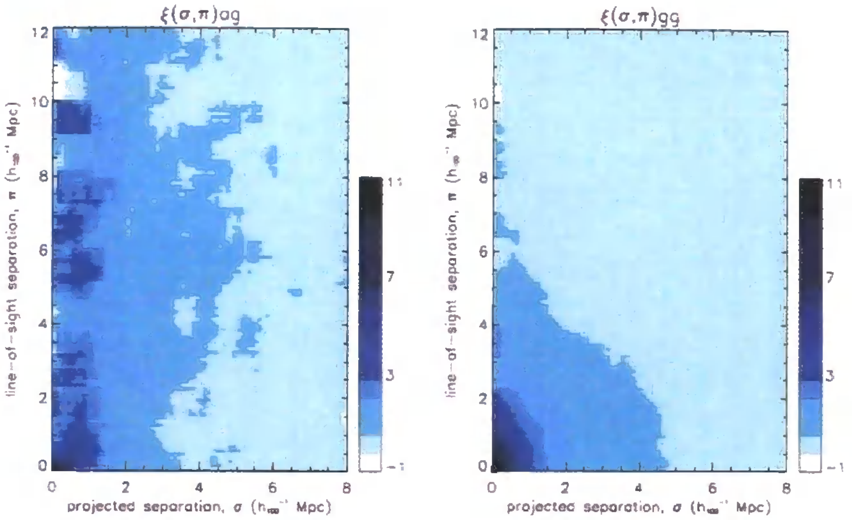


Figure 2.2 $\xi_{\rm AG}$ (σ,π) (left) and $\xi_{\rm GG}$ (σ,π) (right) were computed and binned in these density plots on a scale of $\Delta\sigma = \Delta\pi = 0.1$ h$_{100}^{-1}$Mpc then smoothed. Jackknife errors predict variations in $\xi_{\rm AG}$ and $\xi_{\rm GG}$ of ≤ 0.5 and ≤ 0.2 respectively. These plots show that at low redshift $\xi_{\rm AG}$ and $\xi_{\rm GG}$ are similar. This figure is from Figures 2 and 3 from Ryan-Weber (2006).

What the plots in Figure 2.2 show is that at low redshift and small separation there is a strong correlation between absorbers and gas rich galaxies and between the galaxies themselves. $\xi_{\rm AG}$ and $\xi_{\rm GG}$ are also found to be similar in value. This supports the idea of matter clustered within the vicinity of galaxies (within the nodes of the cosmic web) be they absorbing clouds or other galaxies. However unlike the galaxy auto-correlation, that decays radially along the σ and π direction, $\xi_{\rm AG}$ remained high out to 10 h$_{100}^{-1}$Mpc along the line of sight.

This "finger-of-god" effect at small projected separation was probably caused by the increased velocity dispersion between these pairs. The proposed origin of this are large galactic outflows, although a high cross-correlation at 10 h$_{100}^{-1}$Mpc would suggest outflow velocities $\sim$ 1000 km s^{-1}, and gas accretion. The gas falls in from the more diffuse IGM and drains along filaments causing an elongation along the line of sight. The gas rich galaxies that were detected using 21 cm radiation also tend to have a low mass (M $\sim 10^8$ M$_\odot$), and it is galaxies such as these that Kereš et al. (2005) claimed were dominated by this cold method of accretion.

The "finger-of-god" effect is thought to become even more pronounced at low redshift because as the filaments become more structured they collapse to give a smaller cross-section. Therefore as time progresses the projected separation between a galaxy embedded in this collapsing filament and an absorber decreases, while there is an increase in their velocity dispersion.

That an equally strong "finger-of-god" is not observed for the auto-correlation of galaxies is an indication that unlike the absorbers, the galaxies are more static than the gas within the filamentary structure, and it is the absorbers that have a much greater peculiar velocity on top of the Hubble flow.

A simulation of ξ_{AG} at low redshift

The results of Ryan-Weber (2006) were compared with simulations in Pierleoni et al. (2008). Using the SPH codes GADGET-2 and GADGET-3 (Springel, 2005a), Pierleoni et al. (2008) ran a simulation that comprised of 2×400^3 dark matter and gas particles each of mass 4.3×10^7 $\mathrm{M}_\odot$ in a box of width 60 h_{72}^{-1}Mpc. Three simulations were run to a redshift of zero, where each one had a different galaxy wind model. Either no winds, strong winds or extremely strong winds were present.

333 sightlines were fired through the simulation along each Cartesian axis in order to generate the spectra. The absorbers from these spectra that had a column density in the range $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12.41 - 14.81$ were then correlated with $\sim$ 5000 galaxies that resided in haloes of mass $\mathrm{M} = 8 \times 10^{10} - 10^{13.5}$ $\mathrm{M}_\odot$. However contrary to Ryan-Weber (2006), Pierleoni et al. (2008) found that at local redshifts the galaxy auto-correlation is stronger than the absorber-galaxy cross-correlation.

No significant variation in the correlation function was seen for the different wind models. Only in the case of the extremely strong winds was a minor decrease in cross-correlation observed within 0.3 h_{72}^{-1}Mpc from the galaxies.

Figure 2.3 shows ξ_{AG} and ξ_{GG} plotted in redshift space from the simulation that had no winds. Unlike Ryan-Weber (2006) no elongation is observed along the line of sight.

To recreate the observations of Ryan-Weber (2006) further still the correlation function was investigated using only 27 imaginary lines of sight. This caused the sample noise to increase, and so the results of both the cross and auto-correlation functions now became comparable. Bootstrap errors showed a difference between

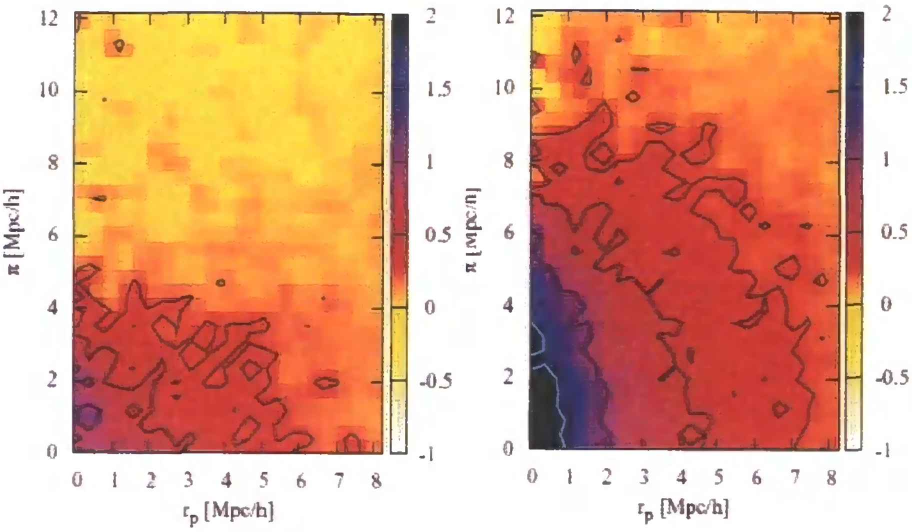


Figure 2.3 The ξ_{AG} (σ,π) and ξ_{GG} (σ,π) from the simulation. The figures have a binsize of $\Delta\sigma = \Delta\pi = 0.4$ h^{-1}Mpc. The contour lines in the cross-correlation (left) are at ξ_{AG}=1,0.5 and 0.25. Those in the auto-correlation (right) are at ξ_{GG}=2,1,0.5 and 0.25. Bootstrap errors gave variations in the cross and auto-correlation functions of ~ 0.2 and ~ 0.08 respectively. These figures are Figures 6 and 7 from Pierleoni et al. (2008).

ξ_{AG} and ξ_{GG} at only a 1σ significance level. Pierleoni et al. (2008) suggest that it is this lack of data pairs from too few sightlines that made the cross and auto-correlation functions indistinguishable at low redshift. They also attribute the large "finger-of-god" in Figure 2.2 to this sparse sampling of absorbers.

It has been argued (Wilman et al., 2007) that a "finger-of-god" may also be caused by a large anisotropy, for example a quasar line of sight that runs parallel and close to to a large filament. This would cause a large elongation along the line of sight. However Pierleoni et al. (2008) ruled out a "finger-of-god" being caused by this geometric effect in their simulation by showing that the ξ_{AG} result remained the same no matter along which orientation the sightlines had been directed.

Observations at low to intermediate redshift

Work by Williger et al. (2006) studied the 60 Lyα absorbers which are at redshifts $z = 0.002 - 0.423$ and are located in the STIS spectrum of the line of sight towards quasar PKS 0405-123.

These absorbers were correlated with galaxies collected from a CFHT survey that had a minimum magnitude of r $\sim$ 23. It was found that the correlation function was heavily dependent on the column density, even though the range in values was small ($\log_{10}(N_{HI}(cm^{-2}))\sim 13.1 - 14.5$). The sub-sample of weak absorbers with $\log_{10}(N_{HI}(cm^{-2}))= 13.1 - 13.2$ only showed significant correlation at separations less than $\Delta v \leq 125$ km s^{-1}. Meanwhile those absorbers with $\log_{10}(N_{HI}(cm^{-2}))= 13.5 - 14.0$ had the most significant correlation out to 250 km s^{-1}, with values that were comparable to the galaxy auto-correlation.

Chen et al. (2005) also studied the sightline towards PKS 0405-123 and compared the positions of 112 Lyα absorbers from STIS and FOS spectra with 482 galaxies. These were collected using a magnitude limited survey down to R$\leq$ 20.

The most striking result found by Chen et al. (2005) was a 2.5σ difference in the level of cross-correlation between galaxies and absorbers that depended on whether the galaxy in question was dominated by emission or absorption spectral lines. The correlation function between galaxies that were dominated by emission and Lyα absorbers that had $\log_{10}(N_{HI}(cm^{-2}))\geq 14.0$ was found to be similar to the emission type galaxy auto-correlation. Both of these remained high out to 250 km s^{-1}, with values of $1+\xi_{AG}= 7.0 \pm 2.0$ and $1+\xi_{GG}= 7.2 \pm 2.0$. No significant cross-correlation was seen out to the same velocity separation for those galaxies dominated by absorption, $1+\xi_{AG}= 2.0 \pm 2.0$. However self-clustering amongst absorption galaxies showed the strongest correlation out of all the samples, $1+\xi_{GG}= 13 \pm 2.0$. The reasons for this suggested by Chen et al. (2005) are that since absorption dominated galaxies are those that inhabit higher mass haloes a stronger auto-correlation is expected. Meanwhile the cross-correlation results are stronger in the emission dominated galaxies because both these and the Lyα absorbers are part of the same halo system.

Morris and Jannuzi (2006) collected a sample of 636 galaxies with redshifts $z \leq 1$ using the CFHT MOS. These were in 16 fields of view and 50% complete down to R=24.4. Along these same fields the positions of 406 Lyα or C IV absorption systems from 17 HST FOS spectra (Weymann et al., 1998, Young et al., 2001) were compared with the galaxy positions. The absorbers had a column density that ranged from $\log_{10}(N_{HI}(cm^{-2}))= 13 - 19$.

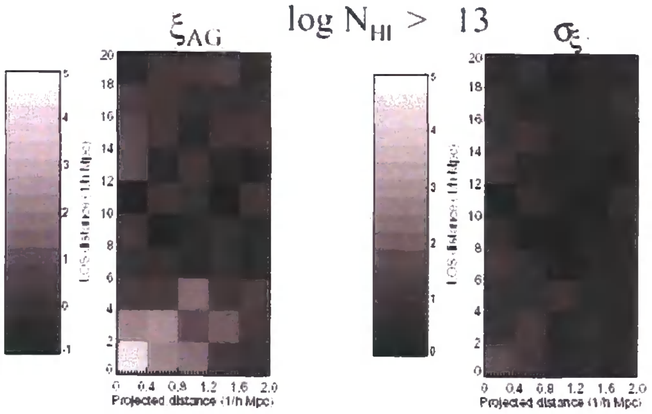


Figure 2.4 These results show $\xi_{\mathrm{AG}}(\sigma,\pi)$ of the dataset from Morris and Jannuzi (2006) binned $\Delta\sigma = 0.4$, $\Delta\pi = 2.0$ h_{70}^{-1}Mpc for all Lyα absorbers with $\log_{10}(\mathrm{N}_{\mathrm{HI}}\,(\mathrm{cm}^{-2})) \geq 13$. This shows that beyond the bin of smallest separation the correlation is very weak between absorbers and galaxies with a sample noise that is comparable with this value. This figure is Figure 1 from Wilman et al. (2007).

Using these catalogues Wilman et al. (2007) then calculated the two-point cross-correlation function by pairing each galaxy with each absorber that were at a distance of $\pi \leq 20$ h_{70}^{-1}Mpc and $\sigma \leq 2$ h_{70}^{-1}Mpc. The maximum of $\sigma = 2$ h_{70}^{-1}Mpc is set by the CFHT field of view at these redshifts. The results for the Lyα absorber-galaxy cross-correlation for all column densities are plotted in Figure 2.4 together with the jackknife error for each bin.

Figure 2.4 shows that there is little evidence for correlation between absorbers and galaxies beyond the central bin of $\pi = 2$ and $\sigma = 0.4$ h_{70}^{-1}Mpc. The error margin for these bins is of the order of the correlation value.

The majority of the pairs from this central bin were made up of those absorbers that had a column density of $\log_{10}(\mathrm{N}_{\mathrm{HI}}\,(\mathrm{cm}^{-2})) \geq 17$. This is shown in Figure 2.5 where the correlation function is shown to be strongly dependent on the column density.

Based on the results in Figure 2.5 Wilman et al. (2007) reached the conclusion that clustering is weak for absorbers up to a column density of $\log_{10}(\mathrm{N}_{\mathrm{HI}}\,(\mathrm{cm}^{-2})) = 16$, with only tentative evidence for an increase in ξ_{AG} at small separation. Above this column density the correlation function increased substantially. However the associated error with this result is still too great to say that an absorber with $\log_{10}(\mathrm{N}_{\mathrm{HI}}\,(\mathrm{cm}^{-2})) \geq 17$ is no longer part of the IGM and should now be considered

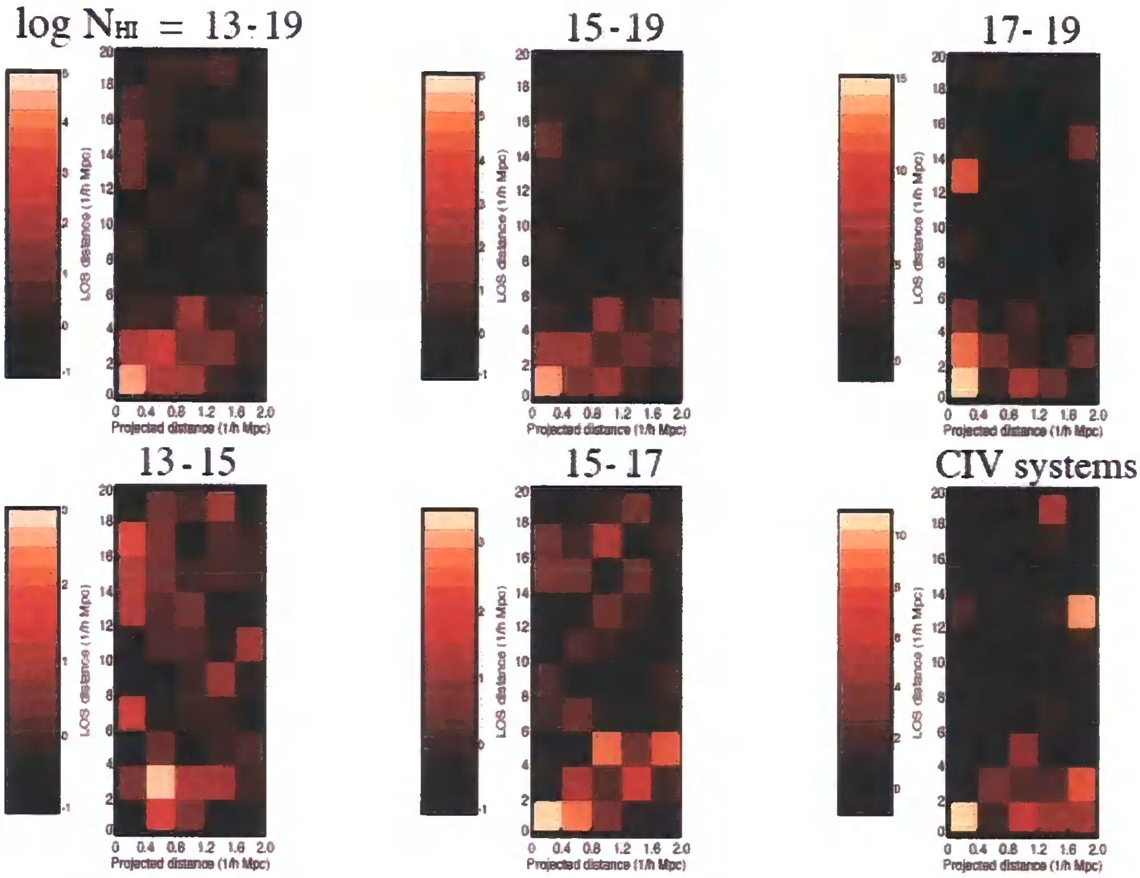


Figure 2.5 These plots are of ξ_{AG} binned $\Delta\sigma = 0.4$, $\Delta\pi = 2.0\ h_{70}^{-1}$Mpc and according to the column density of the absorber. As the column density increases so does the level of correlation in the bin of smallest separation. The sample noise for each plot is similar to that plotted in figure 2.4. The C IV correlation is comparable to the H I-galaxy signal with $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 16.5$. This figure is Figure 2 from Wilman et al. (2007).

to predominantly reside in a galactic halo. In addition to this the column densities had been estimated assuming a fixed b parameter of 30 km s^{-1}, and so it would be naive to introduce a critical column density based on this evidence.

The result for the correlation of galaxies with C IV agrees with those statements made in Section 2.3.3. It was found that C IV absorbers are correlated with galaxies at small separation to the same extent as that found for Lyα systems with $\log_{10}(N_{HI}\,(cm^{-2}))= 16.5$. Therefore C IV is found within and near to galaxy haloes. However this result will have been limited by the minimum detectable EW of the FOS spectra of 0.24 Å.

Unlike Chen et al. (2005) no evidence was found for a difference in correlation depending on the spectral type of the galaxy. However unlike Chen et al. (2005) the dataset of Wilman et al. (2007) used FOS and not STIS spectra. These are of a considerably lower resolution ($\Delta v = 230$ versus 7 km s^{-1} respectively). It is possible that those lines that are closely associated with haloes were most likely blended lines so that many possible pairings at small separation were lost.

No "finger-of-god" was observed in the results of Wilman et al. (2007). This was maybe because those galaxies targeted by the CFHT were of a higher mass, and so rely more on hot accretion. However it is also possible that there were an insufficient number of pairs to show this effect.

Elongation along the line of sight was present when the correlation plots of Wilman et al. (2007) were recreated using SPH simulations. Sightlines were fired through a simulation at a redshift of $z = 0.5$, and those absorbers from the simulated spectra were correlated with galaxies that were of total baryonic mass $M \geq 2 \times 10^{10}$ $M_{\odot}$. It is thought that this elongation in the simulated data arose either because of gas draining along the filaments and falling into the denser structures, or there was a sightline that ran parallel to a filament causing many pairs to appear along the line of sight.

Observations at high redshift

In order to observe the IGM-galaxy relationship at high redshift (where the Lyα absorption has now reached optical wavelengths) Adelberger et al. (2003) correlated the Lyα lines from 8 quasars (which had redshifts z = 3.1 – 4.1) with 431 LBGs in the study that was mentioned in Section 2.2.2

Target galaxies were selected that are at a redshift of $z = 3 \pm 0.25$ using the Lyman-break technique. The galaxy field was observed using U, G and R filters. At these redshifts the U filter is blue-wards of the Lyman-limit so that the galaxies were absent in this image. The flux is strongest in the G band and then slightly less in R. Therefore with an appropriate colour cut a catalogue of galaxies could be identified for follow-up spectroscopy.

It was expected that because H I follows the density structure that the greater concentrations of gas would be located near to the galaxies. This was found to be so because gas with a greater optical depth was found to lie within $\sim$ 10 h^{-1}Mpc. However evidence arose of an anti-correlation within $\sim$ 0.5 h^{-1}Mpc. It was hypothesised that this was caused by super-winds. These cleared the immediate vicinity of the galaxies and heated the surroundings, thereby decreasing the neutral fraction and so increasing the photon transmission. It is thought that these winds can travel up to speeds of $\sim$ 600 km s^{-1} and can escape the potential of the LBG (Pettini et al., 2001).

With a larger data set of 23 sightlines and 1044 UV selected galaxies (z = 1.8 – 3.3) Adelberger et al. (2005) found that the anti-correlation at small separation disappeared.

Instead a result was reached where the optical depth of the neutral gas increased as the separation from the galactic centre decreased from 4 to 0 h^{-1}Mpc. However in a third of the cases only weak absorption was detected around some galaxies, and so the theory persists that these galaxies are surrounded by an expanding wind. It is thought that these exceptions occur when the sightline has passed through the pathway of least resistance, a route that these escaping winds will also take.

Adelberger et al. (2005) also found that the cross-correlation of LBGs with C IV systems with $\log_{10}(N_{CIV}(\mathrm{cm}^{-2})) \geq 12.5$ was comparable with the galaxy auto-correlation function. This suggests that they are part of the same system, and this was definitely the case for absorbers with a higher column density. C IV absorbers with $\log_{10}(N_{CIV}(\mathrm{cm}^{-2})) \gtrsim 14.0$ were detected within $\sim$ 80 h^{-1}kpc of the star-forming galaxies and travelling at $\Delta v = 260$ km s^{-1}. This suggests that these were a component of an escaping wind. The C IV absorbers that had a column density of $\log_{10}(N_{\rm CIV}\,({\rm cm}^{-2})) > 14.0$ were found within ~ 40 h^{-1}kpc.

These findings support the theory that was raised in Section 2.3.3. That many metals that are observed at local redshift were deposited in the IGM by these powerful winds at earlier epochs.

The large sample noise in all these observations of $\xi_{\rm AG}$ means that more data is needed. However even less results have been obtained for the redshift window $z \sim 1$. The main reason for this is in acquiring a sufficient number of galaxy redshifts at this epoch that are within the same fields of view as quasars for which there is high resolution UV spectral data. This is hard because of the "redshift desert" of spectral lines, and the copious amounts of time that is required on 8 m class telescopes in order to perform spectroscopy on these galaxies.

The aim of this thesis will be to bridge this gap. In order to investigate the correlation function at a redshift of $z \sim 1$, H I absorption was detected in the absorption spectra of the STIS quasars PKS 1127-145 ($z = 1.187$) and HE 1122-1648 ($z = 2.40$) (Bechtold et al., 2002, Carswell et al., 2002a). The redshifts of galaxies in the surrounding $7' \times 5'$ field of view were then collected using multi-object-spectroscopy (MOS) on the FOcal Reducer/low dispersion Spectrograph (FORS2) instrument. The method used to generate the galaxy catalogues and the quasar absorption linelists are described in Chapters 3 and 4 respectively.

Chapter 3

The FORS2 Galaxy Survey

3.1 The Need for a Galaxy Survey

A spectroscopic galaxy survey was required in the field of view around a background quasar in order to show to what extent the H I absorbers found in chosen QSO absorption spectra were correlated with the galaxies. Two $6.8' \times 5.7'$ galaxy surveys towards the quasars PKS 1127-145 ($z = 1.187$) and HE 1122-1648 ($z = 2.40$) (Bechtold et al., 2002, Carswell et al., 2002a) were carried out.

The surveys were carried out using FORS2 (Appenzeller et al., 1998) in mask-exchange-unit (MXU) mode at the VLT telescope Paranal on the 20-22 February 2006 (MJD 53787-53789).[1] The total wavelength range of the E230M grating used in the archived STIS data was between 2050-3100 Å. This meant we could potentially detect Lyα absorbers between the redshifts of 0.68 and 1.55. For this reason it was our aim to observe galaxies within a similar redshift window. This was done by selecting galaxies based on a magnitude limit ($21.5 \leq m_R \leq 24.5$) and targeting those likely to have [O II] emission appearing in the 5600-11000 Å bandpass. It is thought these [O II] emitters, indicative of star-forming galaxies, with their systems of feedback, outflows and photoionisation are responsible for most of the complexity in the gas-galaxy relationship at $z \sim 1$.

---

[1] Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. ESO programme number 076.A-0312.

3.2 An Overview of the FORS2 Instrument and Pixels

3.2.1 The FORS2 Instrument

FORS2 is located on the Cassegrain focus of the VLT UT1 telescope (Antu) at Paranal. The site is located at latitude $-24°37'33.1''$, longitude $70°24'11.6''$ W, 2635 m above sea-level.

Operating in the optical regime, FORS2 offers both imaging and long-slit or multi-object spectroscopy (MOS). What made FORS2 particularly suited to our needs was MOS could be performed with the mask-exchange unit (MXU). Traditional MOS on the FORS2 has a single mask in the focal plane with 19 movable slits. The MXU is a carousel that can hold up to 10 masks. The slits, which have a custom shape and dimension are cut with a LASER cutting machine. This enabled us to target 100s of galaxies in one night, as each mask using the dimensions we chose, could contain up to 50 slits (Jehin and O'Brien, 2007).

The square FORS2 detector consists of two tiled chips, "Thor" and "Belenos", hereafter called chips 1 and 2 respectively. Each of these contains 2048 (vertical) x 4096 (horizontal) pixels of size 15 x 15 μm, with a $\sim$ 480 μm chip gap. This detector was ideal because unlike the sister instrument FORS1, the pixels are particularly sensitive at wavelengths $\geq$ 6500 Å as is shown in Figure 3.1.

We were able to survey the $6.8' \times 5.7'$ field of view located around a quasar using a standard resolution of $0.25''$ pixel^{-1}.

3.2.2 Pixel Statistics

In CCD electronics the number of photo-electrons detected by each pixel is represented by a number in ADU (Analogue-to-Digital Unit). This is related to the count rate by the CCD gain factor K. For the pre-imaging and spectroscopic frames the gain was 1.25 and 0.7 e$^-$ ADU^{-1} respectively. Those frames with a higher gain had a faster read-out time. It was preferable to read at a slow rate as this lead to a smaller read-noise.

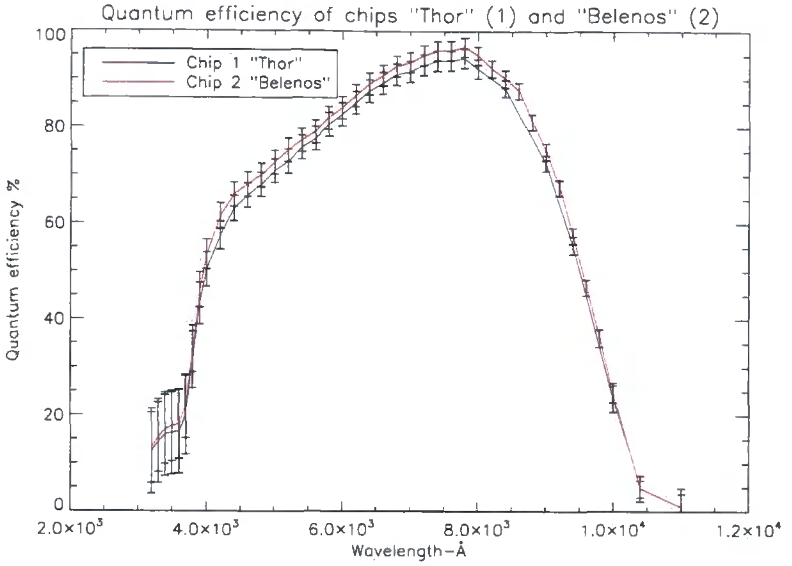


Figure 3.1 The FORS2 MIT pixels have a high quantum efficiency ≥ 6500 Å at the expense of shorter wavelengths. This was ideal as we were targeting galaxies in a 5500 → 10000 Å bandpass. Data is taken from the FORS2 manual (Jehin and O'Brien, 2007).

The counts were subject to Poisson statistics because of the random number of photo-electrons that would strike the CCD during the exposure time. The counts were split between those photons from the source, the sky background and instrument contributions. When combined these gave the signal-to-noise ratio (S/N) of the observation. The sources of this error included:

1. The source noise. If a source had R_* photo-electrons per second then the 1σ error in an exposure time t was $\sigma_* = \sqrt{\mathrm{R}_*\mathrm{t}}$.
2. Sky-background. All images/spectra had the sky brightness/sky-lines subtracted. This was done by sampling the median count value in the background of the image and subtracting this from all pixels. The variations across the image in the size of this background contributed towards the sky-noise. $\sigma_{\mathrm{sky}} = \sqrt{\mathrm{R}_{\mathrm{sky}}\mathrm{t}}$.
3. The read-noise. The level of read-noise depended on the frequency at which the pixels were measured and on the level of binning prior to read-out. The slow 100 kHz rate that was used has a read noise of 3.0 e^- pixel^{-1}. Fast 200

kHz read-out caused a read-noise of 4.1 e^- pixel^{-1}. With no binning each pixel contributed to the read-noise. However as pixels were binned 2 × 2 then the read-noise contribution to the S/N was reduced by a factor of 2^2.

4. Dark current. Thermally generated electrons within the chip contributed to the dark current as they accumulated during the exposure time. The FORS2 chips are maintained at a temperature of 153 K and this way they only produced $\sim 3\ e^-\ \text{pixel}^{-1}\ \text{h}^{-1}$.

Each of these sources of noise could assumed to be independent and so were added in quadrature to give the S/N equation.

$$\frac{\mathrm{S}}{\mathrm{N}} = \frac{\mathrm{R_* t}}{\sqrt{\mathrm{R_* t + R_{sky} t + R_{dark} t + read\ noise^2}}} \tag{3.1}$$

3.3 Pre-imaging and Mask Generation

3.3.1 Pre-imaging

Prior to the Multi-Object-Spectroscopy (MOS) observing run pre-imaging was taken in service mode of the two quasar fields using the Bessel R_SPECIAL+76 filter (see Figure 3.5). Our choice of a red filter was an attempt to target galaxies most likely to have a high m_R at a redshift of $z \sim 1$. Photometric conditions with good "seeing" ($\leq$ 0.8) was requested for each 10 minute exposure in order to attain accurate astrometry and photometry when selecting faint galaxies.

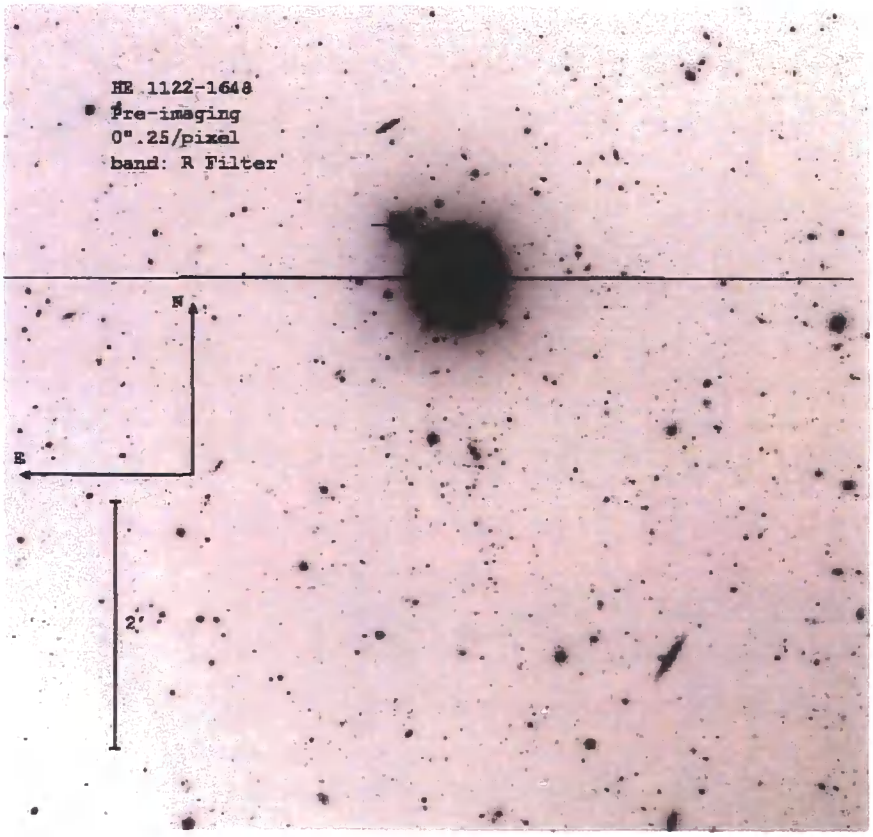
The pixels were binned 2 × 2 prior to readout so each image contained 1024 × 2048 pixels. This was to reduce the effects of read-noise. The areas of sky that were surveyed are in Table 3.1.

The images were reduced using the FORS2 imaging pipeline at ESO headquarters in Garching, Germany. As part of the routine service mode calibration data had also been taken at the same time, this included bias frames, sky flats and photometric standard images. Using this calibration data the raw images were bias subtracted, flat-fielded and cosmic rays were removed.

Table 3.1. The coordinates of the 6.8′ × 5.7′ fields surveyed in the foreground of the quasars HE 1122-1648 and PKS 1127-145

Quasar	RA (J2000)	DEC (J2000)
PKS 1127-145	11 30 21.57 → 11 29 52.50	−14 53 34.84 → −14 45 54.81
HE 1122-1648	11 24 57.47 → 11 24 28.07	−17 09 25.71 → −17 01 45.68

Figure 3.2 The pre-imaging 6.8′ × 5.7′ field surrounding the quasar HE 1122-1648. The software SExtractor was run on this image to generate a galaxy catalogue to select targets for follow-up spectroscopy. Unfortunately a star left a saturation trail where the maximum CCD pixel count (2^{16}) was exceeded and charge leaked across the chip.



The final stages in reducing the science images was then completed locally using the standard IMUTIL tasks within IRAF.[2] The overscan region of each frame was cropped and then images taken from chips 1 and 2 were joined together using IMTILE.

3.3.2 Using Source Extractor

The software Source Extractor (v2.4) (Bertin and Arnouts, 1996) was then run in order to compile a catalogue of all the possible targets for follow-up spectroscopy.

First of all both images were superimposed with a rectangular grid. We employed a cell size of 64 pixels and the median intensity was measured in each cell. Those pixels $\pm 3\sigma$ of this value were clipped and a new median was found. This process was then repeated until all pixels were within 3σ. This final median pixel value was taken to be the local background intensity and was subtracted from each respective grid.

We then defined a detection threshold above which an object was considered. We used a minimum criteria of 5 adjacent pixels, each $\geq 2\sigma$ above the RMS of the sky background that had been subtracted.

Source Extractor then calculated the astrometry of each group of pixels that exceeded this threshold limit. A rectangle was placed over the pixels that made up each object and the barycentre was found. Within each rectangle an elliptical aperture was then defined that covered those pixels over the threshold (see Figure 3.3) with elongation $\frac{\mathrm{A}}{\mathrm{B}}$ and position angle θ.

This way the image coordinates and radius of each target was determined. By using the date, location and pointing information stored in the image header ICRS J2000 coordinates were then calculated.

Source Extractor then went on to calculate the photometry of every target and classified each one according to whether it is a star or galaxy.

Every output parameter available was calculated when Source Extractor was

---

[2] IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation.

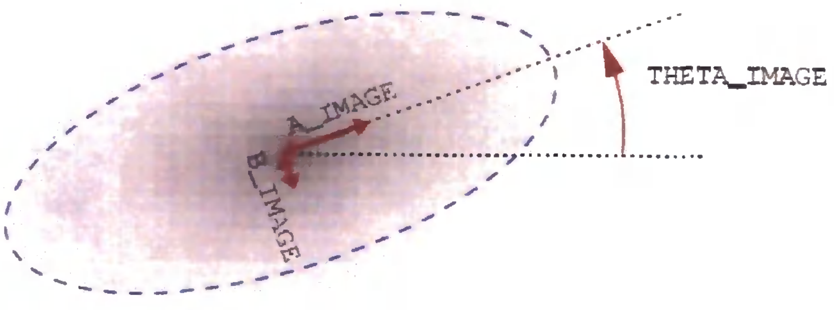


Figure 3.3 Each potential target, defined by a cluster of at least 5 pixels greater than the defined threshold limit of $\geq 2\sigma$ above the local background, was covered by an ellipse. The parameters A, B and θ were then used in the astrometry calculation. This figure is from Bertin (2006).

run. Two of these were used in the selection to produce a catalogue of potential targets and so deserve particular attention.

- MAG AUTO is an estimate of the apparent magnitude of the target. For each object a new ellipse was drawn. The radii of this ellipse were based on the Kron radius (Bertin and Arnouts, 1996), through which $\geq 90\%$ of an object's flux was transmitted. For a pixel distance x (maximum x = R) from the targets barycentre, 2nd moment x^2 and pixel intensity I(x) the Kron radius is given by (Kron, 1980):

  $$\mathrm{R_{KRON}(R)} = \frac{2\pi \int_0^R \mathrm{dx\, I(x)x^2}}{2\pi \int_0^R \mathrm{dx\, I(x)x}} \tag{3.2}$$

  The Kron radius is approximately equal to twice the semi-major axis of the ellipse, 2A (or 2B) (Bertin, 2006). FLUX AUTO was the sum of the pixel values within this ellipse.

  MAG AUTO was chosen over the other magnitude estimators MAG APER, MAG ISOCOR and MAG BEST. MAG APER measures the flux through a standard circular aperture with a fixed radius only. MAG ISOCOR is insufficient because it assumes a symmetric Gaussian intensity profile which is ideal for stars, but the fainter wings of galaxies are neglected. A problem with MAG AUTO occurs if a field is crowded and there is the chance that ellipses will overlap, in which case MAG ISOCOR is used instead. MAG BEST uses MAG ISOCOR for when a neighbouring object will change the predicted magnitude by 10%. As neither field was too populated it was decided to use MAG AUTO only.

The relative magnitude was calibrated by the user defined parameter MAG ZEROPOINT. This was determined by measuring the total counts of exposed standard stars within the image foreground and comparing these flux values with that of their magnitude in the USNO-A2.0 star catalogue.

- CLASS STAR was used in object classification by assigning to each potential target a number between 0 (a galaxy) and 1 (a star). The value of CLASS STAR was found by a simple "Neural-network". This is a program where a computer has been "taught" to assign a particular output having been trained using simulated images (Bertin and Arnouts, 1996). The algorithm is fed 10 parameters, these are the 8 isophotal areas surrounding the barycentre of the target, the intensity of the brightest pixel and the "seeing" at FWHM. By supplying more than one area the certainty in CLASS STAR is improved.

Figure 3.4 shows a plot of CLASS STAR against apparent magnitude (m_R) for all the 2666 objects extracted from the field of HE 1122-1648 using the threshold described above. One can tell that the reliability of CLASS STAR decreased as the targets become fainter. The neural network could no longer differentiate between a star and galaxy, particularly because the 8 isophotal areas were composed of too few pixels. So the most probable value was calculated which tended to 0.5.

Using this plot a cut was made to select those galaxies that would be targeted. In order to minimise the possibility of any stars being selected a conservative cut of CLASS STAR ≤ 0.8 was made (Cristóbal-Hornillos et al., 2003).

This plot also gave an indication as to what would be an effective faint magnitude cut-off. To avoid those targets where CLASS STAR was ambiguous a faint cut of $m_R \leq 24.5$ was chosen. The minimum brightness was also set by what could be observed within the alloted exposure time which was tested using the FORS2 exposure time calculator.

Likewise to remove those foreground galaxies likely to have redshifts too low to correlate with the H I absorbers a bright cut-off of 21.5 was employed.

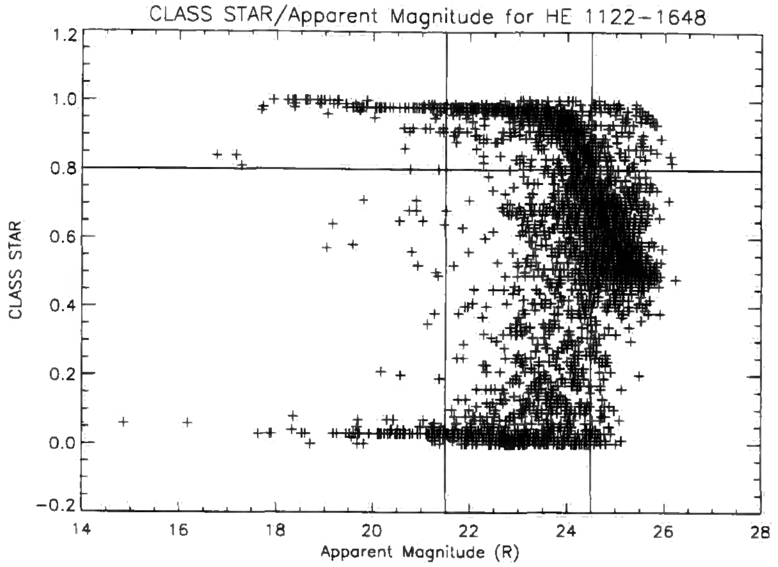


Figure 3.4 This horseshoe-shaped plot of the CLASS STAR classification against magnitude allowed cuts to be made so that galaxies could be selected for use in follow-up spectroscopy. Those selected had $21.5 \leq R \leq 24.5$ and $0 \leq$ CLASSSTAR ≤ 0.8.

3.3.3 Making the Masks

These cuts meant the final catalogues consisted of 1062 and 1122 objects in the fields of HE 1122-1648 and PKS 1127-145 respectively. The lists were then passed to the FIMS software for the FORS2 instrument that designs custom masks for the MXU.

FIMS allows the observer to define the number, size, position and shape of the slits. We designed rectangular slits $1''$ in width by 8 $''$, this meant that all targets took up no more than $^1/_4$ of the length of the slit. This allowed ample space for sampling the sky so that it could be subtracted later. The slit sizes and the galaxy density meant that up to 50 slits could be made per mask.

When investigating 2D 2-point correlation functions, especially if the galaxy survey is far from complete, it is important that the slits are assigned randomly so as not to bias a particular absorber-galaxy/galaxy-galaxy angular separation. Therefore a random slit allocation was used in FIMS to ensure a random distribution of galaxies was selected when designing each mask.

The catalogue was split between those galaxies with $21.5 \leq m_R \leq 23.5$ and $23.5 \leq m_R \leq 24.5$ so that separate "bright" and "faint" masks would be used. They could then be exposed for different periods of time. Altogether 22 masks were designed with each mask containing approximately 40 slits. There were 11 for each field of view, including 5 that were assigned "bright" objects and 6 that had slits over galaxies with $23.5 \leq m_R \leq 24.5$.

These files were then passed to the machine room at Paranal. Here the slits were cut using a LASER prior to the masks being placed in the carousel on the FORS2 instrument.

3.4 Details of the Observing Run

The FORS2 MXU was used in visitor mode using the R_SPECIAL+76 filter and the 200I+28 grism. Figure 3.5 shows the response of these two components. Using a slit width of 1″ and this grism gave a central wavelength of 7450 Å and allowed a spectral range of 5600 → 11000 Å. The dispersion was 2.43 Å pixel^{-1} with a resolution $\frac{\lambda}{\Delta\lambda} = 380$ (Jehin and O'Brien, 2007).

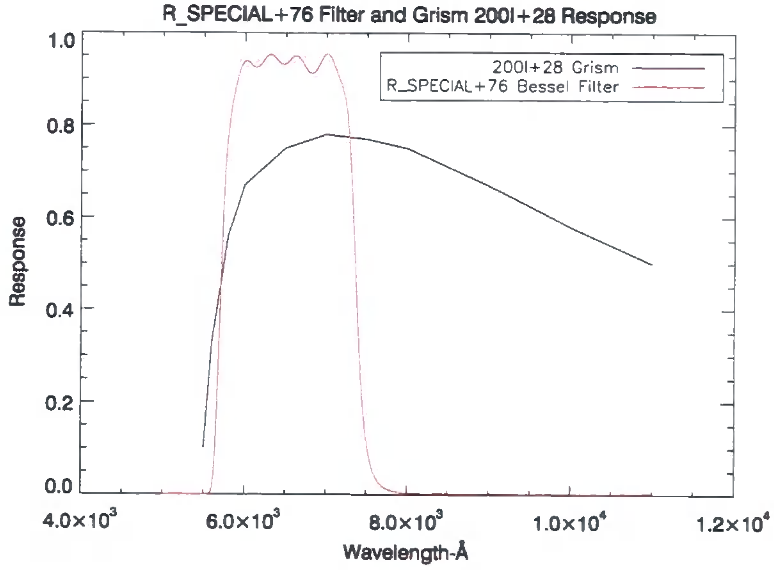


Figure 3.5 In order to target [OII] emission from galaxies with z ∼ 1 the low-resolution grism 200I+28 and the red filter 'R_SPECIAL+76' were used. This data is from Jehin and O'Brien (2007).

The plan for the first night of observation (MJD 53787) was to concentrate on

Table 3.2. The galactic spectroscopy that was carried out

MJD	Mask name	No. of Exposures	Exposure Time (s)	'Seeing'	Comments
53787	1127 "Bright" 1	3	650	1.7″	...
53787	1127 "Faint" 1	2	1650	1.4″	No continuum
53787	1127 "Faint" 2	2	1650	1.0″	Only a Few sources
53787	1127 "Bright" 2	3	650	1.0″	...
53787	1127 "Bright" 3	3	650	0.80″	...
53787	1127 "Bright" 4	3	650	0.80″	...
53787	1127 "Bright" 5	1	650	1.7″	...
53788	1122 "Bright" 1	2	1750	1.0″	TOO
53788	1122 "Bright" 2	2	1750	0.75″	...
53788	1127 "Bright" 1	2	1750	0.65″	...
53788	1127 "Bright" 2	1	1750	0.50″	...
53788	1127 "Bright" 3	1	1750	0.65″	...
53788	1127 "Bright" 4	1	1750	0.65″	...
53788	1127 "Bright" 5	2	1350	0.65″	...
53788	1122 "Bright" 3	2	1350	0.65″	...
53788	1122 "Bright" 4	2	1200	0.70″	...
53789	1122 "Bright" 1	2	1750	0.65″	...

PKS 1127-145 with 5 "faint" masks and 5 "bright". Very poor "seeing" conditions ($\sim 1.4''$) and high winds hampered the first night of observing and it was found that the integration times were insufficient for the "faint" masks in these conditions.

For the second night (MJD 53788) a decision was made to concentrate on the "bright" masks only, so the carousel was loaded with all 10 "bright" masks from both fields of view. The weather improved significantly together with the "seeing" which remained $\leq 1''$. During the first part of the second night there was a target-of-opportunity (TOO) that interrupted the first observation. Therefore it was agreed that 1122 "Bright" 1 mask would be observed again for a third night in service-mode. Two exposures of 1750 seconds each were carried out in good conditions. What was observed is summarised in Table 3.2. The calibration data taken is summarised in Table 3.3.

The pixels were binned 2×2 prior to readout so each image contained 1024×2048 pixels. Higher resolution was unnecessary and avoided as this would have led to a higher read-noise and over-sampling of the slit-spread-function along the dispersion axis.

Table 3.3. The calibration data for the galactic spectroscopy that was taken

Category	No. of Exposures	Exposure time (s)
Bias	5 per night	0.0
Spectroscopic Standard	1 per night	150.0 maximum
Spectroscopic flat	5 per mask per night	2.0
He-Ar arc	1 per mask per night	70.0

3.5 The Data Reduction of the FORS2 Data Set

At the time of reduction no ESO pipeline existed for the swift reduction of FORS2 MXU data so the extraction of galaxy spectra was mainly performed using IRAF, with predominant use of the IMUTIL and NOAO packages.

3.5.1 Preparation and Cleaning of the Frames

- The first step was to determine what area on each chip contained the science data and what area was the overscan region. To prevent negative pixel values each element has a pre-image pedestal charge of ~ 100 ADUs. When the CCD is read the hardware continues to measure the charge of imaginary pixels and it is these that measure the size of this pre-exposure off-set. A bias frame from each chip was studied using IMPLOT and this region was found to be in the top 5 rows of every frame. The IRAF program COLBIAS in the NOAO package was used to remove this off-set in every exposure. The median pixel value in each column of the overscan region was subtracted from every pixel in that column. Each frame was then cropped cutting off the overscan.

- The next step was to remove the variation in the bias level across the chip. Each night and for both chips there were 5 bias images taken. Each set of 5 were co-added to give their mean value using IMCOMBINE. 3σ clipping was used when combining the images to avoid anomalous/extreme pixel values. Each master bias frame was then subtracted from every other associated image.

No dark calibration images were taken. The dark current of the FORS2 detector is only $\sim$ 3 e^- $pixel^{-1}$ h^{-1}, so any dark e^- would have added an insignificant contribution to the pixel count.

It was also necessary to clean the images. The software DS9 was used to inspect the frames and check for any defective pixels. This led to the creation of a pixel mask, an array of 1s and 0s marking out those pixels that needed removing. On all of the frames the IRAF program FIXPIX then interpolated over these regions with the median value of the pixels either side of the blemish. For example, in chip 2 a section of row 311 had "hot-pixels". This was a defect where the saturation level of each pixel was very low. This horizontal line was interpolated over along the vertical direction.

The decision to split up science frames into several sub-exposures and then stack the results was mainly taken to minimise contamination by cosmic rays. The majority of these were removed using COSMICRAYS from the CRUTIL package. The identification of rays and their removal was controlled by 4 parameters.

- Window size - This was the size of the box in which the mean pixel value was calculated for cosmic ray detection. A box length of 7 pixels was used.
- Threshold - This was the minimum value above the mean value of the surrounding pixels at which a cosmic ray was identified within a window. This value depended on the noise of the image and how weak the cosmic rays were. Based on trial and error an optimum value of 10 ADU was chosen.
- Fluxratio - This parameter was the ratio of the mean neighboring pixel to the cosmic ray pixel and set the level at which a possible cosmic ray, found by the threshold limit, was rejected. Again a method of trial and error was implemented, where "before and after" images were blinked in DS9 in order to see what value to use. A ratio of 10 was sufficient to remove the majority of the cosmic rays.
- Number of passes - This was the number of times that the 7 x 7 pixel window was passed over the image while recursively removing cosmic rays. 8 passes

were made. This number was a compromise between the proportion of rays removed and the time the process took to complete.

This procedure proved adequate in removing most of the cosmic rays. However significant cosmic rays persisted in regions with strong emission features. Rather then interfere with the science, particularly any absorption or emission lines, these cosmicrays were left in. Any feature that made it to the final 1-D extraction could later be identified and removed by interpolation.

3.5.2 Flattening, Distortion Correction and Wavelength Calibration

Once the images were cleaned the IRAF task IMTILE was used to join both files for the two chips in every frame. This included each science frame together with the 5 corresponding flat and Helium-Argon arc lamp exposures. The 5 dome-flat exposures for each mask were then co-added using the median option in IMCOMBINE to make a master flat. The median pixel rather than mean was used to avoid the result being effected by any extremely high or low values. Each of these frames was also divided by their exposure time to give the pixels in ADU s^{-1}.

Each science exposure together with a flat and arc frame were then sent to a generic spectroscopy reduction pipeline written by D. Kelson (Kelson, 2003).[3]

First of all this software divided the science images, containing the 2D spectra, by a master flat. This was done to correct for the illumination variation across the spectrum caused by pixel-to-pixel gain variations, the blaze angle and the wavelength sensitivity.

The 2D spectra were then transformed to correct for any distortion along the spatial and dispersion axes. These were caused by defects, such as the instrument flexure under gravity, or the optical design that produce curved spectra.

Using the He-Ar arc frame the program then found the dispersion relation and so gave a wavelength calibration for the 2D spectra.

---

[3] D. Kelson, The Observatories, Carnegie Institute of Washington, 813 Santa Barbara Street Pasedena, California 91101 USA.

All that was required was a data file defining the location of the apertures for each mask, and a line-list with the standard wavelengths and relative line strengths of the arc frame. This was available from the ESO FORS usermanual (Jehin and O'Brien, 2007). Figure 3.6 shows the wavelengths and relative intensities of the He-Ar lines used when calibrating the dispersion solution.

Standard IRAF tasks, for instance APALL, do exist for these procedures with more control permitted for the user at intermediate stages. However the Kelson software was more effective at removing the x/y-distortions and providing a wavelength calibration. For example within IRAF the distortion correction is done slit by slit by fitting a low-order polynomial to the exposed spectrum. The continuum in my data was very faint, therefore fitting an appropriate function was very difficult and time consuming. Another advantage was that numerous frames for the same mask could be processed at once without the need to manually trace the continuum for each aperture or redefine the wavelength calibration.

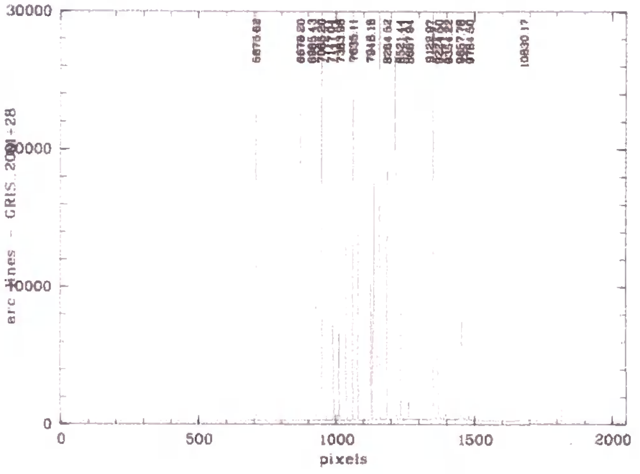


Figure 3.6 The wavelengths and relative intensity of the He-Ar arc-lamp emission lines used in the wavelength calibration. The lack of emission lines to the left of 700 pixels was not a problem. The line at ~ 700 pixels is at 5875.62 Å and this was very close (~ 300 Å) to the blue cut-off of 5600 at Å. This image is from Jehin and O'Brien (2007).

Where APALL proved to be more useful was in the sky-subtraction and 1-D extraction. So it was decided that parts of both codes would be used in the reduction of the full data set. The Kelson pipeline software produced the distortion corrected, wavelength calibrated 2-D spectra, APALL was then used to sum along the spatial axis to produce 1-D spectra.

3.5.3 Fitting the Continuum, Sky-subtraction and 1-D Extraction

The calibrated 2-D science frame was sent to the APALL program in IRAF. This has the capability to run automatically with a vast choice of pre-set parameters. An initial extraction was run interactively to decide these values.

This process was run for each slit on an individual basis. First of all we had to locate the pixels containing the galaxy spectra. The summed pixel value as a function of the spatial direction across each spectrum for a narrow bandpass was plotted. This generally contained a dominant central peak caused by the galaxy continuum, ~ 5 pixels across, with a plateau either side that were the sky background counts.

We were then asked to give the lower and upper bounds of this peak in the spatial direction in order that the pixels containing the continuum could be identified. To each side of this peak two regions of ~ 5 pixels were also defined to estimate a median value for the sky background. This median was then subtracted from all pixels along the spatial axis to remove the sky-background level and sky lines.

A plot was then drawn that traced the galaxy continuum along the dispersion axis. The aim was to define a function that would fit these points. Linear functions were chosen as any distortion had already been corrected. The pixels containing the galaxy continuum were then summed vertically, collapsing along the spatial direction producing a 1-D spectrum.

APALL was set to do a variance weighted extraction. Within the image header were details of the read-noise and gain. Using these IRAF was able to find the associated Poisson error of each pixel using equation (3.1). When the pixels were summed along the spatial axis their contribution to the flux was weighted by this estimate in the error. This was of particular benefit as it enhanced emission and absorption features, while suppressing those pixels with a low S/N.

The output of APALL for each slit were 4 1-dimensional arrays. The weighted ob-

Table 3.4. The standard white dwarf stars observed for flux calibration

MJD	Standard[a]	RA (J2000)	DEC (J2000)	V_{mag}	Spectral type
53787	SA95-42	03 53 43.6	−00 04 33.0	15.61	DA
53788	GD50	03 48 50.1	−00 58 33.8	14.06	DA2
53789	LTT4816	12 38 49.4	−49 47 59.6	13.79	DA

[a] The stars catalogue data is from Hamuy et al. (1994), Oke (1990)

ject spectrum, the unweighted object spectrum, a plot showing the sky background that was subtracted and a σ- spectrum, all as a function of wavelength.

3.5.4 Flux Calibration

As well as the galaxies spectra of standard white dwarf stars with known photometry were also observed using the 200I grism and R_SPECIAL filter. This was to determine the instrument response in order that a flux calibration could be carried out. The details of those stars observed over the 3 nights are in Table 3.4.

The spectra for each standard star was extracted using the exact same technique described above. This was important as the aim was to measure the sensitivity of the instrument by comparing measured standard spectra with their literature profile. This function, how absolute flux changes with wavelength, was then applied to the galaxies to give flux calibrated spectra.

Figure 3.7 shows the SEDs (Spectral Energy Distribution) of the 3 standard stars used. The 1-D standard spectra were imported into the IRAF task STANDARD together with these literature curves. The flux through a set bandpass (the width of 1 pixel) of the standard spectra was then summed and compared with the same wavelength range of the standard SED. Parameters used within the program included the airmass, exposure time of the observation and the magnitude and effective temperature of the standard star. This program set the absolute scale of the flux by defining a zero-point per unit wavelength of 3.68×10^{-20} ergs $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{\AA}^{-1}$.

This calibration data was then passed to the program SENSFUNC. The principle of SENSFUNC was to find the function that would transform the uncalibrated

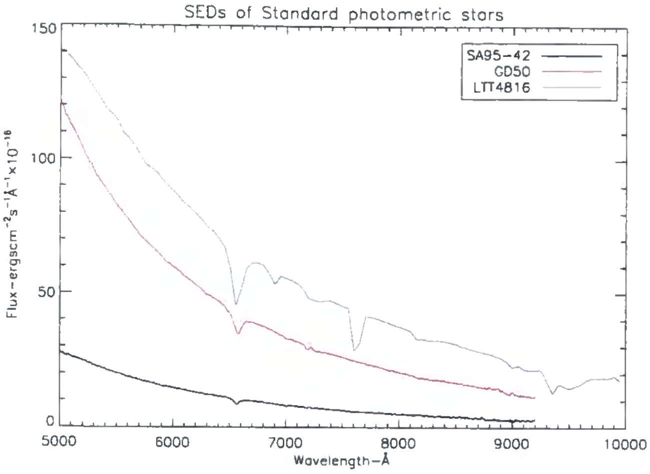


Figure 3.7 The spectra of the 3 standard White Dwarf stars observed, the data from this plot was used in the IRAF task STANDARD in order to flux-calibrate the galaxy spectra. Details for the star SEDs are from Hamuy et al. (1994), Oke (1990).

star spectrum into that found by the output from STANDARD. A 5th order spline function was chosen to do this. The inverse of this function was then the sensitivity function of the FORS2 detector. The best possible function was found by combining the data points from all 3 standard observations and plotting a best fit-curve. Each 1-D galaxy spectrum was then multiplied by this inverse-function to produce flux calibrated spectra.

3.5.5 Redshift Determination of the Emission and Absorption Galaxies

The final step before identifying spectral lines was to combine the separate exposures. Only spectra from certain masks were combined because of the large difference in the quality of the data taken over the 3 nights. For example the 2nd "bright" mask of the field PKS 1127-145 had 3 short 650 s exposures taken on the first night, these produced no usable data. Only the 1 exposure of 1750 s taken on the second night was used.

Spectra were co-added using a variance weighted algorithm. In the case where 3 or more spectra were being combined 3σ clipping was also used to avoid extreme values. It was at this stage that any cosmic rays, left in the exposure since the images were cleaned, were removed by linear interpolation. The spectra were imported into

Table 3.5. The emission and absorption line wavelengths used when determining galactic redshifts

Emission species	Rest-wavelength (Å)	Absorption species	Rest-wavelength (Å)
[O II]	3727.3	K	3933.7
Hβ	4861.3	H	3968.5
[O III]	4959.0	CaI	4226.7
[O III]	5007.0	G	4304.4
[N II]	6548.1	Hβ	4861.3
Hα	6562.8		
[N II]	6583.0		

the package SPLOT and anomalous peaks were cut by setting the flux level equal to the continuum.

The σ-spectra were also combined, their contributions summed by equation (3.3).

$$\sigma^2_{\rm total} = \left(\Sigma^{\rm n}_{\rm i=1}\left(\frac{1}{\sigma^2_{\rm i}}\right)\right)^{-1} \qquad (3.3)$$

Identification of emission lines

The plotting package SPLOT and FITS viewer DS9 were then used to identify emission lines by eye in the 1-D and 2-D spectra. Table 3.6 shows the rest-wavelengths of those lines used in calculating the galaxy redshifts.

Upon successful identification of any spectral features the galaxy was sent to the NOAO program RVIDLINES. Each feature was fitted with a Gaussian profile that was centred on the emission peak. The FWHM of the profile was set to a default 4 pixels (9.72 Å), with a 5 pixel minimum gap between 2 features.

A variance weighted mean redshift was then computed using all of the features identified. A RMS error was also calculated using the redshift of each feature $z_{\rm feature}$ and $\bar{z}$ in equation (3.4).

$$\sigma_{\rm z} = \sqrt{\frac{\sum^{\rm n}_{\rm i=1}(z_{\rm feature} - \bar{z})^2}{\rm n - 1}} \qquad (3.4)$$

Therefore no RMS error was defined for redshifts determined using a only a single line.

Within the header of the 1-D spectra were the universal time of the observation and the location of the Paranal observatory. RVIDLINES used these values to compute a heliocentric correction to the redshift. All of this information was then printed to a log file so that all the data could be passed to the final galaxy catalogue. Figure 3.8 shows the combined spectra of exposures 4 and 5 of slit 11 from PKS 1127-145 "bright" mask 1. The heliocentric redshift $z = 0.94896 \pm 0.00024$ was calculated using the [O II], H_β, and [O III] doublet emission lines.

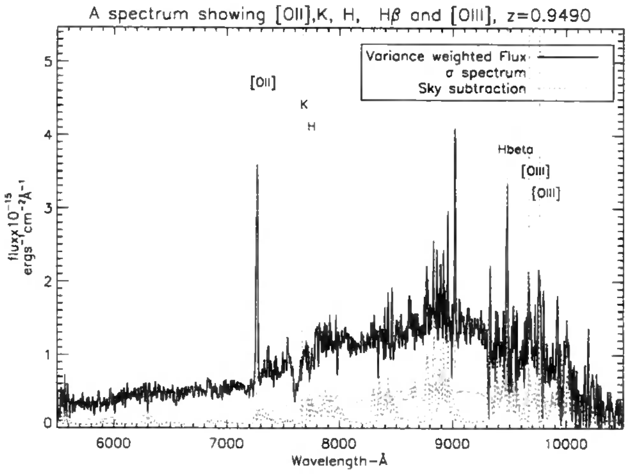


Figure 3.8 A variance weighted combined spectrum of slit 11 from "bright" mask 1 of sightline PKS 1127-145. The redshift of this emission line dominated galaxy was determined using [O II], H_β, and the [O III] doublet. The feature at ~ 7600 Å is atmospheric O_2 absorption. The anomalous lines at ~ 9000 Å are a product of poor sky-subtraction. The flux calibration beyond $\lambda \geq 10000$ Å is also not reliable. The σ-spectrum (red) and sky value that was subtracted (green) are also shown. The value of the sky subtraction has been divided by 100 to scale it to the source intensity.

Using this method the redshifts of 187 emission line dominated galaxies were found.

Identification of absorption lines

The problem of poor signal-to-noise (S/N) and very weak continua meant identification of absorption dominated galaxies was not possible by inspection. Instead the IRAF program XCSAO was used (Kurtz and Mink, 1998). This is where absorption features were identified by cross-correlating with a template absorption spectrum using a fast-Fourier transform (FFT) algorithm. XCSAO is an updated IRAF program originally devised by Tonry and Davis (1979). Figure 3.9 shows the high S/N template.

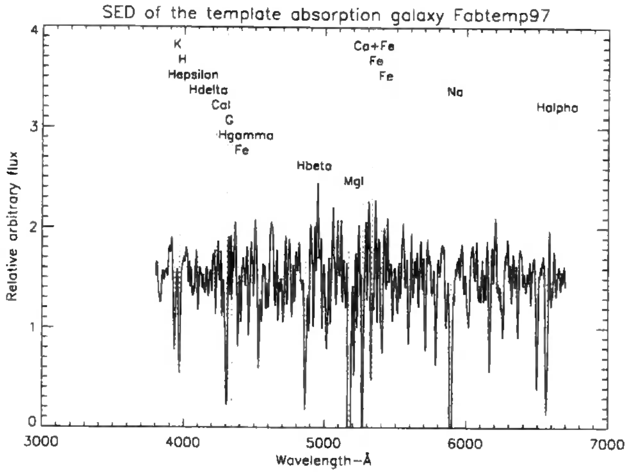


Figure 3.9 The spectrum for the template galaxy fabtemp97. This template is the mean of 1489 galaxies taken with the FAST spectrograph at the Mt. Wilson telescope, they have been co-added with their continua and emission lines removed (Kurtz and Mink, 1998). This was used when cross-correlating absorption lines using XCSAO. The template has been transformed to zero redshift and the common absorption lines have been marked.

When cross-correlating a potential absorption galaxy the following steps were taken.

- Cropping the object spectrum. All of the galaxy spectra suffered from decreasing S/N above 9000 Å and near to the 5500 Å blue cut-off, so a bandpass of 6000 Å $\leq \lambda \leq$ 9000 Å was chosen. Any emission lines were also removed and the continuum subtracted.

- Apodisation. This is where the flux at the ends of the spectrum were smoothly set to zero by multiplying by a cosine function. Any abrupt cut-off in the spectrum would produce high frequency components in Fourier space as the intensity dropped to zero over an infinitesimal bandpass. The ends of the spectrum were given a diminishing flux 5% from each end.

- Both the template and object spectrum were then rebinned to a common dispersion solution comprising of 2048 pixels and then Fourier transformed.

Table 3.6. Confidence levels and criteria needed for the galactic redshifts

Confidence Level	Comment	No. of galaxies
1	Multiple lines (3+) with strong lines (usually AGN)	54
2	Multiple lines (2+) with moderate intensity	34
3	Multiple lines (2+) faint lines	28
4	Multiple lines (2+) ambiguous lines	59
5	Single line redshifts (usually [O II])	25

- Each spectrum was then apodised in Fourier space so that extreme frequencies would not dominate. A high frequency cut-off (parameter nrun) was used as this removed data that contributed to the Poisson noise. A low-frequency cut-off was made (low_bin) as this removed signal that was dominated by the large-scale structure of the continuum. The parameters top_low and top_nrun described the frequencies at which the power tapered to zero using a cosine filter. The frequencies to use for the 4 parameters, low_bin, top_low, top_nrun and nrun were decided on a trial basis. The tests were done using an absorption galaxy with an accurately known redshift, NGC 4884 (z= 0.021632).
- The Fourier transforms T and G were then multiplied together to give the Fourier-transform of the cross-correlation function $t \cdot g$
  $\mathcal{F}(t_{\text{template}} \cdot g_{\text{galaxy}}) = T_{\text{template}} * G_{\text{galaxy}}$.

The peak in the correlation plot gave the velocity difference between the object and template spectrum, while the relative height of this peak gave the confidence level of the solution. A confidence of $R \geq 3$ was accepted as a reliable velocity difference. Unlike RVIDLINES, XCSAO does not produce a heliocentric velocity correction as the result is the velocity separation between the template galaxy and object. Therefore XCSAO was used to identify any absorption lines. These absorption lines were then used in RVIDLINES to give redshifts for the 13 absorption line dominated galaxies that were included in the final table.

Each of the 200 galaxies that made up the final table were then assigned a confidence level based on the number and strength of any features. The criteria for each level are shown in Table 3.6.

3.6 The Luminosity Distance and K-Correction

Using the standard ΛCDM cosmology stated in Section 1.2.5 the luminosity distance to each galaxy (D_L) was found using Mattig's formula (Peacock, 1999).

$$D_L(z) = \frac{2c}{H_0\Omega^2}\left(\Omega z + (\Omega - 2)(\sqrt{1+\Omega z} - 1)\right) \tag{3.5}$$

The absolute magnitude M_B of each galaxy was then computed.

$$M_B = m_R - 5\log_{10}\left(\frac{D_L(z)}{10pc}\right) + K_{correct} \tag{3.6}$$

m_R took the value of MAG AUTO that was computed when Source Extractor was run on the pre-imaging. As the pre-imaging was only exposed through the red Bessel filter a theoretical K-correction had to be computed using the response curves of two filters and the spectral energy distributions (SEDs) of template galaxies. Details of the code used for the colour correction are described in Bruzual and Charlot (2003).

Figure 3.10 shows the response curves of the two filters. The red filter is the ESO R_SPECIAL+76, the blue Bessel filter B-Bessel+74 shown alongside is also located on the FORS2 instrument. The rest-frame B-band magnitude was selected because it is closest to the observed R-band for galaxies at a redshift of $z \sim 0.80$, close to the median redshift in our galaxy catalogue.

The SED templates are shown in Figure 3.11. These are described in Bruzual and Charlot (2003) and cover 5 spectral types, from the passive elliptical galaxies Es0 (black SED), to the star-forming Sdm type (pink SED) where most of the flux is transmitted in the UV.

To compute the B − R K-correction for each spectral type the theoretical flux that would be transmitted through the rest-frame B filter was computed. Each SED was then transformed to a higher redshift in steps of $\Delta z = 0.005$, and the flux through the red filter was integrated. The output from the code was a table, with a value for the K-correction listed as a function of redshift and spectral type. Galactic evolution was not taken into account when transforming the SEDs. With no clear indication from the galaxy spectra as to what spectral type each galaxy

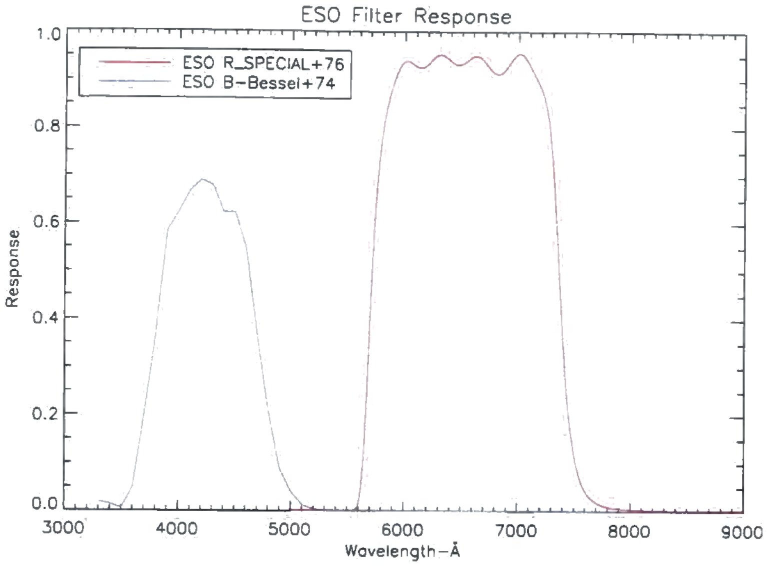


Figure 3.10 The filter responses of two FORS2 Bessel filters. The red filter is the R_SPECIAL+76 used in the FORS2 spectroscopy. The blue filter, B_BESS+74, describes the filter used for the synthetic B − R colour. Data is from Jehin and O'Brien (2007).

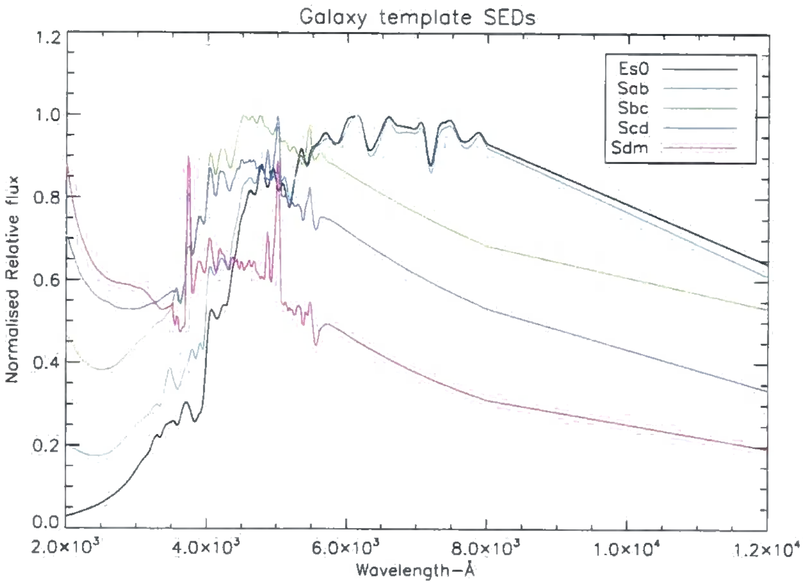


Figure 3.11 The SED templates for the 5 galaxies used in the K-correction code. The spectral types range from the elliptical Es0 galaxies described by the black line, to the star-forming Sdm galaxies (pink line) where most of the flux is transmitted at UV wavelengths. These templates are summarised in Bruzual and Charlot (2003).

was the B – R colour for every template was considered. Therefore we were able to tabulate 5 results, giving a minimum and maximum possible value for M_B. Figure 3.12 shows the mean value of M_B for all of the 200 galaxies in our catalogue plotted as a function of redshift.

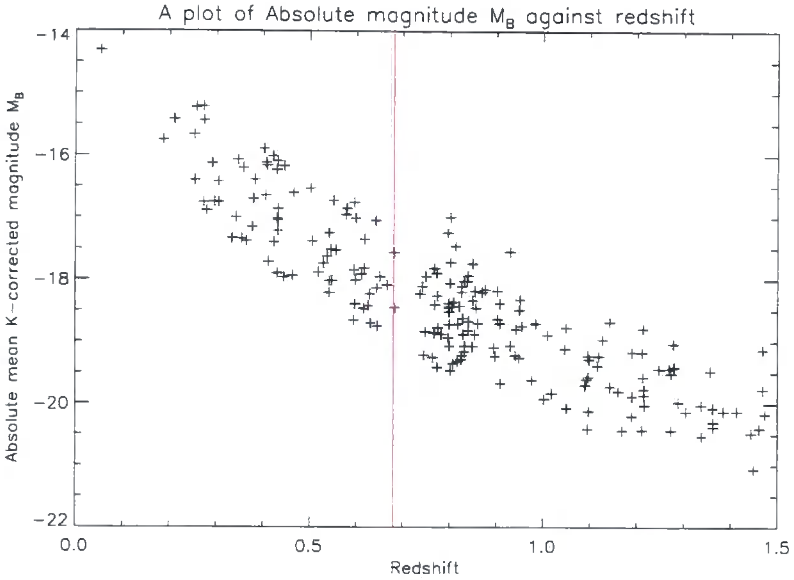


Figure 3.12 The mean absolute magnitude M_B, calculated from the 5 SED galaxy templates and theoretical K-correction. The median galaxy redshift of the catalogue was z = 0.803.

3.7 The Final Galaxy Catalogue

Altogether the spectroscopic redshifts of 200 galaxies were found and were added to a master catalogue, with a median redshift of z = 0.803. 200 galactic redshifts is lower than what we would have preferred, after all we are trying to maximise the number of galaxies within the redshift range 0.68 and 1.55 in order to correlate their position with Lyα absorbers. The number of galaxies that could be used for this was only 127. The lower number of galaxies at high redshift can be attributed to the poor "seeing" conditions when observing the faint $23.5 \leq m_R \leq 24.5$ targets. The details of all 200 galaxies in both quasar fields are given in Appendix A.

Chapter 4

The Lyα-Absorbers in UV and Optical Spectroscopy

4.1 The Reason for UV Spectroscopy

The Lyα absorbers used in the cross-correlation function were detected in the absorption spectra of the background quasars PKS 1127-145 (z = 1.187) and HE 1122-1648 (z = 2.404) (Bechtold et al., 2002, Carswell et al., 2002a). These quasars were selected because high resolution (R= 30000) STIS echelle spectroscopy exists within the HST archive that covers the wavelength range 2050-3111 Å, the range of wavelengths needed when locating Lyα lines at a redshift of z $\sim$ 1. With such a high dispersion it was hoped that low column density ($\log_{10}(N_{HI}\,(cm^{-2}))\sim$ 13) H I absorption lines could be resolved and Voigt profiles fitted in order that their column density and b-parameter could be determined.

HST data is required because in order to examine the Lyα-galaxy relationship at low to intermediate redshifts one must look at UV spectroscopy (as was utilised in Morris and Jannuzi (2006)). Here the correlation was investigated by comparing 635 galaxy redshifts out to z $\leq$ 1 with 379 Lyα absorbers located along 16 lines-of-sight. These were extracted from low-resolution HST FOS spectra using the 'high' resolution gratings G190H and G270H (resolving power $\sim$ 1300) and the G160L grating (resolving power $\sim$ 250). However with such a low resolution the b-parameter must be fixed, a mean b-parameter for all absorbers of 30 km s^{-1} was assumed.

The advantages of using STIS echelle rather than the FOS long-slit data were;

the instrumental broadening decreased from 231 to 10 km s^{-1}, this allowed more absorbers per spectra of a lower column density ($\log_{10}$ (N_{HI} (cm^{-2}))$\lesssim$ 13.5) to be detected. As metal lines have a narrower intrinsic b-parameter more and a greater variety of metals (C IV, O VI) could be found which allowed a more complete line-list. The Doppler broadening could also be measured which permitted a more realistic model when cataloguing the large variety of Lyα absorbers in the IGM.

4.2 The Instruments and Archive Data Used

4.2.1 The Archive Data

The only Lyα absorbers that were used in the correlation function were detected using the high resolution instrument the Space-Telescope-Imaging-Spectrograph (STIS) in echelle mode with the E230M grating. In addition to these, to help identify lines and to compile a more complete line-list, spectra were also analysed that were taken by STIS in long-slit mode, the low-resolution Faint-Object-Spectrograph (FOS) and the ESO UVES (Ultraviolet and Visual Echelle Spectrograph) instrument [1] (Dekker et al., 2000).

Table 4.1 lists the data products used in the investigation that were collected from the STScI (Space Telescope Science Institute) and ESO (European Southern Observatory) archives (Kaufer et al., 2007, Keyes et al., 1995, Kim Quijano et al., 2007).

4.2.2 The Long-slit and Echelle Spectrographs

All spectrographs are based on the same principle. Incoming photons go through the entrance slit which is at the focal point of the telescope. After this point the light will diverge, but it is passed through a collimator before it then strikes the grating. The interference pattern then passes through a camera before impacting on the detector.

---

[1] Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. ESO programme numbers 067.A-0567(A), 068.A-0570(A) and 069.A-0371(A).

Table 4.1. Data that was used from the ESO and HST archive for the QSOs HE 1122-1648 and PKS 1127-145

Instrument + Grating	wavelength (Å)	Exposure time (s)	Resolving power $\frac{\lambda}{\Delta\lambda}$
		HE 1122-1648	
STIS E230M [a]	2270-3111	4x2877	30000
STIS E230M [a]	1574-2382	6x2712	30000
STIS G430L [a]	2900-5700	2x1000	1040
FOS G190H [b]	1571-2311	3x1873	1300
UVES Blue/Red [c]	3200-5765	15x5400	45000
		PKS 1127-145	
STIS E230M [d]	2270-3111	14x5798	30000
FOS G190H [e]	1571-2311	4x2160	1300
FOS G160L [f]	901-2508	1x1450	250
UVES Blue/Red [g]	3200-6809	10X3060+4x4800	45000

[a] PI Baldwin HST proposal 9885

[b] PI Reimers HST proposal 5950

[c] PI Savaglio (Schaye et al., 2007)

[d] PI Bechtold HST proposal 9173

[e] PI Deharveng HST proposal 3483; (Deharveng et al., 1995)

[f] PI Rao HST proposal 6577; (Pettini et al., 1999)

[g] PI Lane,PI Savaglio (Kacprzak et al., 2007)

Gratings are characterised by a length L, the distance from the centre of each facet to the next facet centre (d) and the facet width (s). Consider a pair of incoming light rays at an incident angle α to the grating normal that reflect at an angle β. Parallel reflections whose points of incidence on the grating were a distance d apart with a phase-difference $\phi_\mathrm{d} = \mathrm{n}\pi$, $(\mathrm{n} \in \mathbb{Z})$ would have maximum constructive interference (Figure 4.1a).

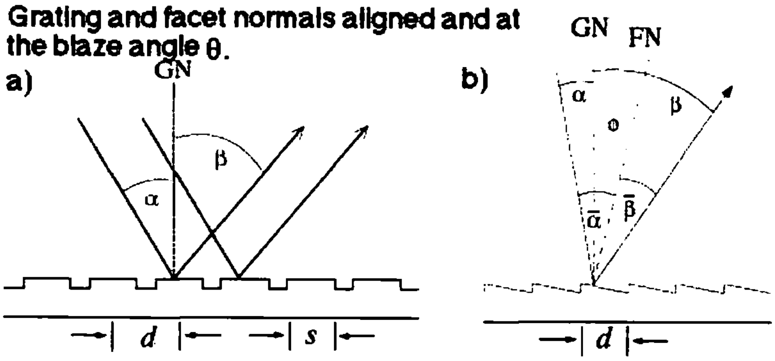


Figure 4.1 When the normal to both the facets and grating are aligned (figure a) then the intensity of the interference pattern will decrease for spectral order n as $1/\mathrm{n}^2$. By cutting the facets at the blaze angle θ (figure b) $\bar{\alpha} = -\bar{\beta}$ for all n. This means the interference pattern will not diminish when trying to observe a higher spectral order. These figures are from Churchill (2007).

The grating equation (4.1c) gives the incident angle, the reflecting angle and the grating size d required to target a particular spectral order (n) and wavelength regime (Churchill, 2007). For a phase difference ϕ_d:

$$\phi_\mathrm{d} = \frac{\pi \mathrm{d}}{\lambda}\left(\sin\alpha + \sin\beta\right) \tag{4.1a}$$

$$\text{So for the maxima when } \phi_\mathrm{d} = n\pi \tag{4.1b}$$

$$\frac{\mathrm{n}\lambda}{\mathrm{d}} = \sin\alpha + \sin\beta \tag{4.1c}$$

For mono-chromatic light the interference pattern would be a periodic series of maxima that are a phase difference $\lambda/\mathrm{d} = \mathrm{n}\pi$ apart and have a width of λ/L. This function is plotted in Figure 4.2a.

This is of course simplified as not only will the light have passed through a

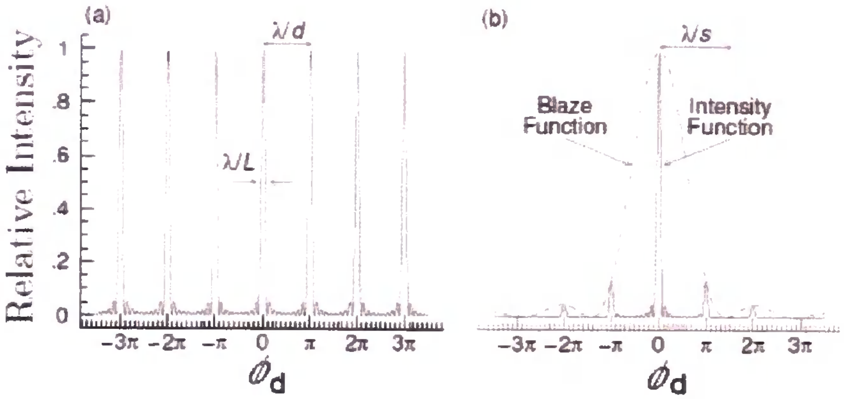


Figure 4.2 (a) This is the facet centre to facet centre interference function derived from equation (4.1c). Each peak (drawn much wider here than for a practical grating) is caused by rays with a phase difference $\phi_\mathrm{d} = \mathrm{n}\lambda/\mathrm{d} = \mathrm{n}\pi$ apart each of width λ/L. (b) The facet centre to facet edge interference creates the blaze function with minima at $\phi_\mathrm{s} = \mathrm{n}\lambda/\mathrm{s}$. This has also been plotted as a function of ϕ_d. The overall intensity now decreases as $1/\mathrm{n}^2$ so the maximum intensity is only at zeroth order. These figures are from Churchill (2007).

bandpass and not be mono-chromatic, the interference function will also have a term gained from the phase difference (ϕ_s) between the centre and the edge of each facet (over the distance s).

This is the non-periodic blaze function described in Figure 4.2b. The maximum is at zeroth order, $\phi_\mathrm{s} = \phi_\mathrm{d} = 0$ when $\alpha = -\beta$ and minima are at phase-differences of $\phi_\mathrm{s} = \mathrm{n}\lambda/\mathrm{s}$.

One cannot use the zero order spectrum which would provide the maximum intensity as there would be zero dispersion. However the relative intensity of the n^{th} order now goes as $1/\mathrm{n}^2$ which greatly reduces the efficiency of the instrument. Equation (4.2c) gives the overall expected intensity for order n (Churchill, 2007).

$$I_\lambda \propto I_{\mathrm{n\ Interference}} \cdot I_{\mathrm{n\ Blaze}} \tag{4.2a}$$

$$\propto \frac{\sin^2\left(\frac{\mathrm{L}}{\mathrm{d}}\phi_\mathrm{d}\right)}{\sin^2\phi_\mathrm{d}} \cdot \frac{\sin^2\phi_\mathrm{s}}{\phi_\mathrm{s}^2} \tag{4.2b}$$

$$\text{So for order n: } I_\mathrm{n} \propto \frac{\sin^2\left(\mathrm{n}\pi\frac{\mathrm{s}}{\mathrm{d}}\right)}{\left(\mathrm{n}\pi\frac{\mathrm{s}}{\mathrm{d}}\right)^2} \tag{4.2c}$$

What is needed is a phase shift in the blaze function which would translate the function in Figure 4.2b to the required order. The equation that defines the function in Figure 4.2b is only valid if the normal to both the facets and the grating are aligned. If the grooves are cut with an angle θ, called the blaze angle, relative to the plane of the grating then the two normals will no longer lie along the same axis. This is illustrated in Figure 4.1.

The blaze angle θ between the two normals in Figure 4.1b is the necessary phase shift. The angles of incidence and reflection relative to the facet now obey the relation $\bar{\alpha} = -\bar{\beta}$, where $\bar{\alpha} = \alpha - \theta$ and $\bar{\beta} = \beta - \theta$. Therefore $\phi_{\mathrm{s}} = 0$ for all n and I_{n} will have the maximum possible intensity.

The differentiation of equation (4.1c) gives:

$$\frac{d\beta}{d\lambda} = \frac{\mathrm{n}}{\mathrm{d}\cos\beta} = \frac{\sin\alpha + \sin\beta}{\lambda\cos\beta} \tag{4.3}$$

Therefore it is the constant term n/d that sets the dispersion relation and hence their choice sets the instrument resolution. So the ratio n/d is small for long-slit spectra which can only cover a short band-pass and target low orders where $\mathrm{n} \leq 3$.

A further problem arises when doing spectroscopy. Equation (4.1c) has multiple solutions with $(\mathrm{n}-1)\,\lambda_{\mathrm{n}-1} = \mathrm{n}\lambda_{\mathrm{n}}$ and all of these orders will be superimposed. To remove unwanted wavelengths when doing long-slit spectroscopy the light passes through a narrow bandpass filter before it enters the spectrograph.

Echelle spectrographs separate these orders using a secondary grating, a cross-disperser set at 90° to the first. This separates the superimposed spectra along the spatial axis so a higher dispersion is possible with a longer wavelength coverage. For this reason Echelle spectrographs target high orders $20 \leq \mathrm{n} \leq 100$ with a lower grating density (Kaufer et al., 2007).

The Faint-Object Spectrograph (FOS)

The main difference between the FOS, STIS and UVES spectrographs that were used was in the detectors. UV photons were captured using a cathode to generate photo-electrons, while the optical detectors were CCDs.

The UV ($\sim$ 900 $\rightarrow$ 2500 Å) FOS spectra were detected using the blue (for PKS 1127-145 observations) and red (for HE 1122-1648) Digicon detectors (Keyes et al., 1995). Incoming UV photons strike a 2D photocathode, this serves to ensure that there is a minimum frequency cut-off, because for each UV photon there will be orders of magnitude more photons at optical wavelengths.

The photo-electrons are then accelerated to a 1D array of 512 diodes. The number of detectors per resolution element for the high-resolution ($\lambda/\Delta\lambda$ = 1300) FOS data was 1-6 Å diode^{-1}. This increased to 25 detectors for the low-resolution (250) data.

The FOS data was collected in ACCUM mode. This is where the electrons gathered in the pixel bins for up to 4 minutes, (this depended on the magnitude of the target and the background count rate so that the detector did saturate) before being summed to the total integrated count rate. Not being a CCD there was no readnoise, but there was a dark count rate. This was approximately 0.010 and 0.007 counts s^{-1} diode^{-1} for the red and blue detectors respectively, and this set a lower limit for the minimum detection threshold.

The Space-Telescope-Imaging-Spectrograph (STIS)

STIS is capable of performing both UV and optical spectroscopy (Kim Quijano et al., 2007). The much higher spatial resolution and detection limit of STIS when compared to FOS is made possible because of the MAMA (Multi-Anode Microchannel Array) detectors (Joseph et al., 1995). These are a 2D array of photon-counting detectors, similar in principle to a photo-multiplier tube.

In the UV detector a photon strikes the Cs_2Te photo-cathode which frees a single photo-electron. This is then accelerated towards a microchannel plate, a piece of lead-oxide glass which is comprised of curved individual tunnels. Within this channel an electron cascade causes a pulse of $\sim 4 \times 10^5$ electrons to then strike the detector. The detector is made up of 2048 × 2048 pixels which clock the electron count rate. Like the FOS raw data STIS data was also collected in ACCUM mode. This raw count rate was then binned 2 × 2 to give an array of 1024 × 1024 pixel values, which in the same manner as a CCD were stored as a 16-bit numbers. Unlike a CCD this photon-counting method did not have a read-noise. The instrument dark current, primarily caused by impurities in the detector was $\leq 1.7 \times 10^{-3}$ counts sec^{-1} pixel^{-1}.

The obvious advantage over the Digicon detectors was that STIS could observe along the spatial dimension. This allowed $2.5'' \times 2.5''$ UV imaging, and meant pixels were available for order separation when doing echelle spectroscopy.

The STIS instrument also has a CCD detector which permits detection from the near UV all the way up to 11000 Å. However the quantum efficiency below 3000 Å is less than 20%. This was used for the long-slit G430L spectrum. The CCD is composed of 1024×1024 binned pixels with a dark current of 0.0009 e$^-$ sec^{-1} pixel^{-1}, a read-noise of 5.4 e$^-$ pixel^{-1} and a gain of 1.0 e$^-$ ADU^{-1}.

The Ultraviolet and Visual Echelle Spectrograph (UVES)

Finally the ESO UVES instrument is made up of two separate spectrographs, with a blue and red arm that cover the wavelength ranges $3000 \rightarrow 5000$ Å, and $4200 \rightarrow 11000$ Å respectively (Kaufer et al., 2007).

The large 76° blaze angle, high values of n and relatively low grating densities allows for very high (up to 100000) resolution spectroscopy over the entire wavelength range. To manage this large bandpass UVES has 3 CCD detectors each with 4096×2048 pixels. The one MIT-LL CCD that detects blue wavelengths has a dark current of 0.0001 e$^-$ sec^{-1} pixel^{-1}, a low-gain read-noise of 4.1 e$^-$ pixel^{-1} and a gain of 1.84 e$^-$ ADU^{-1}.

The two mosaiced red EEV CCDs (with a chip gap that covers 1 order at approximately the central wavelength) have a dark current of 0.0003 e$^-$ sec^{-1} pixel^{-1}, a read-noise of 3.8 e$^-$ pixel^{-1} and a gain of 1.6 e$^-$ ADU^{-1}.

4.3 Reduction of both Long-slit and Echelle Spectra

4.3.1 Reduction of the Raw Data Suite

The preliminary calibration and reduction of data for each instrument was performed by STScI or ESO 'on-the-fly' by an instrument specific pipeline. The steps performed in the initial reduction of raw data depended on the detector.

FOS calibration

The details for the FOS pipeline CALFOS are from Keyes et al. (1995) and they differ significantly from both the STIS and UVES pipelines as the data is already 1-dimensional. The pipeline began by computing the statistical error in the raw data, this was then carried forward through each reduction step and used in the variance spectrum. This was analogous to the detection of optical photons using a CCD as the variance in the flux count obeys Poisson statistics.

The total photon count was then divided by the exposure time to give the count rate.

Next the background sources of noise were subtracted. This included the dark current generated by the instrument and scattered light from the optics.

The variations in sensitivity with wavelength from diode-to-diode and the illumination correction were then compensated for by dividing each pixel by the flat-field response.

After each observation FOS would perform an arc exposure using the internal Pt-Cr-Ne lamp in order to find the wavelength zero-point. Vacuum wavelengths were used, and then the pipeline using the date of observation performed a heliocentric velocity correction.

Finally using the data from a spectrophotometric observation the pipeline flux calibrated the object spectrum, converting the count rate to ergs $\mathrm{s}^{-1}\ \mathrm{cm}^{-2}\ \mathrm{\AA}^{-1}$, thereby correcting for the telescope and spectrograph transmission and the detector efficiency.

STIS calibration

The initial stage of the pipeline CALSTIS varied depending on whether the data was taken using the MAMA or CCD detector. A detailed description of both pipelines is covered in Kim Quijano et al. (2007).

The first step when reducing UV data was to bin the pixels 2 × 2 from the raw 2048 × 2048 array to 1024 × 1024. Similar to the FOS pipeline, the next process was to calculate a statistical error, this error was then propagated through the reduction pipeline. One could model this on equation (4.5) below only the read-noise and bias terms do not exist.

The MAMA detectors saturated above 200000 photons s^{-1} across the detector. Thus if the count rate is too high one UV photon will produce less than the expected number of electrons in the cascade, and so this trend was corrected for at this stage. The dark current was then subtracted and the 2D spectra were divided by a flat field exposure created by a lamp within the spectrograph.

The STIS CCD data was reduced in the same manner as generic CCD optical data. The necessary steps are described in detail within Section 3.5.1. The bias level was subtracted and the overscan trimmed while cosmic rays were identified and removed. Cosmic rays were not removed for the UV data sets.

The images were then divided by the flat field and the amount of scattered light from the instrument was subtracted from the total. The final pixel count was found using equation (4.4).

$$[\mathrm{DATA}] = \frac{([\mathrm{RAWCOUNTS}] - \mathrm{OVERSCAN}) - <\mathrm{BIAS}>}{<\mathrm{FLAT}>} - [\mathrm{SCATTER}] \quad (4.4)$$

The pixel counts in ADUs was converted to electrons by multiplying by the detector gain to give the flux in electrons per pixel. The variance in this value was calculated by the Poisson error using the signal-to-noise (S/N) relation (see equation (4.5)).

If RN, R_{ij}, B_{ij}, and I_{ij} are the read-noise and gain corrected pixel counts for the raw science frame, bias frame and flat frame (that has been normalised to 1) respectively. Then variance (Churchill, 2007)

$$\sigma_{ij}^2 = \frac{1}{I_{ij}^2}\left(R_{ij} + RN^2 + \sigma_{Bij}^2 + \left(\frac{R_{ij} - B_{ij}}{I_{ij}}\right)^2 \sigma_{Iij}^2\right) \tag{4.5}$$

In which the $\sigma_{B,I}^2$ relations are the variance in the bias and flat frames for a box-sample of size n that is centred on pixel ij.

$$\begin{aligned} \sigma_{Bij}^2 &= \frac{1}{n}\sum (B_{ij} - < B_{ij} >)^2 \\ \sigma_{Iij}^2 &= \frac{1}{n}\sum ((I_{ij} - 1)^2 + \sigma_{Bij}^2) \end{aligned}$$

For both UV and optical data the STIS pipeline then used an arc exposure to calculate the wavelength dispersion solution with the necessary heliocentric velocity correction. Using a spectro-photometric star the spectra were also flux calibrated to give counts in ergs s^{-1} cm^{-2} $Å^{-1}$. The STIS pipeline produced two data products, both rectified 2D images and extracted 1D long-slit and echelle spectra, though the separate echelle orders had not yet been combined.

UVES Calibration

The pipeline steps for the UVES raw data followed the same procedure as the STIS optical data save for two main differences. UVES is a ground based instrument hence the sky lines needed to be subtracted when extracting the 1D spectra. Sampling the sky when extracting spectra is described in Section 3.5.3.

When we analysed the absorption lines of the UVES spectra we used vacuum wavelengths so that they would correspond to the absorption lines of the STIS data. Therefore the dispersion solution needed to undergo an air-to-vacuum wavelength correction first before transforming to the heliocentric reference frame.

1D extraction

According to Kim Quijano et al. (2007) the pipeline STIS spectra were not extracted using variance weighting, so this algorithm described below applies to the UVES spectra only. The STIS spectra were extracted by a regular summation across the spatial axis.

Optimal extraction (Horne, 1986) increases the S/N by weighting the pixels according to their variance (σ_{ij}^2) and distance along the spatial axis. For each column j along the dispersion axis a Gaussian function is fitted to N pixels that all contribute to the illumination profile. The flux from each pixel (I_{ij}) is scaled by the parameter P_{ij} according to its position along this curve, where over the N pixels along the spatial axis, $P_j = \sum_{i=1}^{i=N} P_{ij} = 1$. The pixels are weighted by the formula, $\mathrm{weight_{ij}} = \mathrm{P_{ij}^2/\sigma_{ij}^2}$. Where σ_{ij}^2 is the variance in both the science data counts plus the variance in the background. The 1D spectrum is then extracted by binning the pixels that contain the exposed spectrum along the spatial axis, thus the flux per pixel along the dispersive axis is now,

$$I_j = \frac{\sum_{i=1}^{i=N} \frac{P_{ij}}{\sigma_{ij}^2} \left(I_{ij} - \mathrm{Background_{ij}}\right)}{\sum_{i=1}^{i=N} \frac{P_{ij}^2}{\sigma_{ij}^2}} \tag{4.6}$$

and has an optimally extracted variance per pixel of,

$$\sigma_j^2 = \left(\sum_{i=1}^{i=N} \frac{P_{ij}^2}{\sigma_{ij}^2}\right)^{-1} \tag{4.7}$$

4.3.2 Post Calibration FOS and STIS Long-slit Spectroscopy Reduction

Combining multiple exposures

The final steps in reducing the spectra were completed locally. The procedures that remained were to combine each exposure using variance weighting and fitting a QSO continuum. A different method was used depending on whether the spectrum was long-slit data or from echelle spectroscopy.

The nature of digicon, MAMA and CCD detectors means that, prior to any combining of exposures, each of the spectra have a wavelength relation that need not be the same or constant as $\Delta\lambda \propto \lambda$ (see equation (4.3)). So the first step before combining multiple exposures was to inspect the dispersion and then re-sample each of the spectra to a common solution.

Consider the i^{th} pixel in spectrum j of N exposures each containing P_j pixels. This pixel has a width of $\Delta\lambda_{ji}$ and central wavelength λ_{ji} (Churchill, 2007). The pixel bin straddles the wavelengths:

$$\begin{aligned}\lambda_{ji}^{-} &= \lambda_{ji} - \frac{1}{2}\Delta\lambda_{ji} \\ \lambda_{ji}^{+} &= \lambda_{ji} + \frac{1}{2}\Delta\lambda_{ji}\end{aligned} \tag{4.8}$$

Assume that the new re-sampled constant pixel size is to be $\Delta\lambda_k$ and there are to be M pixels for each spectrum. The pixels, now with an index k, of the re-sampled spectra have a central wavelength, minimum wavelength and maximum wavelength of,

$$\begin{aligned}\lambda_k = \lambda^{\text{min}} &+ (k-1)\,\Delta\lambda_k \quad ; k = 1, M \\ \lambda_k^{-} &= \lambda_k - \frac{1}{2}\Delta\lambda_k \\ \lambda_k^{+} &= \lambda_k + \frac{1}{2}\Delta\lambda_k\end{aligned} \tag{4.9}$$

So that flux is conserved when re-sampling the following co-efficient η_{ji} is calculated depending to what extent λ_{ji} and λ_k overlap.

$$\eta_{ji} = \begin{cases} \frac{\lambda_{ji}^{+}-\lambda_k^{-}}{\Delta\lambda_{ji}} & \text{for } \lambda_k^{-} \leq \lambda_{ji}^{+} \leq \lambda_k^{+} \\ \frac{\lambda_k^{+}-\lambda_{ji}^{-}}{\Delta\lambda_{ji}} & \text{for } \lambda_k^{-} \leq \lambda_{ji}^{-} \leq \lambda_k^{+} \\ 1 & \text{for } \lambda_k = \lambda_{ji} \\ 0 & \text{otherwise, where there is no overlap} \end{cases} \tag{4.10}$$

For each spectrum (j), the new flux and variance resampled to pixel basis k are calculated to be:

$$I_{jk} = \frac{\sum_{i=1}^{P_j} \eta_{ji} I_{ji}}{\sum_{i=1}^{P_j} \eta_{ji}} \tag{4.11a}$$

$$\sigma_{jk}^2 = \frac{\sum_{i=1}^{P_j} \eta_{ji} \sigma_{ji}^2}{\sum_{i=1}^{P_j} \eta_{ji}} \tag{4.11b}$$

Now that all the spectra have the same dispersion solution and number of pixels they can be combined using variance weighting (Bevington and Robinson, 2002).

$$I_k = \frac{\sum_{j=1}^{N} \frac{I_{jk}}{\sigma_{jk}^2}}{\sum_{j=1}^{N} \frac{1}{\sigma_{jk}^2}} \tag{4.12a}$$

$$\sigma_k^2 = \left(\sum_{j=1}^{N} \frac{1}{\sigma_{jk}^2} \right)^{-1} \tag{4.12b}$$

The problem with equation (4.12a) is that if the variance spectra fluctuate rapidly then they will produce a rapidly fluctuating spectrum. Therefore the variance spectra were box-car smoothed over W pixels using equation (4.13) before combining.

$$\sigma_i^2 = \frac{1}{W} \sum_{l=0}^{W-1} \sigma_{i+l-W/2}^2 \tag{4.13}$$

Inspection of the G190H data for PKS 1127-145 showed that a minimum smoothing box size of W = 5 pixels was necessary in order to prevent the variance spectra from causing a rapidly fluctuating combined spectrum. This was only necessary at wavelengths $\leq$ 1700 Å where the S/N per pixel was $\leq$ 2. All the other long-slit spectra were combined using an un-smoothed variance.

The spectra taken with the FOS G190H grating along both lines-of-sight already had a common dispersion solution so these spectra did not need re-sampling. The flux and variance spectra were combined using equations (4.12a) and (4.12b) with an IDL program.

The two data sets that were taken using the STIS G430L grating towards HE 1122-1648 did not have the same dispersion solution. Inspection of this solution showed that the central wavelength of each equivalent pixel bin was different by $\sim$ 2%. The IRAF program DISPCOR was used to set a common dispersion relation that conserved the flux using equations (4.8) to (4.11b) and the IDL program was used to combine them.

Fitting the QSO continuum

When analysing absorption lines the measured observable is the equivalent width (EW), and to obtain this the flux needed to be divided by the QSO continuum. As both the b-parameter and column density were derived from the minimum and FWHM of the absorption profile an accurate placement of the continuum was essential because any fault in the level would be carried forward as a systematic error.

The continua for the long-slit spectra were fit using the tools in SPLOT. The first step when estimating a continuum was to crop the spectrum by ~ 100 Å at both the red and blue ends. This was to remove bad data with a S/N per pixel ≤ 2. Any emission features or large absorption profiles that deviate from the smooth continuum line were also clipped. In particular this included the Lyα and Lyβ emission lines if they were present. The continuum of a QSO spectrum was fitted by sampling the spectrum at points devoid of absorption, usually in regions redwards of the Lyα emission, this was then extrapolated to lower wavelengths. The continuum is expected to have a smooth variation in the level of flux with wavelength, and only break when it reaches the Lyman-limit (Kim et al., 2007).

To ensure a smooth continuum line across the spectrum that would not deviate because of a dense Lyα forest, or abruptly change at an emission feature, a chebyshev function of order 2 was fit to the data using SPLOT.

The final spectra were then saved in a FITS format for use in RDGEN, this is a program that would normalise the spectra by dividing through by the continuum and detect absorption features. RDGEN is part of the VPFIT[2] package that was used when modelling Voigt profiles, this technique is described in Carswell et al. (1991).

---

[2] ©2007 R.F.Carswell, J.K. Webb, M.J. Irwin, A.J. Cooke, Institute of Astronomy, Madingley Rd, Cambridge, CB3 0HA,Uk.

4.3.3 Post Calibration STIS E230M and UVES Echelle Spectroscopy Reduction

Unlike the long-slit spectra the echelle data was composed of many exposures. These not only needed to be re-sampled and combined, but also the separate overlapping orders needed to be joined end-to-end while simultaneously avoiding aliasing or losing flux when re-sampling. For this reason a C++ program called UVES Popler (version 0.17) (Murphy et al., 2007) was used. This was devised and written by M. T. Murphy [3] with the intention of reducing post-pipeline UVES echelle data. The input for UVES Popler was all the optimally extracted echelle orders from the STScI CALSTIS and UVES pipelines and their associated variance spectra. Combining these orders, resampling the data and setting a common dispersion solution was then an automated process, where UVES Popler employed the algorithms in equations (4.8) to (4.12b).

UVES Popler

Values for parameters with the keywords listed below were submitted to UVES Popler when combining echelle data. These defined the limits when σ clipping and what constraints to use when fitting the continuum. A QSO continuum across the full wavelength region of interest was expected to be a smoothly varying function of low order. Therefore based on this description what values to use for the continuum parameters were found on a trial basis.

- -ftyplya/-ftypred: This enabled the user to set the desired function for the continuum both below (ftplya) and above (frtpred) the Lyα emission line. A chebyshev was selected for both regions.
- -cordlya/-cordred: This set the order for the continuum, again separately above and below the Lyα line. As the QSO continuum was expected to have little variation a quadratic chebyshev was chosen for both regions. If present

---

[3] M. Murphy, Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Mail 39, PO Box 218, Hawthorn, Victoria 3122, Australia

the Lyα emission line had been clipped to ensure easy interpolation between the functions.

- -pctllya/-pctlred: It was the high column density absorption lines that prevented an initial accurate estimate for the level of the continuum. These lines were ignored by UVES Popler by rejecting any data that was in the lower 50% of the flux values below the Lyα emission line. Flux values that made up the lower 30% of the data above Lyα were also ignored. A higher tolerance was permitted as there were only a few metal lines.

- -rsiglyal/-rsiglyau/-rsigredl/-rsigredu: The level to set the continuum was then iteratively improved by neglecting data that remained 1.2σ below a provisional continuum line and any emission/noise features that existed 3σ above. The value of 3σ was the default setting and proved adequate as no emission lines were present. A low value of 1.2σ was chosen because it was our wish to find absorbers with a minimum detectable equivalent width, so ignoring flux levels that are fractionally below a possible continuum level ensured that these low column-density absorbers were not excluded. It was not possible to go lower still because of the large margin of error that was needed when estimating the position of the line. If this level were set too low then many more noise lines could be mistaken for absorption profiles.

- -vclya/-vcred/-vlya: The continuum level was determined by sampling the data at regular intervals in velocity space and then interpolating between these points. Care had to be taken when doing this because of the large fluctuations in the level of a quasar spectrum in the Lyα forest. In order that the continuum line would not deviate from what is expected the flux was sampled at 10000 kms^{-1} (vclya) intervals below the Lyα line. This was decreased to 2500 kms^{-1} (vcred) at wavelengths above.

The higher UV ionising background that surrounds the QSO (the line of sight proximity effect (Bajtlik et al., 1988)) lowers the neutral fraction, thereby decreasing the density of Lyα absorbers. The continuum had been fit with different parameter values redwards of Lyα emission because fewer lines are expected. Therefore the beginning of the red section of the continuum line was fit 5000 kms^{-1} (vlya) to the left of the Lyα emission wavelength.

Unfortunately no procedure other than inspection presented itself to ensure that we had the level of the continuum correct. However one can regard this systematic error in the column densities and b-parameter as negligible when compared to the margin for error that must be considered when modelling Voigt profiles on data with a $S/N \leq 2$.

Both HE 1122-1648 and PKS 1127-145 were fit with the same parameters for the STIS and UVES spectra. UVES Popler allowed the user to edit each echelle order separately and only re-sampled, combined the data and set the continuum when the file was saved. Therefore it was possible to use the software to remove individual bad pixels. Between each echelle order there was approximately 200 Å overlap, and because of edge effects the last ~ 10 pixels of each order were very noisy. These pixels were clipped before the spectra were combined. UVES Popler then normalised the spectrum by dividing through by the continuum.

The output from UVES Popler was a multi-dimensional fits file that included the wavelength dispersion solution, the spectrum flux, the normalised flux, the continuum (set to 1.0) and the variance spectrum. The normalised spectra were then plotted in RDGEN to detect absorption features.

The figures of the normalised QSO spectra for the STIS E230M data that contained absorption lines used in the correlation function are in Appendix B.

4.4 Producing a Quasar Absorption Line-list

Locating the Lyα absorbers that would be used in the cross-correlation function at a redshift $z \sim 1$ required the STIS E230M data. The other spectra were used for completeness, the identification of metals and to corroborate the identity of STIS echelle lines.

The Lyα and Lyβ emission lines of quasar PKS 1127-145 are located at 2658.67 Å and 2243.25 Å, therefore only the Lyα series lines were within the wavelength range of the E230M data. This meant PKS 1127-145 was far easier to investigate than quasar HE 1122-1648 which is at a redshift z = 2.404. Lyα, Lyβ and Lyγ emission lines for this quasar were at 4133.28, 3487.46 and 3229.13 Å respectively. Hence many Lyman series absorptions were located in the UVES spectrum. Locating these lines was essential so that higher order Lyman lines could be identified all the way down to a wavelength of 2050 Å. Any remaining lines could then only be Lyα absorptions (or metals) in the convoluted E230M STIS spectrum.

As they are of lower resolution, the FOS and STIS long-slit spectra were only used as an aid when identifying lines. These spectra were not used to fit Voigt profiles that would be included in the line-list.

4.4.1 Locating and Measuring the Equivalent Width of the Absorption Lines

An automatic process for locating absorption lines whatever their species was completed using RDGEN. The output from this was a table that included the wavelength, an EW measurement, the error associated with these parameters and the significance of the line. The significance level (SL) of each line was defined as, for error in the EW σ_{EW} (Janknecht et al., 2006).

$$\mathrm{SL} = \frac{\mathrm{EW}}{\sigma_{\mathrm{EW}}} \tag{4.14}$$

RDGEN used the variance spectrum when calculating the error in the EW, no term introduced by errors in the level of the continuum were included.

At this stage all the lines within the STIS spectra where RDGEN had estimated an EW significance $\sigma \geq 1$ were passed forward to VPFIT.

4.4.2 Fitting Voigt Profiles using VPFIT

Absorption lines were fitted with Voigt profiles using VPFIT (version 9.3). On the first line of input to the program was the FWHM velocity of the instrument. These velocities were calculated using the resolving power listed in Table 4.1, and they were; 10.0 km s^{-1}, 6.7 km s^{-1} and 231 km s^{-1} for the STIS E230M, UVES and FOS G190H gratings respectively (Instrumental profile = $\frac{c}{R}$.)

Using this instrumental velocity VPFIT would attempt to model the absorption line. This was achieved by a χ^2 minimisation when comparing the real and theoretical Voigt profiles. The parameters of the lines were then returned together with 1σ errors. The input to the program was a table of the absorbers detected using RDGEN, that included the ion species and an approximate initial guess for the redshift, column density and b-parameter. The atomic data required was taken from Morton (2003).

We analysed the UVES and STIS spectra in chunks of $\Delta\lambda$ = 200 Å at a time, unless there was a known higher order Lyman series line that was part of the same absorber. In this case both positions, no matter what their separation in wavelength would be submitted, and as the column density and Doppler parameter for both lines are dependent a tighter constraint could be made when estimating the 1σ errors. As our E230M spectra had a low S/N, and we already knew where the potential absorbers lay, we did not use the automatic line detection and fitting program available as this proved to be less reliable than manual input.

The following steps were made when identifying and then modelling the lines. First the Lyman series emission lines were found using the known redshift of the quasar.

Next the galactic lines were identified and flagged as any redshifted line that was blended with these had to be ignored. These galactic lines were included in the line-list for completeness and are listed together with their rest wavelengths in Table 4.2. No significant galactic lines ($\sigma_{\mathrm{EW}} \geq 3$) from species other than those listed in Table 4.2 could be detected.

The third step was to identify redshifted metal lines above Lyα emission. The metals within the IGM are more likely to be found in the vicinity of DLAs, and it was the DLA system at z = 0.313 along the line-of-sight to PKS 1127-145, (where the Lyα and C IV lines are visible in the FOS spectra (Rao et al., 2003b)) which are associated with the Mg II and Fe II systems in the UVES spectrum. The HE 1122-1648 FOS G190H spectrum exhibited a damped system at z = 0.681 ($\log_{10}$ (N_{HI} (cm^{-2}))= 20.53$\pm$0.03), Mg II and Fe II lines in the UVES spectrum were associated with this DLA. No heavy metal lines were detected at any wavelength

Table 4.2. The redshift zero galactic lines identified in STIS E230M and UVES data

Species	wavelength (Å)[a]
Fe II	2600.17, 2586.65, 2382.77, 2374.46, 2344.21
Mg II	2803.53, 2796.35
Mg I	2852.96
Ca II	3969.59, 3934.78
Na I	5897.56, 5891.58

[a] Rest wavelengths and oscillator strengths are from Morton (2003).

that were not also associated with a damped ($\log_{10}(N_{HI}\,(cm^{-2}))\geq 20$) system.

The other method employed was to identify metal lines such as C IV and Mg II by their doublet wavelength separation.

Now metal lines were found within the Lyman-series forest, in particular the doublets C IV ($\lambda\lambda 1548.20, 1550.77$) and O VI ($\lambda\lambda 1031.93, 1037.62$). As this was done by eye only two ways were known to discriminate between metals and possible Lyα lines. First, if both lines could be spotted there was the doublet spacing. Plus for metal lines the b-parameter is almost always $\leq$ 10km s^{-1} while for Lyα b is seldom $\leq$ 15 km s^{-1} and cannot be < 10 km s^{-1} (Janknecht et al., 2006) . So an inspection of the FWHM gave a good indication as to the nature of the species.

The poor S/N of the STIS E230M spectra meant that only 18 O VI systems could be detected all of which were in the spectrum of HE 1122-1648.

With the metal absorption lines identified it could now be assumed that those left were Lyman lines. First of all any remaining unidentified lines within the Lyα forest were Lyα lines. RDGEN was then used to plot the output of VPFIT and the location of the Lyβ lines was highlighted. VPFIT was then rerun if there was a positive match so that both lines could be used to constrain the errors. Any line remaining within the Lyβ forest could then only be Lyα. This process was repeated towards shorter wavelengths until the blue end of the STIS echelle spectrum was reached at 2270 and 2043 Å for the STIS spectra of quasars PKS 1127-145 and HE 1122-1648 respectively. The damped system at z = 0.681 and very poor S/N prevented further

Lyα absorbers being detected at shorter wavelengths in the sightline towards HE 1122-1648.

The minimum detectable column density

A minimum cut-off was imposed, below which the EW of a line was too low and the output from VPFIT too unreliable. An effective way to decide the minimum detectable column density for the Lyα lines was to invert the spectrum using RDGEN. Using the column density and b value of the smallest known Lyα absorber, we attempted to fit an 'absorption line' to the largest noise profiles. When the column density and 1σ error to the largest noise profile became indistinguishable from that of the smallest known real absorber this gave the minimum reliable value. For the PKS1127-145 and HE 1122-1648 STIS spectra the minimum column densities were $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}}))= 13.02 \pm 0.28$ and $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}}))= 13.15 \pm 0.14$ respectively. For the UVES HE 1122-1648 spectrum, where the S/N was ~ 120 pixel^{-1} at the central wavelength (Dall'Aglio et al., 2008) this minimum decreased to $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}}))= 11.70 \pm 0.50$.

VPFIT would sometimes attempt to fit these weak lines with unphysical b-parameters. For thermal broadening the Doppler parameter is proportional to $\sqrt{\frac{2kT}{m}}$. Thus any Lyα line that was fitted with $b \leq 10$ km s^{-1} was dropped from the list, where we have adopted the same cut that was used in Janknecht et al. (2006).

The UVES spectra of the two quasars had already been reduced and line-lists made before and these were used as an aid in identifying lines and when estimating the initial values for VPFIT. These lists included lines that had been ignored when using RDGEN as they had an EW significance of $\sigma \leq 1$. The details for HE 1122-1648 came from Carswell et al. (2002b). The recommendation from R.T. Carswell (private communication) to re-fit the lines was because of an updated version of VPFIT. The UVES line-list for PKS 1127-145 was adapted from Ryabinkov et al. (2003).

4.4.3 Reduction of Data With very poor Signal-to-noise

Fixing the variables

Within regions of the STIS E230M data where the S/N per pixel was ~ 1 VPFIT would often halt as no adequate solution could be found. To achieve an acceptable fit the initial estimates used at input could be frozen, for example in the case where the exact location of a line was known. This could be because it was part of a doublet system, a higher order Lyman-series line or was a metal line that was associated with a known high column density Lyα absorber. When this occurred we could keep z constant and that would help constrain the errors in the other line parameters.

There were cases where a known absorber was swamped by noisy data either side of the line, and that prevented an accurate fit. If this was the case the continuum could be forced to interpolate over the noisy data in order that VPFIT could be used to fit a single profile.

We set VPFIT to terminate if the 1σ error in the column density was greater than $\Delta\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 0.5$. Whenever this point was reached we would fix the b-parameter to the most common b value, this was 26.5 km s^{-1}. Figure 4.3 shows a histogram of the b-parameter for the red spectrum STIS data. When binned by 2 km s^{-1} this results in a modal value of 27.0 km s^{-1}. The finer binning of 1 km s^{-1} gave the b=26.5 km s^{-1} that was used when fitting absorbers in the blue spectrum.

The blue HE 1122-1648 STIS E230M spectrum

The S/N of this spectrum was very poor and below 1 for most wavelengths. The wavelength range was meant to cover 1574 → 2382 Å but the data became unusable below the DLA at 2043 Å. An original plan had been to ignore this spectrum, and concentrate on the Lyα absorber-galaxy pairs at redshifts $z \geq 0.87$. However inspection of the galaxy redshift distribution showed that a lot of galaxies were found at $z \sim 0.8$. This redshift regime was covered by this spectrum and it was our wish above all else to maximise the total number of pairs available for the cross-correlation function.

The solution was to sum the absorbed flux below the continuum line over a

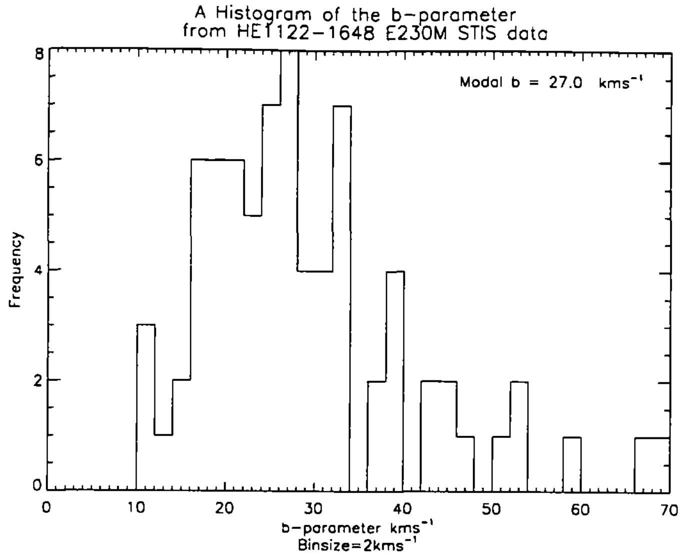


Figure 4.3 This histogram is of the b-parameter measured using VPFIT on the red STIS E230M spectrum. Only profiles that had a significance $\geq 3\sigma$ and did not need a parameter fixing (b value or the redshift) were included in this figure. In this example the binsize is 2 km s^{-1} which produces a modal b-parameter of 27.0 km s^{-1}. In order to constrain the error in the column density for profiles of very low S/N the b-parameter was fixed using a modal value of 26.5 km s^{-1}. This value was found using a binning of 1 km s^{-1}.

narrow bandpass and then plot this as a function of wavelength. The position of the absorption lines that were lost in the noise could then be determined. Knowing where these lines were meant that Voigt profiles of a fixed b-parameter could then be fit.

Firstly the spectrum was cropped to remove the damped system. A flux deficit (d_i) per pixel i over a window of size M+1 was then summed using the following algorithm that was computed using an IDL program. C_i is the continuum level normalised to 1 for all i and F_i is the flux.

$$d_i = \sum_{j=-\frac{M}{2}}^{\frac{M}{2}} (C_{i+j} - F_{i+j}) \times (\lambda_{i+j+1} - \lambda_{i+j}) \tag{4.15}$$

The variance of this result was also computed, where the values were summed in quadrature.

$$\sigma_{\rm i} = \sqrt{\sum_{\rm j=-\frac{M}{2}}^{\frac{M}{2}} \left[\sigma_{i+j}\left(\lambda_{\rm i+j+1} - \lambda_{\rm i+j}\right)\right]^2} \quad (4.16)$$

The summed flux deficit as a function of wavelength is plotted in Figure 4.4a. The first difficulty was in finding a value for M. If M were too small then absorptions would remain hidden, too large and we would lose resolution. A minimum value for M was the expected FWHM of the absorption lines assuming a particular b-parameter. Equation (1.29) related the b-parameter to the FWHM of an absorption profile b = FWHM/$2\sqrt{\ln 2}$. Using the modal b value of 26.5 km s^{-1} from Figure 4.3 gave a FWHM of 44.1 km s^{-1}. The central wavelength (2175 Å) had a dispersion of 0.03 Å pixel^{-1}, which implies FWHM = 11 pixels. However a window of 11 pixels was too low, and the flux deficit spectrum was still too noisy to identify absorption profiles from the spectrum.

Two identifications had been made in this spectrum. These were two Lyβ lines associated with the systems at z = 1.229 and z = 1.234, these can be seen in Figure 4.4a at ~ 2290 Å. The smallest profile had a width of 24 pixels, so assuming this profile, which had a significance σ = 8.7, was typical of the minimum width we could hope to identify this was chosen to be the window size.

The next parameter to find was a flux deficit cut for a real absorber. For this we adapted the earlier method of inverting the spectrum to sample the noise. Equation (4.15) was applied to the inverted spectrum using the same window size, this is plotted in Figure 4.4b. The peak inverted summed flux had a value of 0.54. Using this cut on the real summed spectrum meant only 28 absorbers would be included. So by using the peak noise value there was a danger that many real Lyα absorbers would be lost.

The red HE 1122-1648 STIS E230M spectrum had a Lyα line density of $\frac{dN}{dz} = 195$, and all results indicate that the line density distribution is almost flat at these redshifts (Kirkman et al., 2007) so a similar value is expected. The conservative cut that left 28 absorbers implied a line density of only $\frac{dN}{dz} = 165$. This cut-off needed to be reduced by 20% to a minimum flux-deficit of 0.43 before a similar line density with 33 Lyα absorbers could be observed. As Figure 4.4b shows, such a minimum

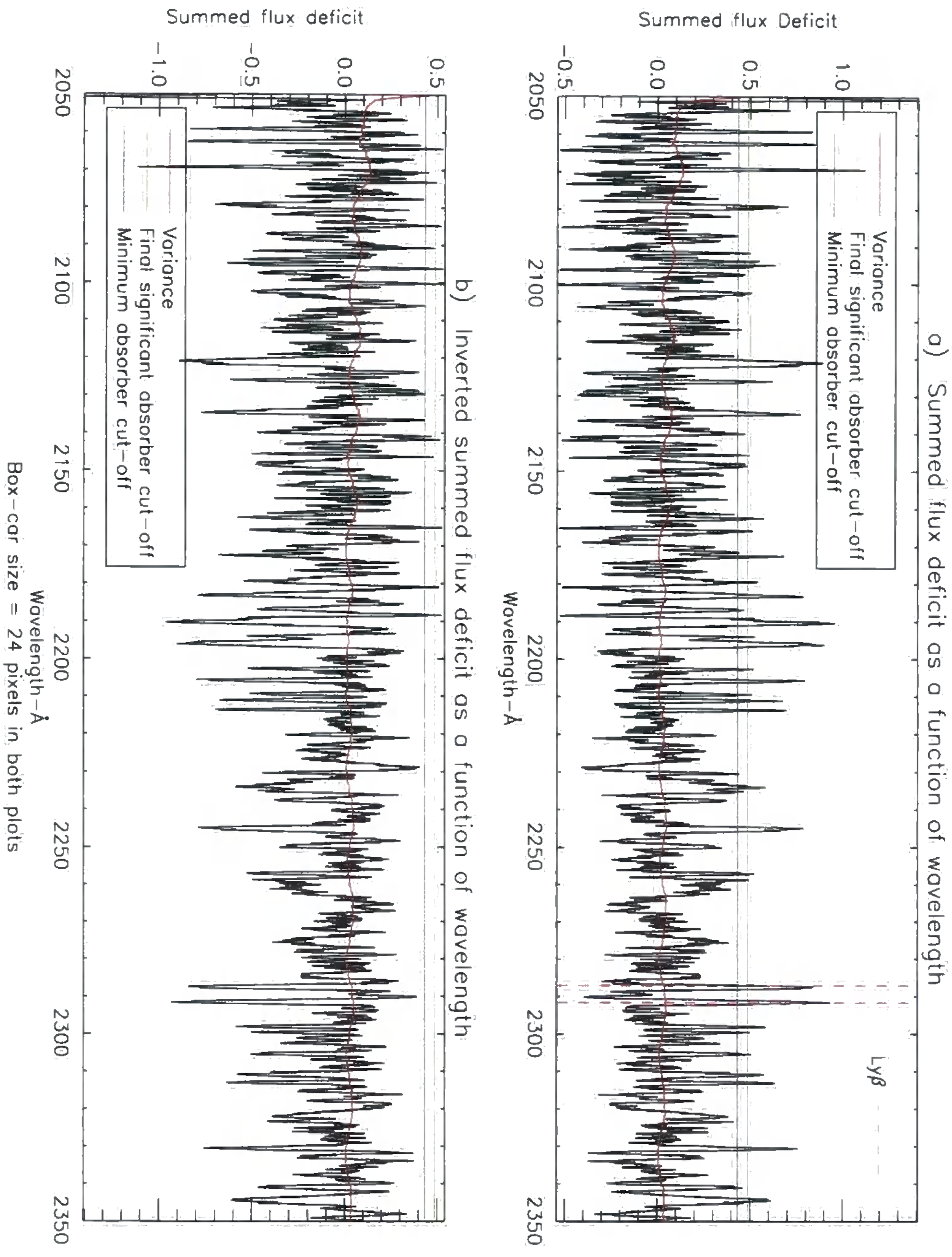


Figure 4.4 (a) In order to locate Lyα absorbers in the blue STIS E230M spectrum of HE 1122-1648, the absorbed flux was summed over a window of 24 pixels using equation (4.15). This way any absorbers were enhanced and could be modelled using VPFIT. (b) The 'significant absorber' cut-off line was set by inverting the spectrum and sampling the noise. To mimic the line-density at longer wavelengths would require a cut in the flux deficit of 0.43, however many noise features would then be included (blue line). A more conservative cut of 0.49 was chosen (green line) which meant 38 additional Lyα absorbers were identified.

cut-off would allow a lot of false identifications from smoothed noise.

Therefore the cut was varied between 0.43 and 0.54 in order to see what proportion of lines would be excluded when compared to the 33 Lyα lines that are expected. Eventually the cut-off in Figure 4.4b was set to exclude all but the highest noise peaks at 0.49. In doing this 46 absorbers were identified. An IDL program was then written that measured the wavelength of these peaks. Those that were higher order H I lines, or had an EW significance less than 3 were removed. This left 37 Lyα absorbers that could be measured using VPFIT.

The original spectrum and variance were then smoothed using a box-car window of 11 pixels, (the FWHM of a Gaussian profile assuming a fixed b-parameter of 26.5 km s^{-1}), and this was used in VPFIT. The initial estimates in the input file now contained the redshift of the absorbers with a b-parameter fixed at 26.5 km s^{-1}. VPFIT was now successful in modelling the Voigt profiles with a 1σ error in the column density of $\Delta\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) < 0.5$, and these absorbers were added to the final line-list.

4.4.4 Variance Weighted Equivalent Width and the Continuum Error

All Lyα lines with an equivalent width significance $\sigma \geq 3$ were included in the cross-correlation function, this is the level that was adopted in Schneider et al. (1993). However allowances had to be made, this included fixing the b-parameter for 57 out of 217 Lyα lines in the STIS spectra and assuming the continuum was correct. Therefore the EW and significance of the lines were measured again. This time a Gaussian weighting was used because the line centre should contribute more to the EW and error calculation than the profile wings. Changes in the significance of a line because of errors in the continuum were also considered. The formulae used to calculate the EW and significance were adapted from Schneider et al. (1993).

Consider an absorption line that covers J pixels and is centred on pixel i with the continuum C set to 1. The EW of the profile is equal to:

$$\mathrm{EW} = \frac{\Delta\lambda_{\mathrm{i}} \sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{j}} \left(\mathrm{C}_{\mathrm{i+j-J/2}} - \mathrm{F}_{\mathrm{i+j-J/2}}\right)}{\sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{j}}^2} \tag{4.17}$$

$\Delta\lambda_{\mathrm{i}}$ was assumed to be constant across the absorption, and η_{j} is the Gaussian weighting coefficient where $\sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{j}} = 1$. The standard deviation (σ) of the Gaussian profile for each absorber was based on its b-parameter.

$$\sigma = \frac{\mathrm{FWHM}}{2\sqrt{2\ln 2}} = \frac{\mathrm{b}}{\sqrt{2}} \tag{4.18}$$

The amount that the variance spectrum contributed to the error in the EW was computed using:

$$\delta\left(\mathrm{EW}_{\mathrm{var}}\right) = \frac{\Delta\lambda_{\mathrm{i}} \sqrt{\sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{i}}^2 \sigma_{\mathrm{j}}^2}}{\sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{j}}^2} \tag{4.19}$$

The variance weighted significance of each line was then, like equation (4.14), the ratio of equations (4.17) and (4.19).

Our approach when testing the validity in the level of the continuum was to test how robust the significance levels were when an error $\delta\left(\mathrm{C_j} - \mathrm{F_j}\right)$ of 10% was considered. VPFIT was not rerun to find new values for the b-parameter or the column density.

$$\left(\delta \mathrm{EW}_{\mathrm{j\,cont}}\right)^2 = \left(\frac{\partial \mathrm{EW_j}}{\partial\left(\mathrm{C_j} - \mathrm{F_j}\right)}\right)^2 \left(\delta\left(\mathrm{C_j} - \mathrm{F_j}\right)\right)^2 \tag{4.20a}$$

$$\therefore\ \left(\delta \mathrm{EW}_{\mathrm{cont}}\right)^2 = \sum_{\mathrm{j=1}}^{\mathrm{j=J}} \left(\frac{\Delta\lambda_{\mathrm{i}}\eta_{\mathrm{j}}}{\sum_{\mathrm{j=1}}^{\mathrm{j=J}} \eta_{\mathrm{j}}^2}\right)^2 \times 0.01 \tag{4.20b}$$

New errors in the EW, that considered both the continuum and variance, were calculated using equations (4.19) and (4.20b) summed in quadrature. It was these latest significance levels that were used in the cross-correlation cut and are listed in Appendix C.

It was found that even with a 10% error margin in the continuum the Gaussian weighted significance of all the STIS absorption lines remained the same to 4 decimal places, and no additional lines were lost because of the 3σ cut. This was because of the relatively high minimum observable Lyα column density ($\log_{10}(\mathrm{N_{HI}}(\mathrm{cm}^{-2}))=$ 13.02). All of the STIS absorptions were of a sufficient depth below the continuum, that changing the level of this line by 0.1 had little effect on their significance.

Many of the UVES lines were so shallow that a 10% change in $(\mathrm{C_j} - \mathrm{F_j})$ made a very large difference to the EW. However the UVES data was of such high S/N that all of the absorptions with $\sigma \geq 3$ before the error in the continuum was considered remained above this significance cut afterwards.

4.5 The Final Line-Lists

The final line-list for HE 1122-1648 contained 849 absorption lines that were detected in the STIS E230M and UVES echelle data. This included 655 Lyα lines each with a b-parameter ≥ 10 km s^{-1} and 194 metals. Of these Lyα lines only 160 came from the two STIS E230M spectra, and only 103 of these had a significance ≥ 3.

Fewer Lyα lines could be detected in the STIS spectrum of PKS 1127-145 due to the lower quasar redshift. 57 Lyα lines were found together with 35 metal lines in both the STIS and UVES spectra. Of these 38 Lyα lines had a significance ≥ 3.

Therefore over a total path length of $\Delta z = 1.1$ we detected 141 Lyα absorbers that could be used in the cross-correlation with the galaxies. Lists containing all of the significant absorbers from the STIS E230M, together with the fitting parameters and 1σ errors are in Appendix C.

What this exercise has shown is that, despite noisy data, it was possible to generate a linelist of potential Lyα absorbers. However sacrifices in resolution and the minimum detectable column density had to be made.

It was hoped that the minimum detectable column density of the STIS data would have been a considerable improvement upon the FOS dataset of Morris and Jannuzi (2006). A greater redshift density of Lyα lines of low column density, $\log_{10}(\mathrm{N_{HI}}(\mathrm{cm}^{-2}))= 13 \rightarrow 15$ were detected. However the minimum column density was still greater than $\log_{10}(\mathrm{N_{HI}}(\mathrm{cm}^{-2}))= 13.0$ and not significantly lower than the minimum column density from the FOS dataset. Similarly even though C IV and O VI lines were able to be identified and resolved, their number density and significance were not sufficient in order to use them in a study with the galaxies.

Despite these drawbacks, the relationship of line density with redshift given in equation (5.1) is similar to that from the literature (Janknecht et al., 2006). So even though there were significant errors in both the b-parameter and column density we remain confident in the position and number density of those Lyα lines used in the cross-correlation function.

Chapter 5

Results of the Galaxy-Lyα Absorber Correlation Function

5.1 Analysis of the Galaxy and Absorber Samples

5.1.1 The Galaxy Redshift Distribution

In order to explain any results that may stem from the 2D 2-point correlation function it was first necessary to consider the redshift distribution of those galaxies and absorbers that were paired. This is because not only was it predicted that the result be redshift dependent, but the number of Lyα absorbers and galaxies available for pairing turned out to rely to a great extent on the redshift.

Figure 5.1 shows the distribution of all the galaxy spectroscopic redshifts for all confidence levels found in the FORS2 survey. We collected Lyα absorbers from the STIS E230M grating that had redshifts between $0.68 \leq z \leq 1.5$. Our method of using the magnitude limit $21.5 \leq R_{Vega} \leq 24.5$ when selecting targets meant we were able to detect galaxies, of which the majority (139 out of 200) were within this same redshift window. The cut-off line at $z = 0.68$ shows the minimum redshift at which a galaxy may be paired with a Lyα absorber. The redshift bin at $z = 1.5$ contains the furthest galaxies from the survey. This depth in redshift was set by the 1 hour integration time allowed for the spectroscopy, and by the fact that there were no significant emission or absorption features with a shorter wavelength than [O II] that were redshifted into the FORS2 R-filter bandpass.

One problem was that we were not successful in getting a good range of galaxy types. All but 13 were emission dominated galaxies, and the modal redshift at

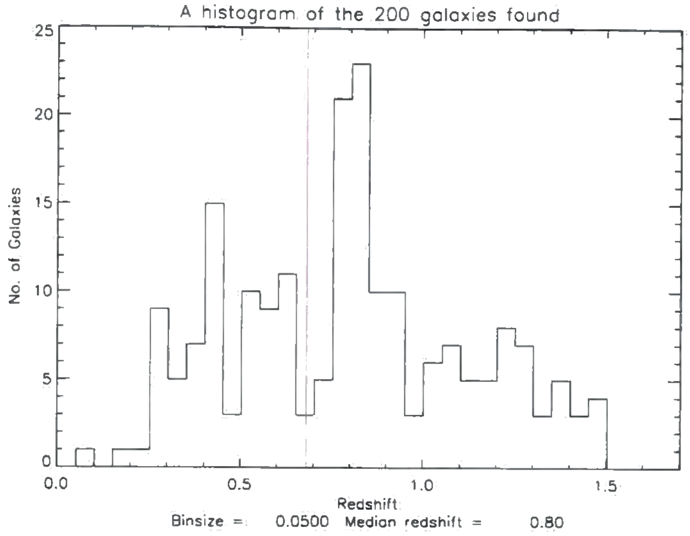


Figure 5.1 The FORS2 survey galaxy redshift distribution along the two QSO fields of view. The line at z=0.68 shows the cut-off above which galaxies may be correlated with the STIS E230M Lyα absorbers. The highest redshift was at z = 1.47. No galaxies were detected above this because of the lack of significant emission features at shorter wavelengths than the [O II] emission line. When above this redshift [O II] was also lost to atmospheric and [OH$^-$] absorption lines.

z $\sim$ 0.8 may have been caused by the high frequency with which [O II] was detected when performing a magnitude limited sample in a wavelength window of 5500–10000 Å.

The number of galaxies within the E230M redshift range could be improved by allocating slits based on a B − R, R − I colour selection that was used in Coil et al. (2004). This technique, recommended by the ESO Telescope Allocation Team, was utilised in later investigations. This is discussed in Section 8.2.

5.1.2 The Galaxy Impact Parameter Distribution

When sampling the galaxies in the QSO field of view the slits were assigned in a random fashion. Therefore the sample of values for the impact parameter between the absorber and galaxy for each pair was also random. However the maximum possible impact parameter is a function of redshift and was constrained by the $6.8' \times 5.7'$ field of view of the FORS2 detector. Figure 5.2 shows a plot of the maximum possible impact parameter. As we detect galaxies at deeper redshifts we see a larger possible projected separation, and hence distance over which we can measure the Lyα absorber-galaxy correlation. Using the stated cosmology, the size of the FORS2 detector and the maximum galaxy redshift meant that galaxy and absorbers were paired out to a projected distance of $\sigma = 5\ h_{70}^{-1}$Mpc.

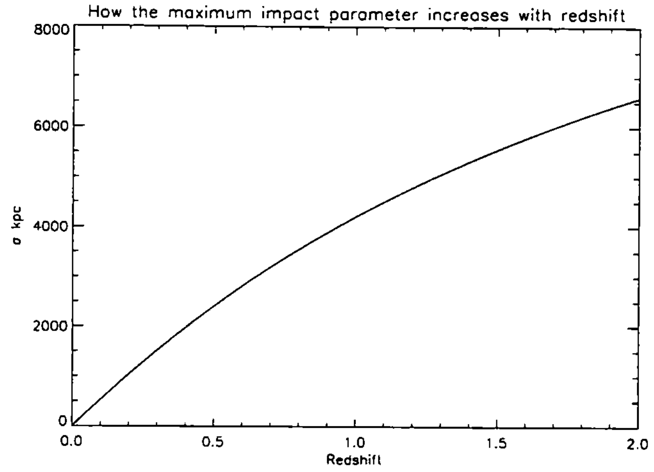


Figure 5.2 The curve in this diagram shows the increase in the maximum possible observable impact parameter when one uses the 6.8′ × 5.7′ detector on the FORS2 instrument. This plot was derived using our stated cosmology and assuming that the galaxy was located in the corner of the detector, the angular separation $\delta\theta^2 = 3.4'^2 + 2.85'^2$ from the QSO at the centre of the chip. The constraints of the size of the field and the maximum galaxy redshift meant we were able to correlate galaxies with absorbers out to a projected distance of $\sigma = 5\ h_{70}^{-1}$Mpc.

A test to investigate how our galaxy redshift distribution may effect the correlation function was to plot the impact parameter distribution. The solid line in Figure 5.3 is a histogram of the projected separation between all 200 galaxies and the central quasar sightline.

We see the expected initial shape because the number of galaxies located within each concentric cylinder centred on the quasar increased.

This distribution then peaked along a plateau at a projected separation of $\sigma = 1000-2000\ h_{70}^{-1}$kpc, before rapidly declining above $\sigma = 3000\ h_{70}^{-1}$kpc. This flattening of the distribution reflects the peaks in the histogram of Figure 5.1 and the expected range in impact parameter at these redshifts (see Figure 5.2).

According to Figure 5.1 most of the galaxies were detected at a redshift z $\sim$ 0.8. Therefore the peak in Figure 5.3 must occur below $\sigma \sim 3000\ h_{70}^{-1}$kpc. However the number of galaxies per bin had already started to decline by $\sigma = 2000\ h_{70}^{-1}$kpc. This was caused by a decrease in the radial density of slits as the edge of the rectangular

$6.8' \times 5.7'$ detector was reached.

The number of galaxies above $\sigma = 4000$ h_{70}^{-1}kpc was so low because the size of the chip meant that a minimum redshift $z \gtrsim 1$ was required to reach this impact parameter. Our galaxy sample was heavily skewed with 73.5% of the redshifts ≤ 1.

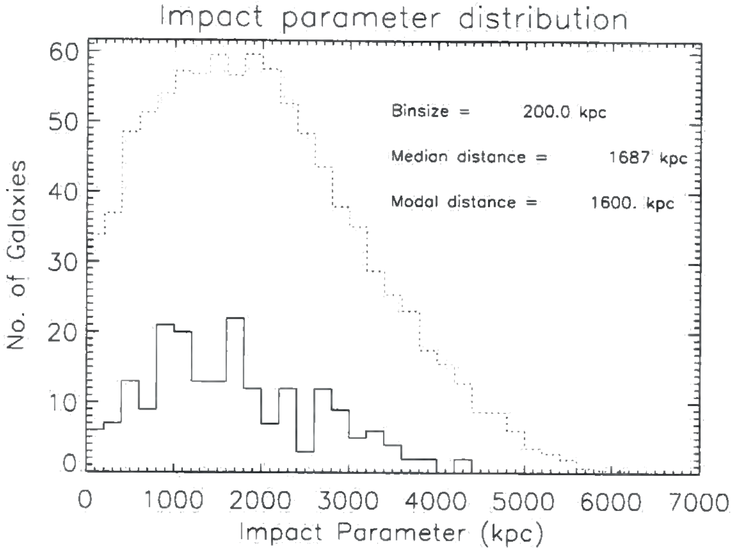


Figure 5.3 The solid line histogram shows the number of galaxies surveyed as a function of impact parameter binned by 200 h_{70}^{-1}kpc. The number of galaxies observed has an expected increase, peaking at 1000-2000 h_{70}^{-1}kpc as the survey volume increased. The edge effects of the chip and the drop off in the number of galaxies at high redshift then caused a decrease in the number of galaxies beyond 3000 h_{70}^{-1}kpc. The dotted line represents the expected number of galaxies in a complete magnitude limited survey ($21.5 \leq R \leq 24.5$) where the data is from Cohen et al. (2000), Hogg et al. (1997). This plot shows that even though we were able to sample the correct impact distribution for the galaxies the survey was far from complete.

The dashed histogram is a completeness test that shows the expected separation had every galaxy within the same magnitude range ($21.5 \leq m_R \leq 24.5$) been targeted. Mock galaxies were placed with a uniform spacing over the same $6.8' \times 5.7'$ area. The required surface density for this magnitude cut came from Figure 4 of Hogg et al. (1997). Each mock galaxy was then assigned a random redshift taken from the Hubble Deep Field North Survey redshift distribution, the details of which are from Table 4 of Cohen et al. (2000). The only unknown was the redshift cut-off for the mock galaxies, for which a fiducial value of 1.8 was chosen. We then calculated the impact parameter of each mock galaxy from an imaginary sight line located at the centre. This was repeated using 1000 mock galaxy catalogues in order to calculate the expected impact parameter distribution.

Figure 5.3 shows that our galaxy survey follows the same shape as the expected distribution. The mock catalogue also peaks in the range $\sigma = 1000 - 2000$ h_{70}^{-1}kpc before decreasing above $\sigma = 2000$ h_{70}^{-1}kpc because of edge effects. However the figure also shows that our survey was far from complete, even when considering any cosmic variance in the galaxy density between the surveys in Hogg et al. (1997) and our quasar fields. An incomplete survey will have effected the significance of the correlation function as there will be many absorber-galaxy pairs unaccounted for that would have occupied a particular separation.

5.1.3 The Lyα Absorber Distribution

The Lyα absorber redshift distribution

Regardless of how complete the galaxy survey was, in order to produce enough pairs there also had to be sufficient Lyα absorbers along each QSO sightline. Figure 5.4 shows a histogram of all the Lyα absorption lines detected with an equivalent width (EW) significance $\sigma \geq 3$ from the STIS and UVES echelle spectra across both lines of sight.

All of the Lyα absorbers below the blue cut-off at z = 1.51 were found in the STIS data and are the 141 lines that were used in the cross-correlation function. Those above are from the UVES spectrum and hence are all from the HE 1122-1648 sightline. Those below the red cut-off at z = 0.86 were found in the blue STIS data set that covered the wavelength range 2050 – 2382 Å and so are also only from HE 1122-1648. The green asterisks mark the redshifts of quasars PKS 1127-145 and HE 1122-1648 at z = 1.187 and z = 2.4 respectively.

Three factors changed the shape of this histogram from the predicted curve. The line-density distribution was expected to be flat out to redshifts z $\sim$ 1.6 and then to increase (Janknecht et al., 2006) for reasons discussed in Section 1.4.5. However this was not seen in this case. This is partly because the data was from only two sightlines, one of which terminated at z = 1.187, but also the deficit of lines at z $\sim$ 1.5 is attributed to the decreasing S/N and low sensitivity at the extreme red

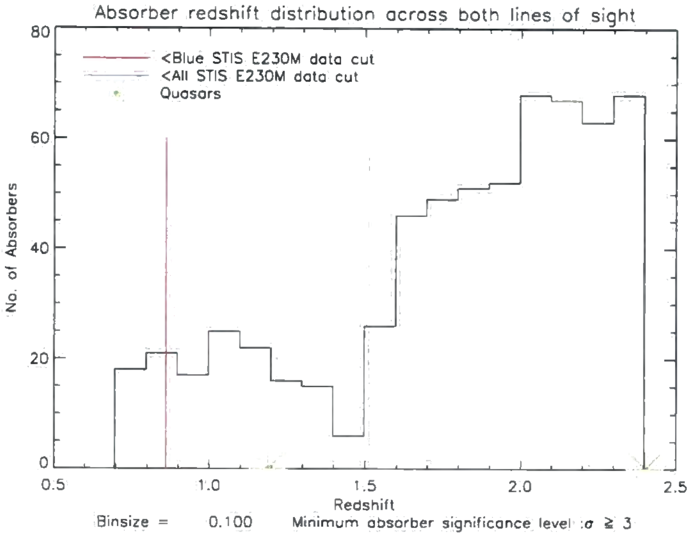


Figure 5.4 This histogram binned by 0.1 in redshift shows the distribution of all Lyα absorbers with significance $\sigma \geq 3$ detected across both lines of sight and both the STIS and UVES instruments. The blue line marks the redshift cut-off between the two instruments and Lyα absorbers above this were only found along the sightline of HE 1122-1648. Lyα absorbers below the red line were found in the low wavelength HE 1122-1648 dataset, a damped system prevented any more absorbers being detected with a redshift below z = 0.68. No echelle STIS data below a redshift of z = 0.896 existed for QSO PKS 1127-145. The green asterisks mark the redshift of quasars PKS 1127-145 at z= 1.187 and HE 1122-1648 at z= 2.4.

end of the STIS spectra of HE 1122-16478. The increase in the line density above z = 1.5 was partly due to the increased sensitivity and S/N of the UVES instrument and not only that predicted from the literature.

The Lyα absorber column densities

Figure 5.5 shows a histogram of the column densities of those Lyα absorbers that were detected in both the STIS and UVES spectra. From this we can deduce that almost all the absorbers detected had a column density $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\leq 15$, those of which that are expected to occupy the diffuse filaments of the IGM and are representative of the mean column density of a Lyα absorber from the model in Section 1.3.4. The median column densities were $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 13.9$ and $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12.9$ from the STIS and UVES spectra respectively.

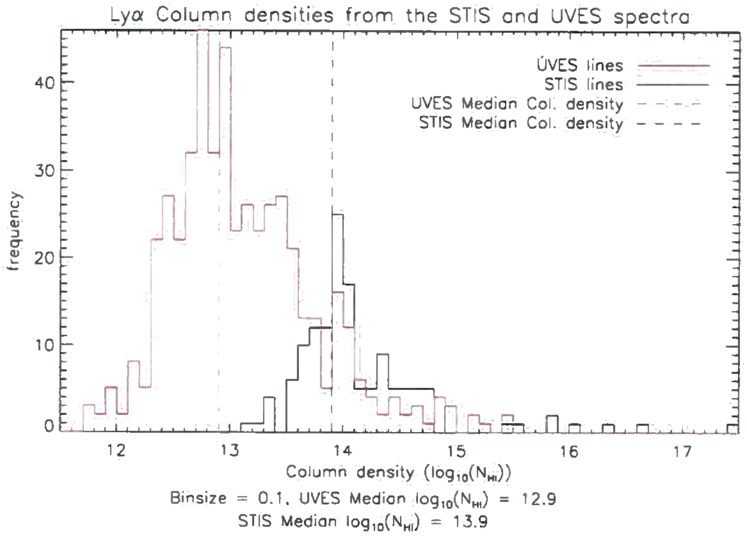


Figure 5.5 A histogram showing the Lyα absorber column densities from the UVES and STIS spectra. The median column densities were $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 13.9$ and $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 12.9$ from the STIS and UVES spectra respectively.

5.1.4 The Line-density Evolution

When investigating the line-density evolution it is useful to plot $\frac{dN}{dz}$ above a minimum equivalent width using equation (1.34). The greater sensitivity and resolution of the UVES instrument over STIS allowed for a much lower equivalent width limit.

Compare the minimum detected EW when using UVES of 0.011 Å with the minimum value of 0.13 Å when using STIS. Therefore a minimum limit of 0.13 Å was adopted at all wavelengths. Figure 5.6 shows the evolution in the number density of Lyα lines along the sightline to both quasars that uses data from both the STIS and UVES instruments.

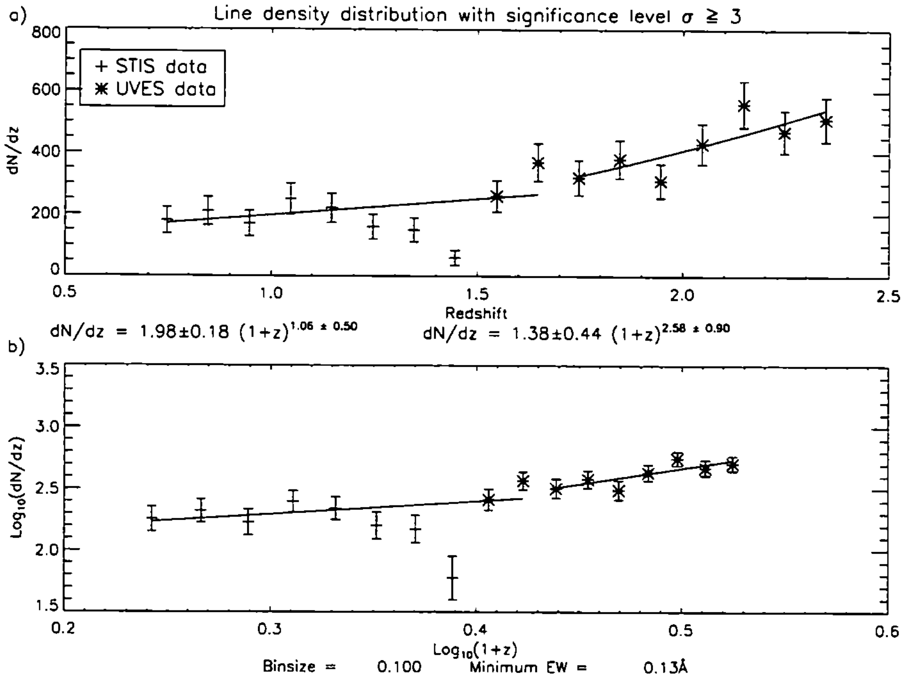


Figure 5.6 The line-density evolution of Lyα absorbers along the sightline to HE 1122-1648 and PKS 1127-145. The minimum equivalent width is the minimum observed in the STIS data of 0.13 Å. Using a least squares method two lines of best fit were plotted in (b) to fit the expected flat and increasing line densities either side of $\mathrm{Log}_{10}(1+z) = 0.43$. These lines were then used to plot the power laws in Figure (a). Up to a redshift of $z \sim 1.5$ the STIS data shows a decreasing rather than flat evolution with redshift. A flat evolution was produced if the two UVES data points at $z \sim 1.6$, that are from a spectrum with a higher S/N, were included. The decline in absorbers at $z \sim 1.5$ was not assumed to be because of a decrease in the line density. There was a low number count because of poor sensitivity at the red end of the STIS spectrum.

Although there should be one smooth relation that describes the evolution of the Lyα absorber line-density from intermediate to low redshifts ($z = 4.0 \rightarrow 0$), it is common practice (Dobrzycki et al., 2002, Janknecht et al., 2006, Scott et al., 2000, Weymann et al., 1998) to attempt to fit two separate power laws above and below a redshift of $z \sim 1.7$, because at this point there is a marked increase in line-density that was described in Section 1.4.5. Linear best fit lines were drawn in plot (b) of

Table 5.1. Literature results for the Lyα absorber line-density evolution

Dataset	Spectrograph	No. of sightlines	Redshift range	Exponent γ	Minimum EW
Weymann et al. (1998)	HST FOS	63	0-1.5	0.26 ± 0.22	0.24
Dobrzycki et al. (2002)	HST FOS	165	0-1.7	0.54 ± 0.21	0.24
Scott et al. (2000)	HST FOS	39	≥ 1.7	1.88 ± 0.22	0.36
Janknecht et al. (2006)	STIS, UVES, HIRES	9	0.5-1.9	1.74 ± 0.31	0.1
Kirkman et al. (2007)	HST FOS	74	0-1.6	1.01 ± 0.11	n/a

Figure 5.6 that produced the constant A and exponent γ for a line-density evolution power law $\frac{dN}{dz} = A(1+z)^{\gamma}$ for these two regions.

For lines with redshift $z \leq 1.7$ $$\left.\frac{dN}{dz}\right|_{EW=0.13\,\mathring{A}} = 1.98 \pm 0.18(1+z)^{1.06\pm0.50} \quad (5.1a)$$

For lines with redshift $z > 1.7$ $$\left.\frac{dN}{dz}\right|_{EW=0.13\,\mathring{A}} = 1.38 \pm 0.44(1+z)^{2.58\pm0.90} \quad (5.1b)$$

The error bars in plots (a) and (b) of Figure 5.6 were determined assuming Poisson noise in the number of lines. Therefore because of the small number of absorbers per bin there was a large 1 σ error margin in the exponents of equations (5.1a) and (5.1b).

As this data was from two lines of sight only the result can not be used to make a definitive statement about the number evolution of absorbers in the universe. However this result was similar to other studies from the literature (Dobrzycki et al., 2002, Janknecht et al., 2006, Kirkman et al., 2007, Scott et al., 2000, Weymann et al., 1998). These previous investigations can be split between those that used the HST FOS spectra collected during the Key Project investigation and the Janknecht et al. (2006) paper that only analysed high S/N echelle data from the STIS, UVES and HIRES spectrographs. The data from Kirkman et al. (2007), rather than count lines, measured the mean flux absorbed over a fixed wavelength bin using the method outlined in Section 1.4.4. Table 5.1 shows the results for these investigations.

All these previous studies show the line-density evolution up to z ~ 1.7 is flat. The result from equations (5.1) agreed with this conclusion. Had the UVES results at z ~ 1.7 been neglected then it could be argued that the line-density evolution has increased since z = 1.5. However there was the problem of poor sensitivity at these wavelengths in the STIS spectra. Most other investigations point to a flat evolution up to at least z = 1.3.

The evolution of lines in our data set above z = 1.7, measured from Figure 5.6 was $\gamma = 2.58 \pm 0.90$. This result used only UVES data, so all points are from the sightline towards HE 1122-1648. Again this trend agreed with the sudden increase in the line-density that was predicted and explained in Section 1.4.5 and illustrated in Figure 1.7.

5.1.5 Combining the Lyα Absorber and Galaxy Datasets

The redshift distribution of both galaxies and absorbers are compared in a Pie-diagram. Figures 5.7 and 5.8 show the location of both galaxies and Lyα absorbers across each sight line. Such diagrams are useful as they give an indication of the efficiency of the MOS surveys, how many potential absorber-galaxy pairs can be expected and in what redshift range they occur.

The maximum impact parameter (~ 5000 h_{70}^{-1}kpc) shows the distance out to which the correlation function was calculated, and also explains why so few pairs were found in the highest separation bin. Only those galaxies with a redshift $z \gtrsim 1$ could achieve this separation.

Pairs that were binned for the correlation function had a maximum line of sight separation of 20 h_{70}^{-1}Mpc. At a redshift of z ~ 1 this equates to $\Delta z \sim 0.05$. Therefore each absorber in the Pie-diagrams was only paired with galaxies within a very narrow redshift window, the size of which is too small to resolve in Figures 5.7 and 5.8.

Our intention behind using high resolution STIS data had been to correlate galaxies at z ~ 1 with absorbers that had separate measurements for the b-parameter and were of the minimum detectable equivalent width. However by inspection, one can tell that the majority of pairs for the correlation function came from the lower redshift (z < 0.85) galaxies. These were paired with the 37 Lyα absorbers found in the blue HE 1122-1648 STIS spectrum. It was these absorbers that were only detected after considerable smoothing and all had to have a b-parameter fixed at

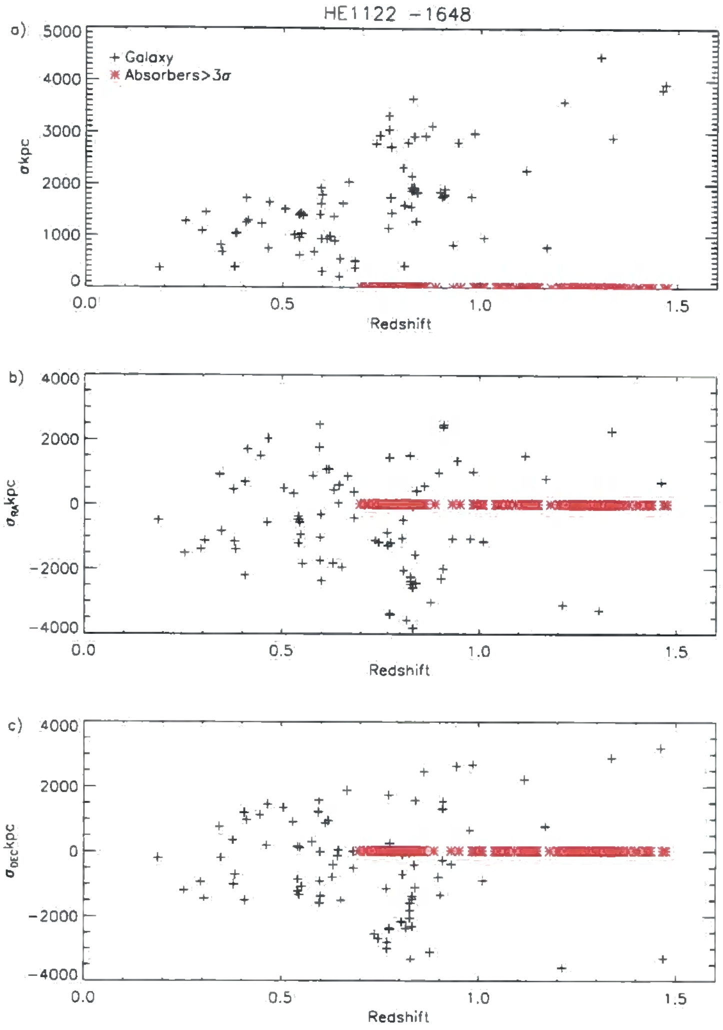


Figure 5.7 The Pie-diagram for quasar field HE 1122-1648. The top panel, Figure (a), shows the impact parameter between the galaxy and the quasar sightline plotted as a function of redshift. Panels (b) and (c) show how this separation is split between the RA and DEC respectively. The increasing maximum separation with redshift, observed in Figure (a) marks the boundary of the survey volume. The maximum redshift plotted only covers the overlap between the two populations, so only absorbers from the STIS E230M spectrum have been included. A lower than expected number of pairs arose because of the lack of galaxies above z = 1 found in the field of HE 1122-1648, even though there were many absorbers found within the STIS spectrum at these redshifts.

26.5km s^{-1} (see Section 4.4.3). Therefore even though the minimum equivalent width of the Lyα absorbers used in the cross-correlation was 0.13 Å, many of the binned pairs (90 out of 194) were from absorbers that had an equivalent width greater than 0.3 Å from this smoothed spectrum.

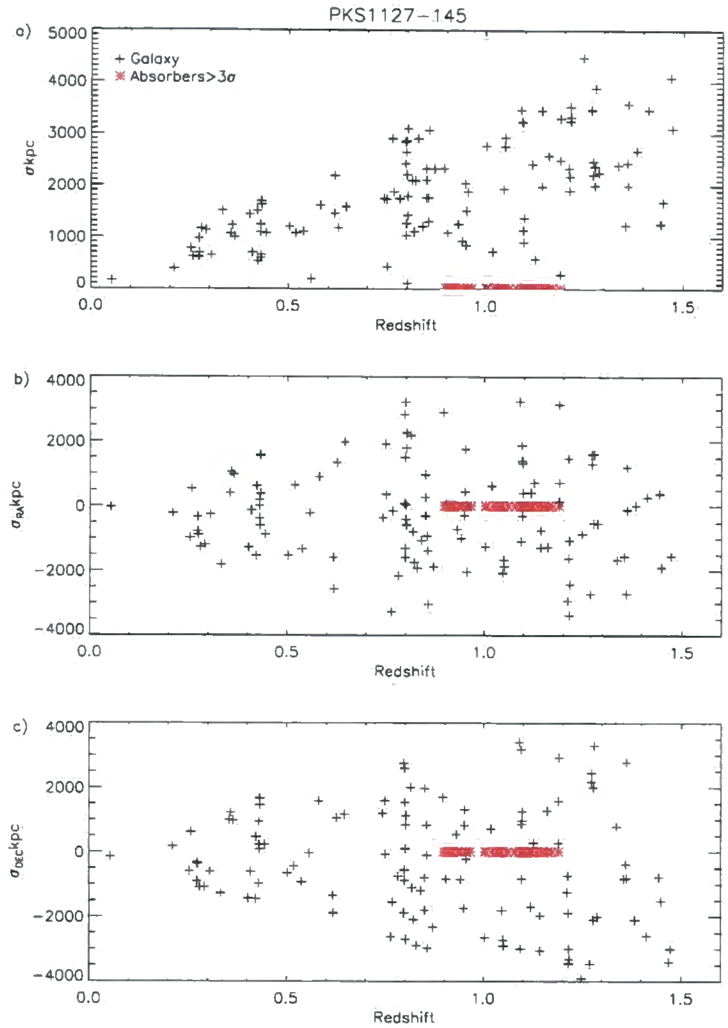


Figure 5.8 The Pie-diagram for the quasar field PKS 1127 showing the impact parameter as a function of redshift and the separation along the RA and DEC. The much narrower absorber redshift window when compared to Figure 5.7 is because of the lower redshift of quasar PKS 1127-145 z = 1.187. Therefore many successful galaxy redshift identifications were superfluous as the population of galaxies with z $\gtrsim$ 1.2 could not be used in the 2-point correlation function.

Both quasars also exhibited a fall in the number of both absorbers and galaxies at z $\sim$ 1. Rather than assume this is because the two are inextricably correlated, we suspect this is more down to the lack of galaxy redshifts determined at z $\geq$ 1 and the poor signal-to-noise of the STIS spectra.

5.2 The 2D 2-Point Correlation Function

The calculation for the 2D 2-point correlation (see Section 2.4.1) was based on the estimator devised by Davis and Peebles (1983) and was calculated using equations (2.2a) and (2.2b). To do this we paired up all the Lyα absorbers that had an equivalent width significance ≥ 3 with every galaxy from our catalogue within the same field of view. We then calculated the line of sight and projected separation between every pair.

5.2.1 Locating the Real Pairs

To collect the real pairs the final galaxy catalogue and Lyα absorber line-list were sent to an IDL program. This computed the line of sight and projected separation of every galaxy from each absorber. When calculating distances we adopted the flat ΛCDM cosmology used throughout this investigation (see Section 1.2.5). The line of sight co-moving distance ($\mathrm{D_C}$) to both galaxies and Lyα absorbers from an observer at redshift $z = 0$ was calculated using equation (5.2).

$$\mathrm{D_C} = \frac{c}{\mathrm{H_0}} \int_0^z \frac{dz'}{\sqrt{\Omega_m (1+z')^3 + \Omega_\Lambda}} \tag{5.2}$$

This is a simplification of equation (1.9) since by this epoch radiation pressure no longer has a significant contribution to Ω. The line of sight separation, π, was then just $\pi = |\mathrm{D_{C\,galaxy}} - \mathrm{D_{C\,absorber}}|$.

The angular separation between 2 points on an image that have been projected from the celestial sphere is given by equation (5.3).

$$\theta_\angle = \arccos\left(\sin\delta_{\mathrm{abs}}\sin\delta_{\mathrm{gal}} + \cos\delta_{\mathrm{abs}}\cos\delta_{\mathrm{gal}}\cos\left(\alpha_{\mathrm{gal}} - \alpha_{\mathrm{abs}}\right)\right) \tag{5.3}$$

The RA and DEC of each absorber were the J2000 coordinates of the quasar in question retrieved from the SIMBAD database. Using the mean line of sight distance towards both the absorber and galaxy, the impact parameter was then calculated to be $\sigma = \theta_\angle \overline{\mathrm{D_C}}$. As long as the π separation was $\leq 20\ h_{70}^{-1}$Mpc, each real galaxy-absorber pair was then written out to a file for subsequent binning. The output from the pairing program included as well as σ and π; the redshift of the objects, the unique galaxy and absorber ID from the catalogue, the absolute magnitude of the paired galaxy, whether it was dominated by emission or absorption, the b-parameter and column density of the paired absorber and the redshift confidence of each object. This way variations in the correlation function could be investigated by requiring objects, whether real or random to fit certain criteria, for example a particular line of sight, confidence, redshift range or column density. The real pairs are denoted DD.

5.2.2 Generating the Random Catalogues

The random galaxies

When producing the random catalogues it was important that the random galaxies had the same position on the sky as the data in order to keep the mask design information. Therefore adopting the technique used in Wilman et al. (2007), all real galaxies that lay in the redshift range of the STIS E230M spectra of that particular line of sight were replaced with 200 fake galaxies. Each of these was then set to a random redshift drawn from a uniform distribution within this range. For example in the case of HE 1122-1648 there were 53 galaxies between redshifts $z = 0.68 \rightarrow 1.51$, making $53 \times 200 = 10600$ random galaxies.

The random absorbers

The random absorbers were generated in the same fashion. Each real absorber was replaced with 200 randoms, each with the same column density, b-parameter and EW significance as the original. The redshift of each random absorber was then chosen by selecting a random position within the same redshift range as the real absorbers in that line of sight.

All real absorbers were then paired with every random galaxy, and the π and σ separations found using equations (5.2) and (5.3). Likewise the separation between all random absorbers and random galaxies were calculated. The real-random and random-random pairs are denoted DR and RR respectively.

Using 200 random targets per real absorber or galaxy was decided on a trial basis. This value was a compromise between limited computing power and ensuring that the DR and RR pair distributions were sufficiently smooth, with little variation along the line of sight at separations of equal impact parameter.

5.2.3 Binning for the 2D Plot

A separate IDL program was then written that selected the real pairs that were separated by less than or equal to 20 h_{70}^{-1}Mpc and 5 h_{70}^{-1}Mpc along the π and σ directions respectively and binned them in redshift space on a grid. Each value in this array was then divided by the associated bin in the DR and RR arrays and 1 subtracted to give $\xi_{AG}(\sigma,\pi)$. Our choice of binning was $\Delta\sigma = 1.0$ h_{70}^{-1}Mpc and $\Delta\pi = 2.0$ h_{70}^{-1}Mpc. This was coarser than the earlier studies of Ryan-Weber (2006) and Wilman et al. (2007), who binned $\Delta\pi = \Delta\sigma = 0.1$ h_{100}^{-1}Mpc and $\Delta\pi = 2.0$, $\Delta\sigma = 0.4$ h_{70}^{-1}Mpc. We required this coarser sampling so we would reduce the effects of shot noise because of the small number of real pairs.

5.2.4 The Error Calculation

In the previous studies of Ryan-Weber (2006) and Wilman et al. (2007), the uncertainty in ξ was calculated by jackknife resampling of all the sightlines. The correlation function was recalculated removing one quasar at a time and an error determined for each bin. Equation 6 from Wilman et al. (2007) gave this error for N sightlines.

$$\sigma_{\xi}^{2} = \frac{N-1}{N}\sum_{i}\left(\bar{\xi} - \xi_{i}\right)^{2} \qquad (5.4)$$

Having only two quasars in this study an alternative 'Bootstrap re-sampling' method was chosen (Pollo et al., 2005), where pseudo galaxy tables and linelists were created with the same number of absorbers and galaxies as the real data. The generation of mock redshifts was identical to that used to generate the random catalogues described above. These were then treated as real data sets and run using the exact same code. 50 of these mock linelists and galaxy tables were produced together with 50 independent estimates for the correlation function.

To avoid confusion 'mock' catalogues are ones treated as real data in order to estimate an error and all have an independent value for DD. Random catalogues, be they from the real or mock data, are used in the denominator of the estimator.

A RMS error based on the sample variance was then calculated using equation (5.5). With N = 50 mock datasets,

$$\sigma_{\xi(\sigma,\pi)} = \sqrt{\frac{\sum_i \left(\bar{\xi} - \xi_i\right)^2}{\mathrm{N} - 1}} \tag{5.5}$$

A problem with this result was that, having been randomly generated, these mock samples did not exhibit any correlation pattern. So they are not a direct measurement of the error in the correlation function with an associated S/N for a particular bin. Rather they are a gauge of the significance of any correlation that may exist and should not be assumed to be equivalent to a Poisson or standard error. Pollo et al. (2005) also claim that 'Bootstrap' errors tend to overestimate the variance. However as we only had a small data set an upper-limit to the actual error was more appropriate.

The Poisson error, found by the square-root of the number of pairs in a particular bin greatly overestimates the significance of the correlation result so was not calculated (Pollo et al., 2005, Wilman et al., 2007).

Using the standard error based on the mock catalogues was invalid because this found the deviation from the mean correlation value, calculated from datasets that were inherently random. Hence these errors, (equal to $1/\sqrt{N}$ those found above in equation (5.5)) also overestimated the significance of any correlation signal.

5.2.5 Results for the 2D 2-Point Correlation Function

Figure 5.9 shows the 194 real Lyα absorber-galaxy pairs out to 20 × 5 h_{70}^{-1}Mpc binned in redshift space with $\Delta\pi$ = 2.0, $\Delta\sigma$ = 1.0 h_{70}^{-1}Mpc. The first inspection confirmed what was already known. Above $\sigma \geq 4$ h_{70}^{-1}Mpc there were very few real pairs detected. For this reason this last column should be ignored in subsequent plots as the Poisson noise on too few pairs introduced very high anomalous results for $\xi_{\rm AG}$ in these bins. Where this occurs the highest value in the colour bar of all of the plots ignores these maxima.

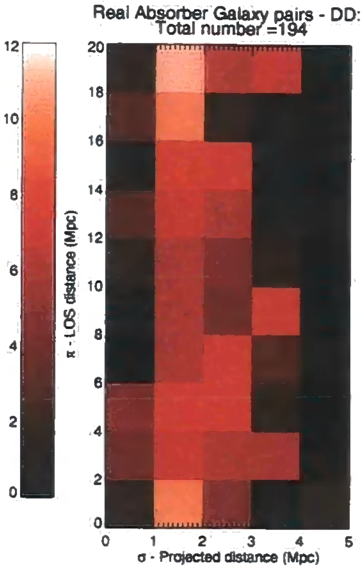


Figure 5.9 The real Lyα absorber-galaxy pairs out to 20×5 h_{70}^{-1}Mpc, binned $\Delta\pi = 2.0$, $\Delta\sigma = 1.0$ h_{70}^{-1}Mpc. All of the galaxies have been correlated in this plot no matter what their redshift confidence, with all Lyα absorbers that had an equivalent width significance $\sigma_{\rm EW} \geq 3$.

Figure 5.10 shows the plots of both the normalised DR and RR counts. The purpose of this plot is to demonstrate the minor differences between using DR or RR samples in the denominator of the correlation function, as both plots have a similar number of pairs in each corresponding bin. With 200 real or random absorbers and 200 random galaxies both plots are smooth along the line of sight. Once normalised the smaller variance in the number of pairs in the RR sample along the line of sight lowered the amount of shot noise in the correlation plots, albeit by an insignificant amount given the margin of error. For example along $\sigma = 0 - 1$ h_{70}^{-1}Mpc DR varies from 1.71 to 2.09 pairs. The number of pairs in all the bins in RR at $\sigma = 0 - 1$ h_{70}^{-1}Mpc have a value of 1.81 ± 0.01.

The DD/DR-1 correlation plot with the RMS error is shown in Figure 5.11. When binned to this level we can conclude the following.

- For low column density absorbers $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2})) \leq 17.4$, there was no evidence for correlation between the Lyα absorbers and galaxies. The maximum

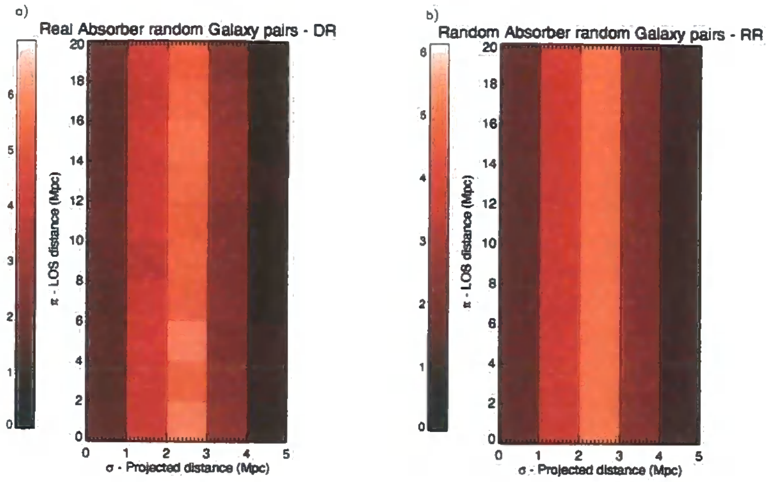


Figure 5.10 These plots show the difference between the DR and RR samples to use in the correlation function. The completely random RR pairs did have a smaller variance along bins of equal projected separation. Apart from this the numbers were very similar, so the difference between using DR or RR did not adversely effect the correlation result.

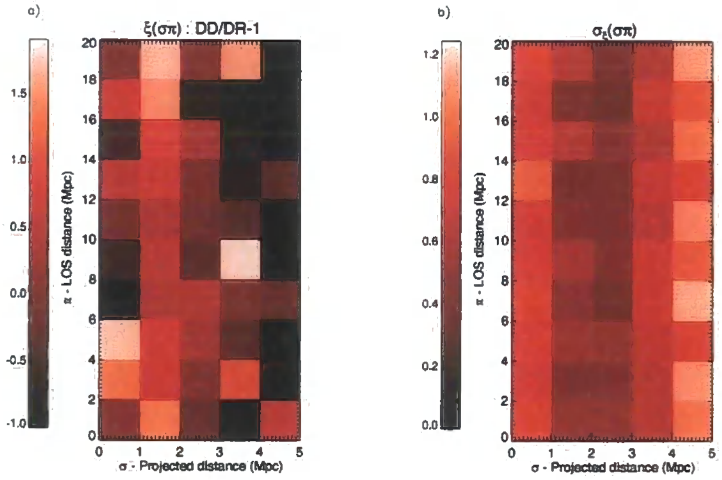


Figure 5.11 The 2D 2-point correlation function $\xi_{\mathrm{AG}} = \mathrm{DD/DR}-1$. There was no evidence for correlation between galaxies and low column density ($\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\leq 17.4$) Ly$\alpha$ absorbers within these distances at this level of binning. A similar result to Figure 2 of Wilman et al. (2007) was that there was tentative evidence that low column density absorbers are more strongly associated with galaxies away from the central bin. However the degree of correlation was the same when the errors were considered. The bin at $\pi = 8-10$, $\sigma = 3-4$ h_{70}^{-1}Mpc had a value $\xi_{\mathrm{AG}} = 1.80\pm0.73$, at the central bin $\xi_{\mathrm{AG}} = 0.07 \pm 0.85$.

value in Figure 5.11 was $\xi_{AG} = 1.80 \pm 0.73$ at $\pi = 8-10$, $\sigma = 3-4$ h_{70}^{-1}Mpc. So within 20×5 h_{70}^{-1}Mpc a 3σ upper-limit for the cross-correlation was $\xi_{AG} = 2.19$.

The main problem was that there were too few real pairs in each bin. With these low numbers the peaks in the correlation function, or bins that exhibited an anti-correlation were not significant. For example the bin with the peak in correlation contained 8 real pairs (2.86 random). 10 bins contained no pairs with $\xi_{AG} = -1$. One of these bins occurred at $\pi = 16 - 18$, $\sigma = 3 - 4$ h_{70}^{-1}Mpc with the minimum value of $\xi_{AG} = -1.0 \pm 0.71$. So there was a $\sim 2\sigma$ variation in ξ_{AG} between the maximum and minimum values along the line of sight at $\sigma = 3 - 4$ h_{70}^{-1}Mpc. A similar variation in ξ_{AG} was also seen at $\sigma = 0 - 1$ h_{70}^{-1}Mpc across the same range in values. The values of ξ_{AG} varied from $\xi_{AG} = -1.0 \pm 0.87$ to $\xi_{AG} = 1.71 \pm 0.82$. Comparable values of ξ_{AG} were also found at $\sigma = 1 - 2$ h_{70}^{-1}Mpc, with a peak of $\xi_{AG} = 1.70 \pm 0.64$ all the way out at $\pi = 18 - 20$ h_{70}^{-1}Mpc. So there were no significant trends observed in the degree of correlation along either axis.

- Like Wilman et al. (2007) a common result was the off-origin maximum in the correlation for low density absorbers ($\log_{10}(N_{HI}\,(cm^{-2})) = 13 - 15$), although this result was also equivalent to 0. The correlation function at the bin of closest separation in Figure 5.11 was $\xi_{AG} = 0.07 \pm 0.85$. This was only 1.4σ lower than the peak along $\sigma = 0-1$ h_{70}^{-1}Mpc. However that both investigations hinted at a drop in correlation at small separation was tentative evidence that low-column density Lyα absorbers are not observed at short distances. Therefore either they are not present, or they are swamped in the vicinity of galaxies by the absorbers with $\log_{10}(N_{HI}\,(cm^{-2})) \geq 17$. As their column density is so low these absorbers are easily photo-ionised and stripped of their neutral gas when they enter the halo. Since only one absorber with a column density of $\log_{10}(N_{HI}\,(cm^{-2})) \geq 17$ was observed in this dataset this supported the idea that low column density absorbers are depleted this way when near to a galaxy.

Figure 5.12 shows the correlation plot of $\xi_{AG} = DD/RR - 1$. As was expected there was no discernible difference in the degree of correlation when one used the DR or RR pairs in the denominator.

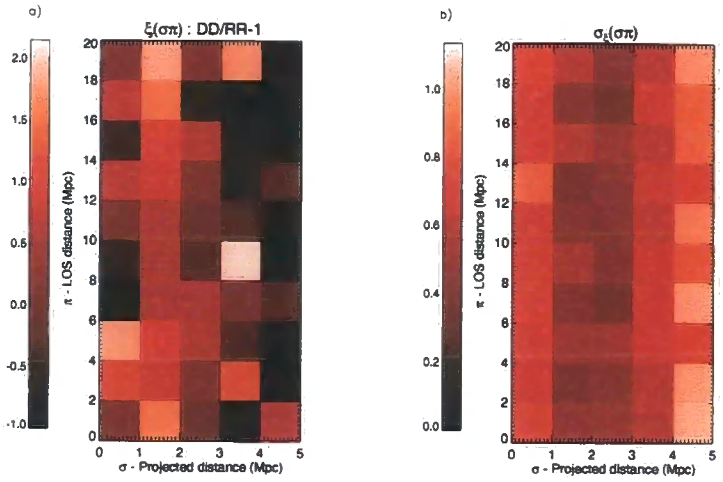


Figure 5.12 The 2D 2-point correlation function $\xi_{\mathrm{AG}} = \mathrm{DD/RR} - 1$. This plot shows that there was no significant difference between using the DD/DR or DD/RR estimator when computing the correlation function. There was no significant evidence of correlation between galaxies and Lyα absorbers of low column density at these distances. A 3σ upper-limit of $\xi_{\mathrm{AG}} = 2.01$ was computed using the same bin as was stated for the peak in Figure 5.11 at $\pi = 8 - 10$, $\sigma = 3 - 4$ h_{70}^{-1}Mpc.

The peak in correlation, which was in the same bin as the peak in the DD/DR plot, now had a value of $\xi_{\mathrm{AG}} = 2.05 \pm 0.67$. This was equivalent to the 3σ upper-limit of $\xi_{\mathrm{AG}} = 2.01$, tentative evidence of correlation away from the central bin. However the same conclusion was reached that there was no significant evidence of any correlation between galaxies and Lyα absorbers of a low column density. As there was no significant difference between using the DR or RR estimator all of the subsequent correlation results discussed will have used equation (2.2b).

When using the Landy and Szalay (1993) estimator in equation (2.5) most of the bins produced the same result as those in Figures 5.11 and 5.12. However the calculation failed in bins where there were very few or no pairs present. All of the correlation estimators normalise the number of pairs so that if DD was equal to 0 then $\xi_{\mathrm{AG}} = -1$. When adapting equation (2.5) for the cross-correlation the result in some bins was less than -1. This implied that once normalised $\mathrm{DD} + 2\mathrm{RR} < 2\mathrm{DR}$.

Therefore we suggest that the normalisation constants used in equation (2.5) are not valid for a correlation that studies two populations. This may have been because the second term $N^i_{dr} = n_q/n_g$ was independent of the number of absorbers.

Invalid results were also returned when we added an additional term that considered the number of random absorbers and real galaxy pairs (RD), so that the adapted version of equation (2.5) became,

$$\xi_{AG}\,(\sigma,\pi) = \frac{N_{rr}DD - N_{rd}RD - N_{dr}DR + RR}{RR}$$

With

$$N^i_{rd} = \frac{n^i_g n^i_p}{n^i_a n^i_g} = \frac{n^i_p}{n^i_a}$$

Regardless of which estimator was used, in all of the bins the errors were considerable because of the small number of real pairs. In addition to this our effective range in column densities was only $\log_{10}(N_{HI}\,(cm^{-2}))= 13-15$, and only 13 absorption dominated galaxies were detected in the survey. Therefore we did not carry out further investigations by testing the cross-correlation dependency on redshift, absorber column density, or galaxy type as this would have sub-divided an already small data sample.

5.2.6 Comparison with the Galaxy Auto-Correlation Function

A method of gauging the strength in correlation between Lyα absorbers and galaxies was to compare $\xi_{AG}\,(\sigma,\pi)$ with the 2D galaxy auto-correlation function $\xi_{GG}\,(\sigma,\pi)$. All of the literature results point to the galaxy auto-correlation peaking at the smallest separation, then decaying in a distorted radial pattern (Coil et al., 2006, Hawkins et al., 2003, Pollo et al., 2005). Figure 5.13 is Figure 4 of $\xi_{GG}\,(\sigma,\pi)$ from Hawkins et al. (2003). This is data from the 2dF Galaxy Redshift survey (2dFGRS) in this instance binned in redshift space out to $\pi = \sigma = 40$ h^{-1} h^{-1}Mpc.

In all the plots of Figure 5.13 the galaxy auto-correlation peaked at $\pi = \sigma \leq 1$ h^{-1}Mpc inside the first contour line with a value of $\xi_{GG} \gtrsim 4.0$. Redshift-space distortions perturbed the radial decay of ξ_{GG}. On small scales elongation along π

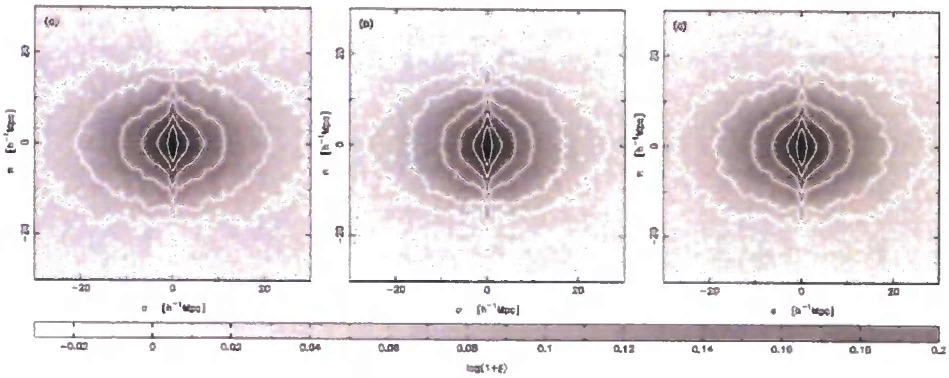


Figure 5.13 These 3 plots show ξ_{GG} (σ,π) from the 2dFGRS catalogue of $\sim$ 220000 galaxies. ξ_{GG} has been calculated using equation (2.5). Plots (a) and (b) are of the two separate fields used in the survey, the South Galactic Pole and North Galactic Pole respectively. The third plot is of the combined data. The error in ξ_{GG} (σ,π) was estimated from mock catalogues but was not plotted in Hawkins et al. (2003). However the S/N was very high on scales less than 20 h^{-1}Mpc. The white contour lines mark out ξ_{GG} at $\xi_{GG} = 4.0, 2.0, 1.0, 0.5, 0.2$ and 0.1. There was the expected peak at the bin of smallest separation and this decayed in a radial fashion that was perturbed by redshift-space distortions. The flattening of the contour lines on large scales along σ was caused by the flow of gas from underdense to overdense regions. The extension along π at small scales was the "finger-of-god" that is caused by peculiar velocities along the line of sight. This figure is Figure 4 from Hawkins et al. (2003).

was caused by peculiar velocities along the line of sight. This is the "finger-of-god" effect.

Figure 5.14 shows our real 266 galaxy pairs, DD binned in redshift space, and the random galaxy pairs RR. The method for producing these pairs followed the exact same procedure that was described in Section 5.2.1.

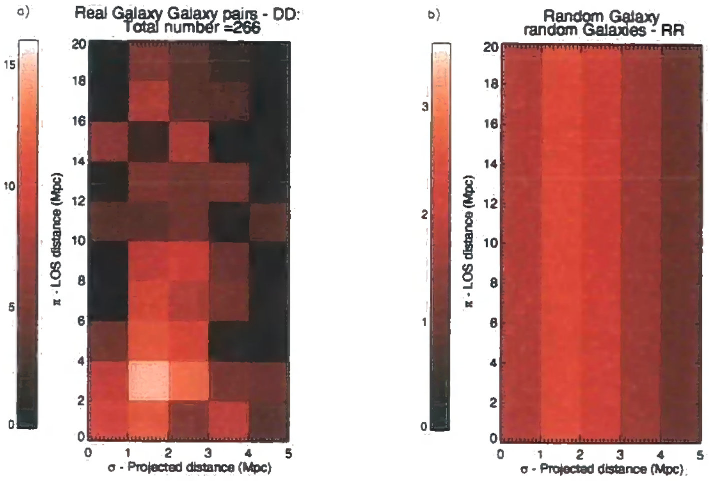


Figure 5.14 The real galaxy pairs (DD) and the random galaxy pairs (RR) used in the auto-correlation ξ_{GG}. The decline in the number of real and random pairs above $\sigma = 2$ h_{70}^{-1}Mpc is because of edge effects that was discussed in Section 5.1.2.

Unlike larger tests of the auto-correlation, such as the $\sim$ 220000 redshifts that were used in the 2dFGRS, we only used 198 galaxies (in the redshift range 0.2-1.5) that gave 266 pairs within the 20×5 h_{70}^{-1}Mpc space. Inevitably this meant that our auto-correlation suffered from Poisson noise.

This was evident in Figure 5.15 which shows the auto-correlation ξ_{GG}. Contrary to the result in Figure 5.13 and Figures 2 and 3 of Ryan-Weber (2006) and Wilman et al. (2007) respectively the closest bin did not contain the peak in the auto-correlation. Instead this occurred at a separation of $\pi = 2 - 4$, $\sigma = 2 - 3$ h_{70}^{-1}Mpc, with a peak of $\xi_{GG} = 4.84 \pm 0.77$. The bin of smallest separation had a value of $\xi_{GG} = 4.41 \pm 1.03$. Therefore this discrepancy was not significant. However the much lower value ξ_{GG} at these distances when compared to Hawkins et al. (2003) does require an explanation if the auto-correlation is to match the result from previous works and provide a way in which to contrast the result with the cross-correlation.

On this small scale and with few pairs we do not see evidence for any redshift-space distortion along the line of sight. However the uncertainty in ξ_{GG} is considerably greater than any effect this may have had.

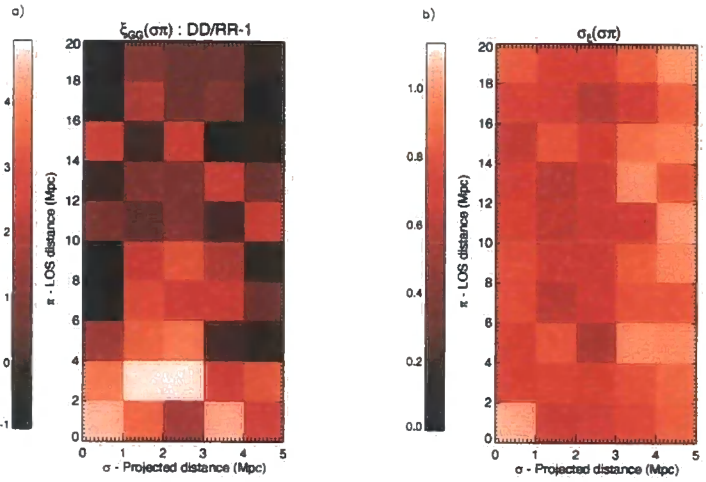


Figure 5.15 The galaxy auto-correlation function, ξ_{GG} (σ,π). The peak in the auto-correlation function was not in the bin of smallest separation. However the value of the peak when compared to the bin at $\pi = 0-2$, $\sigma = 0-1$ h_{70}^{-1}Mpc is not significantly different when we compare the values $\xi_{GG} = 4.84 \pm 0.77$ and $\xi_{GG} = 4.41 \pm 1.03$.

The third estimator in equation (2.5) was also calculated for the auto-correlation, and all of the bins had a $< 1\sigma$ difference to the values in Figure 5.15.

Figure 5.16 is of the 2dFGRS data re-binned to match that in Figure 5.15 so that the values for ξ_{GG} could be compared. Any differences that arose from using the two different estimators, equations (2.2b) and (2.5) were negligible when compared to the error array of Figure 5.15[1].

Not only was the peak at the centre of Figure 5.16 but a significant "finger-of-god" was observed along the line of sight. Binned to this level it was known that the signal-to-noise was greater than 6 (Hawkins et al., 2003) meaning that at the central bin $\xi_{GG} = 17.4 \pm 2.9$. This was a 4σ increase on the value found in the corresponding bin in Figure 5.15. By $\pi = 8-10$, $\sigma = 2-3$ h_{70}^{-1}Mpc the results were

---

[1] Data supplied by Dr. P. Norberg, Institute for Astronomy, University of Edinburgh.

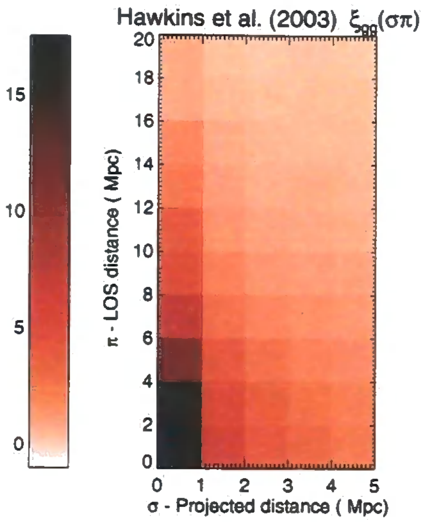


Figure 5.16 The 2D 2 point auto-correlation ξ_{GG} of the 2dFGRS from Hawkins et al. (2003), this was re-binned to $\Delta\pi = 2.0$, $\Delta\sigma = 1.0$ h_{70}^{-1}Mpc. When compared to Figure 5.15 the values for ξ_{GG} were much higher at small σ and π with a central peak of $\xi_{GG} = 17.4 \pm 2.9$. There is also a significant "finger-of-god" redshift distortion along $\sigma = 0 - 1$ h_{70}^{-1}Mpc. Data is from Hawkins et al. (2003).

more comparable with only a 2σ difference.

The main reason why our galaxy correlation was so different was because of the low number of real pairs and in particular those that were close.

Many close pairs would have been removed from our galaxy survey because of the mask designs. Using Source Extractor, 1062 and 1122 potential targets were identified in the pre-imaging of fields HE 1122-1648 and PKS 1127-145 respectively, with approximately 400 galaxies in each field allocated a slit. Of these 20 mask designs only 9 were observed that had useful data. So at most there were 5 masks observed along a particular sightline with large differences in the observing conditions for each one, implying a different success rate. With such an incomplete survey it was not surprising that there were so few galaxy pairs observed that were in close proximity to one another.

A way to overcome losses because of slit collisions would have been to use an additional projected separation dependent weighting factor when calculating the correlation function, a method that was employed by both Pollo et al. (2005) and Hawkins et al. (2003).

An additional problem which will have contributed to the peculiar auto-correlation is illustrated by Figure 5.1. The redshift distribution of the real galaxies was heavily weighted around z ~ 0.8. However the redshift distribution of the random galaxies was constant, because when assigning the random redshifts these could take any value between the bounds of the real data. Therefore because there were many DR and RR pairs at redshifts that could contain only few DD pairs, when DD was divided by DR or RR the correlation function was unevenly weighted.

5.2.7 The Perturbed Redshift Method

To overcome this offset in the galaxy auto-correlation function what was required was a method whereby the random galaxies would have the same redshift distribution as the real galaxies. This was achieved by perturbing each real galaxy redshift by a random amount that varied between $\Delta z = \pm\ 0.01 - 0.1$. These boundaries were chosen so that the large scale galaxy redshift distribution was maintained, but any real galaxy pair that had existed would now no longer appear within the $20 \times 5\ h_{70}^{-1}$Mpc window of the auto-correlation function. This method of generating random catalogues that mimicked the distribution of the real data was also used when generating the DR and RR pairs for the 50 mock catalogues in the error calculation. The plots in Figure 5.17 show the result of this approach.

Plot (a) of Figure 5.17 was a repeat of the DD galaxy pairs from Figure 5.14 that has been redrawn here for convenience. Figure 5.17 plot (b) is of the new RR pairs using the perturbation method. When compared with the RR pairs plot in Figure 5.14 the key changes were the increase at an impact parameter of $\sigma = 2 - 3$ h_{70}^{-1}Mpc, and a decrease at $\sigma = 0 - 1\ h_{70}^{-1}$Mpc. This had the desired effect of reducing the problem of the galaxy distribution. This resulted in the auto-correlation in plot (c) of Figure 5.17, where the peak in the auto-correlation was now at the smallest separation with a value of 3.89 ± 0.65. However despite this improvement ξ_{GG} was still not comparable with the peak observed from Hawkins et al. (2003).

An alternative way to convey the different results from Hawkins et al. (2003)

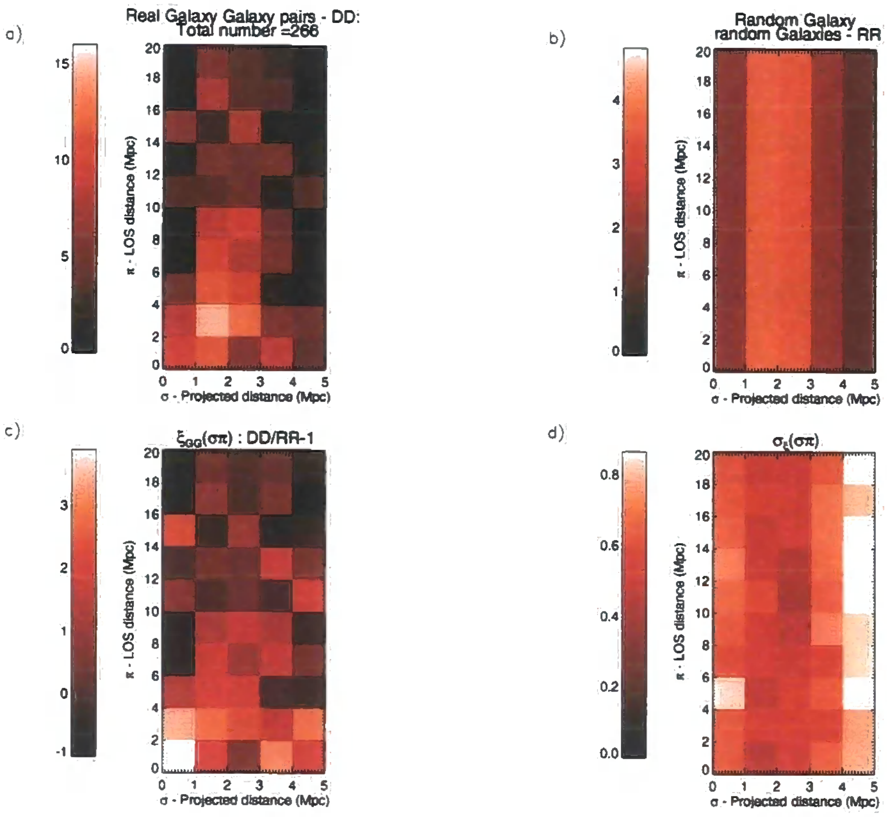


Figure 5.17 Plots (a) and (b) show the DD and RR pairs that were used for the auto-correlation function. The redshifts for the RR pairs were created by perturbing each real redshift by $\Delta z = \pm\, 0.01 - 0.1$. This compensated for the peculiar galaxy distribution. Plots (c) and (d) show $\xi_{\mathrm{GG}}\,(\sigma,\pi)$ and the associated error. Now that we have altered the RR distribution the peak in the auto-correlation was now in the bin of smallest separation. Here the value was $\xi_{\mathrm{GG}} = 3.89 \pm 0.65$. However it was no more significant than the other peaks at $\pi = 2-4$, $\sigma = 0-1$ and $\pi = 0-2$, $\sigma = 3-4$ h_{70}^{-1}Mpc.

and our survey was to plot the redshift-space correlation parameter ξ_{GG} (s), where,

$$s = \sqrt{(\pi^2 + \sigma^2)}$$

Figure 5.18 highlights the immense difference between ξ_{GG} (s) of the 2dFGRS and our survey. (The data has been re-binned and the distance scale amended with h=0.7). Although our results had the correct characteristic shape, Figure 5.18 shows that the auto-correlation was consistently too low when compared with Hawkins et al. (2003) at small separation. Our result for ξ_{GG} (s) only became comparable with the literature result beyond a distance of s = 5 h_{70}^{-1}Mpc.

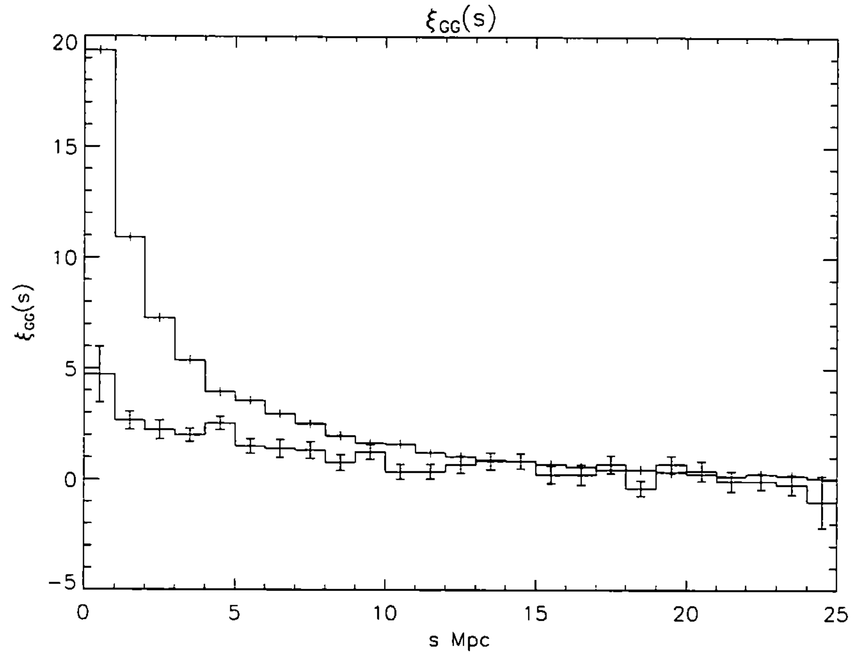


Figure 5.18 This comparison with Hawkins et al. (2003) of the auto-correlation parameter ξ_{GG} (s) shows that our value was only comparable with that of the 2dFGRS result at s $\geq$ 5 h_{70}^{-1}Mpc.

Therefore by using the perturbation method, even though it had to a small extent compensated for the uneven galaxy distribution, it had not weighted those pairs at the smallest separation a sufficient amount. However given the small nature of our galaxy survey a robust result that could be compared with ξ_{GG} from Hawkins et al. (2003) cannot be expected.

A direct comparison with 2dFGRS was also not entirely valid as it is known that the galaxy auto-correlation decreases with increasing redshift (Pollo et al., 2005). Our pairs came from a sample of 198 galaxies that covered the redshift range $z = 0.2 - 1.5$, with most of the pairs occurring at $z \sim 0.8$. The 2dFGRS contained ~ 220000 redshifts from $z = 0.01 - 0.2$.

An alternative comparison was made with the results from the VVDS-Wide survey (Pollo et al., 2005). Figure 5.19 (Guzzo et al., 2008) shows ξ_{GG} calculated using the Landy and Szalay (1993) estimator from a sample of $\sim$ 6000 galaxy redshifts with $0.6 \leq z \leq 1.2$. These have been binned by 1 h^{-1}Mpc along both the π and σ axis and then smoothed with a Gaussian kernel. Here we see a peak of $\xi_{GG} \lesssim 10$ in the central bin and a large "finger of god" along the line of sight. Thus the auto-correlation was less than what was observed in the 2dfGRS survey, but still considerably more than our peak of 3.89 ± 0.65. Like Hawkins et al. (2003), Pollo et al. (2005) also weighted the DD pairs to account for pair loss on small scales.

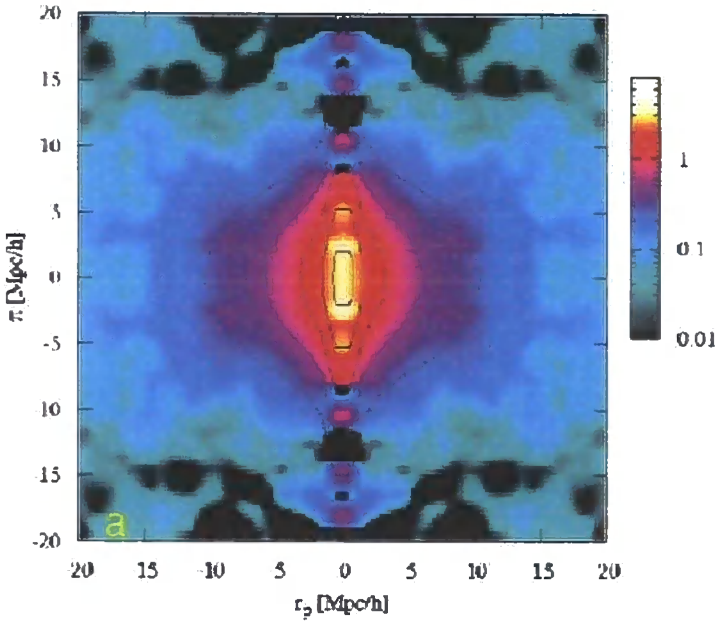


Figure 5.19 To compare with the 2dFGRS and our redshift survey this is ξ_{GG} (σ,π) measured using $\sim$ 6000 galaxy redshifts from the VVDS-Wide Survey. Pairs have been binned 1 h_{70}^{-1}Mpc along both the π and r_p (σ) axis. This sample shows that ξ_{GG} is lower at higher redshift $(0.6 \leq z \leq 1.2)$ than the 2dFGRS sample, with a peak at the centre of $\xi_{GG} \lesssim 10$. This figure is from Guzzo et al. (2008).

Therefore these alternative galaxy surveys provided evidence of strong galaxy-galaxy clustering on scales ≤ 1 h^{-1}Mpc that decayed to $\xi_{GG} \sim 2$, a value comparable with the peak in ξ_{AG} by $\sigma \gtrsim 5$ h^{-1}Mpc.

5.2.8 The Perturbation Technique for the Cross-Correlation

The technique of using perturbed rather than random redshifts was then used to recalculate the cross-correlation in order to look for changes. This time to ensure that all real pairs would be removed from the correlation space the galaxy redshifts were perturbed by a distance between $\Delta z = \pm\ 0.01-0.06$. Each random absorber redshift was also generated by perturbing the redshift of each real absorber $\Delta z = \pm\ 0.06-0.1$. Figure 5.20 shows the results for the new DD/RR estimator.

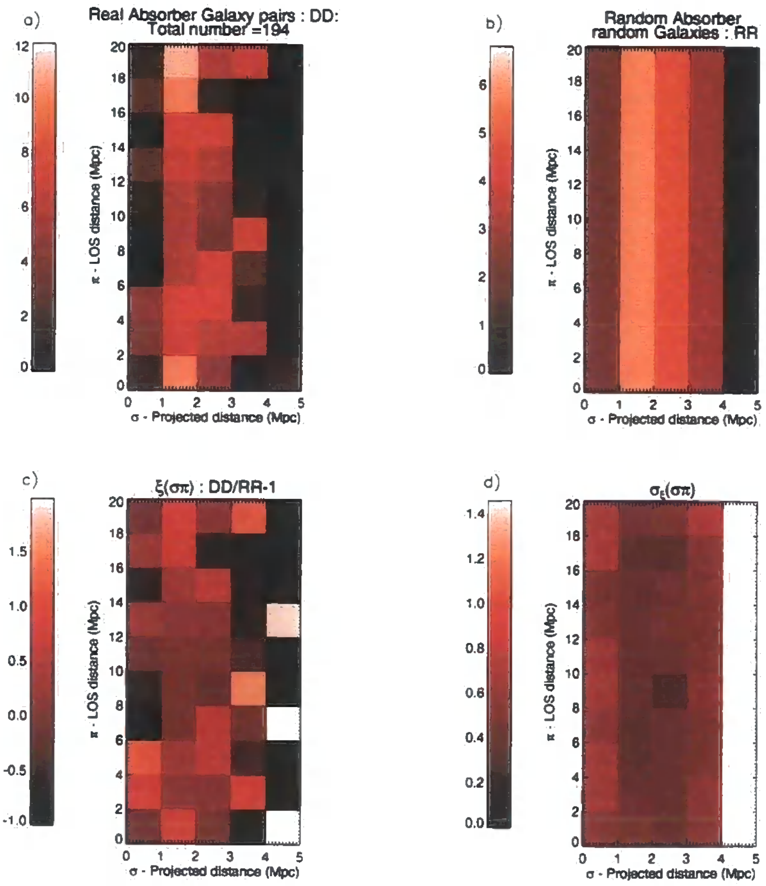


Figure 5.20 Plot (a) is a repeat of Figure 5.9 of the 194 absorber-galaxy real pairs. Plot (b) is of the RR set, where both absorbers and galaxies have been perturbed by values $\Delta z = \pm\ 0.06 - 0.1$ and $\Delta z = \pm\ 0.01 - 0.05$ respectively. This was to observe if using this perturbation method made any difference to the result in Figure 5.12. No significant change in ξ_{AG} was observed in plot (c) to that seen in Figure 5.12. The central bin from Figure 5.12 when using the random method and the RR pairs was $\xi_{AG} = 0.11 \pm 0.78$. When using the perturbed method $\xi_{AG} = -0.11 \pm 0.58$.

Once again the RR distribution of plot (b) in Figure 5.20 should be compared with the RR distribution in Figure 5.10. By using the perturbation method the distribution of the RR pairs had changed. This time the density of perturbed absorber-galaxy pairs in the smallest impact parameter column had increased from 1.81 ± 0.02 to 2.23 ± 0.02. It was the bins at $\sigma = 1 - 2$ h_{70}^{-1}Mpc that have changed the most. When randomised the bins at $\sigma = 1 - 2$ h_{70}^{-1}Mpc had a value of 3.90 ± 0.02, for the perturbed RR pairs this value increased to 5.78 ± 0.02.

Despite these variations in the distribution of the RR pairs when comparing the random and perturbation methods, the changes in the degree of cross-correlation in corresponding bins at all separations were insignificant. Using the perturbation technique the maximum value for the cross-correlation of $\xi_{\rm AG} = 1.53 \pm 0.61$ provides a 3σ upper-limit of $\xi_{\rm AG} = 1.83$.

When $\xi_{\rm AG}$ was compared with $\xi_{\rm GG}$ in Figures 5.20 and 5.17 we could draw the following conclusion. The values of the central bin in both Figures had a 4.6σ difference in the level of correlation. Therefore there was evidence of a difference in the cross-correlation between galaxies and low column density Lyα absorbers and the galaxy auto-correlation within 1 h_{70}^{-1}Mpc. When we compared the cross-correlation with the auto-correlation from the VVDS-Wide Survey and 2dFGRS, this difference in $\xi_{\rm AG}$ and $\xi_{\rm GG}$ dramatically increased beyond 5σ.

In all of these plots where the method of perturbing redshifts was used the DR estimator mimicked the changes observed for the RR pairs so were not included.

5.3 Variation in the Binning and the 1D Correlation Function Along π and σ

5.3.1 Binned 2x2 h_{70}^{-1}Mpc

To improve on the significance of these findings it was necessary to increase the size of each bin. To this end the bin size was increased to $\Delta\pi = \Delta\sigma = 2$ h_{70}^{-1}Mpc. The number of real pairs in each bin was now approximately double the value that it had been previously, however the error in the correlation was still comparable with the data in most bins. All of the pairs were within $\sigma = 5$ h_{70}^{-1}Mpc therefore these new plots only extend out to an impact parameter of 4 h_{70}^{-1}Mpc to avoid too few pairs along the final column.

Figures 5.21 and 5.22 show ξ_{AG} and ξ_{GG} respectively. These used the perturbation technique to generate the random galaxies and the RR estimator has been used.

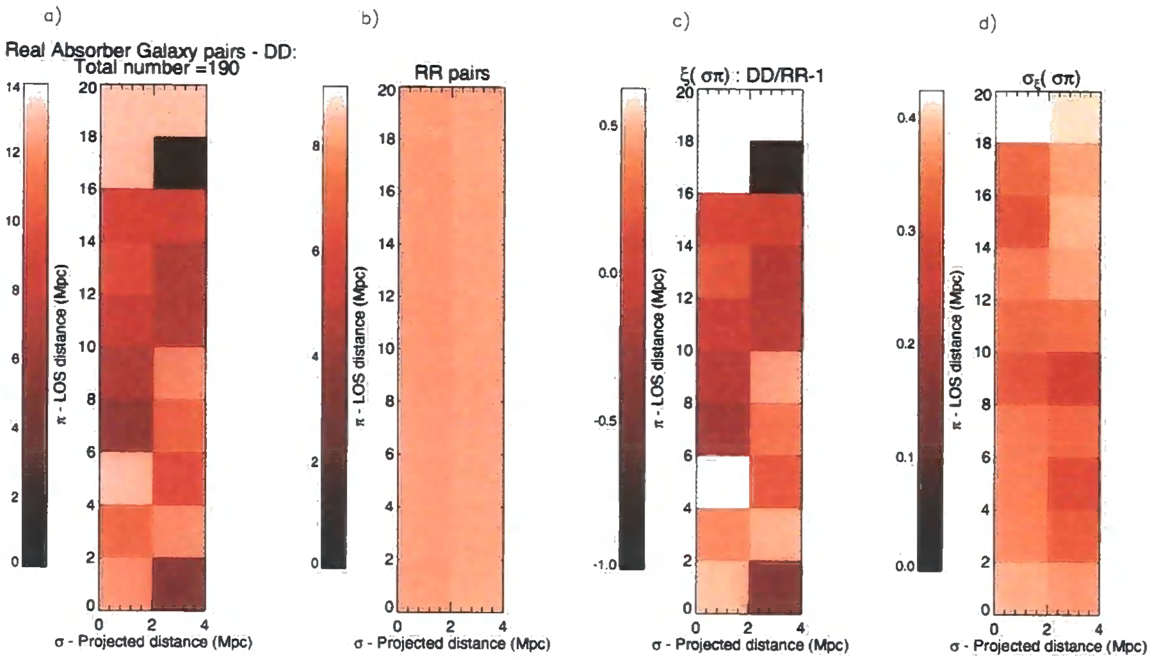


Figure 5.21 Plots a-d are of the DD pairs, RR pairs, ξ_{AG} (σ,π) and error respectively when the level of binning has increased to $\Delta\pi = \Delta\sigma = 2$ h_{70}^{-1}Mpc. The error per bin decreased and the peak in correlation was now 0.62 ± 0.36 at $\pi = 4 - 6$, $\sigma = 0 - 2$ h_{70}^{-1}Mpc. The result was the same as to what was observed in Figure 5.12. This negligible degree of correlation also did not change significantly along the line of sight as the the value varied by 1σ only. Values for ξ_{AG} also did not significantly change when at a projected separation of $\sigma = 2 - 4$ h_{70}^{-1}Mpc.

Figure 5.21 supports the conclusions made earlier. The maximum value, not at the central bin, was now at $\pi = 4-6$, $\sigma = 0-2$ h_{70}^{-1}Mpc with a value of 0.62 ± 0.36. However this level of correlation was comparable with all the bins along $\sigma = 0-2$ h_{70}^{-1}Mpc as ξ_{AG} remained within 1σ of this value. Therefore when binned to this extent there was no correlation detected between Lyα absorbers and galaxies, and this level of association was constant along the line of sight. When binned to this level we found a 3σ upper-limit of $\xi_{AG} = 1.08$.

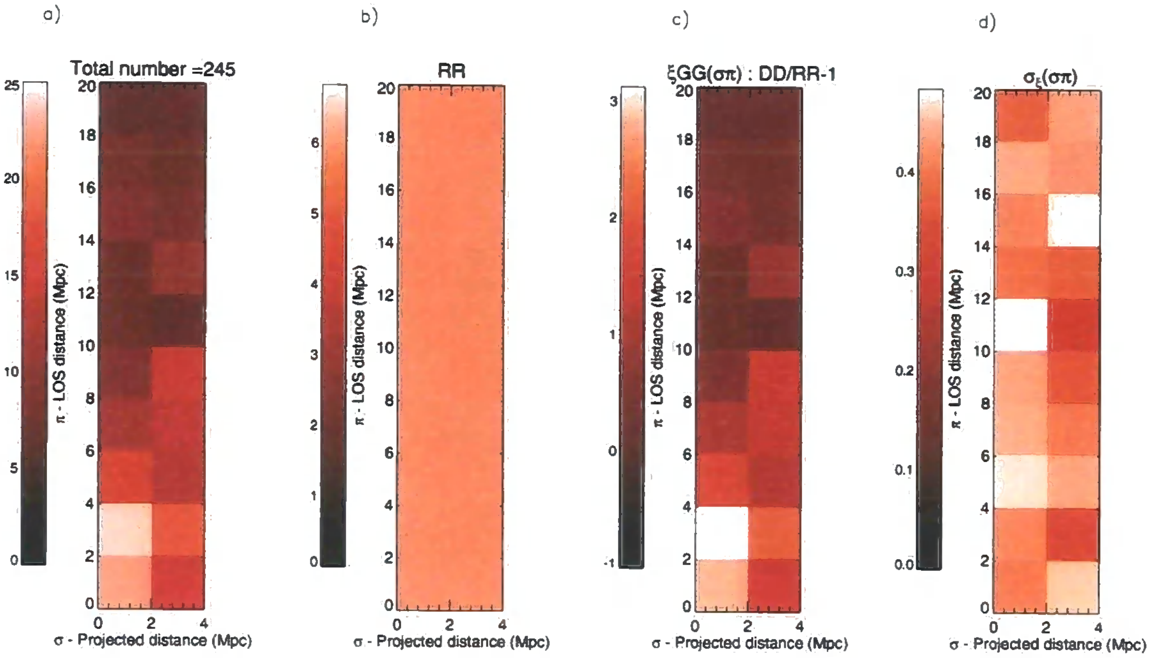


Figure 5.22 When binned $\Delta\pi = \Delta\sigma = 2$ h_{70}^{-1}Mpc the same conclusions could be drawn for the auto-correlation function. Correlation between galaxies decreased along the line of sight, until after $\pi = 6$ h_{70}^{-1}Mpc the correlation signal was comparable with ξ_{AG} (σ,π).

Unlike Figure 5.17 the peak in the auto-correlation of Figure 5.22 was now no longer at the origin. Rather it was along the line of sight at $\pi = 2 - 4$ h_{70}^{-1}Mpc. However this was not a significant change as the two values were 2.78 ± 0.40 and 3.12 ± 0.41, with only 2 extra real pairs at $\pi = 2 - 4$. Even when binned to this level, where averaging over 2 h_{70}^{-1}Mpc along σ has caused ξ_{GG} to decrease, we can still claim that in the central bin there was stronger clustering amongst galaxies than between galaxies and Lyα absorbers at a 4.1σ level of confidence.

5.3.2 The 1D Cross and Auto-Correlation Function

In order to further reduce the Poisson error and increase the certainty in the correlation function we collapsed the pair counts along each dimension entirely. All of the 1D plots shown below used the perturbation method to generate random absorbers and galaxies and used the DD/RR-1 estimator.

Summed along the line of sight

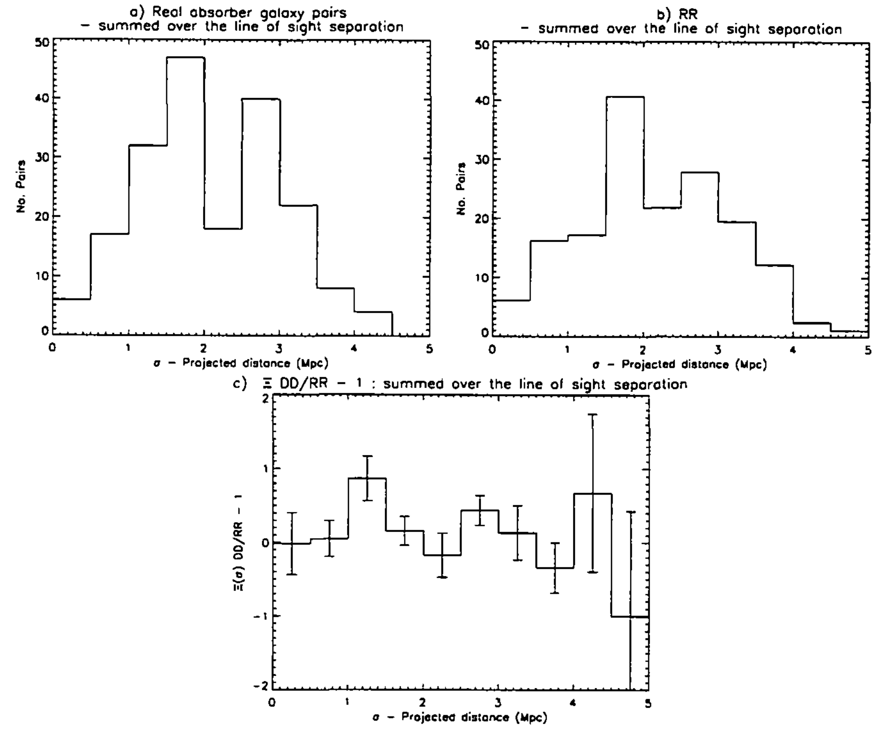


Figure 5.23 Plots (a) and (b) show the number of DD and RR absorber-galaxy pairs when they were binned along the projected separation. The shape of these histograms were a consequence of the edge effects discussed and illustrated in Figure 5.3. The lack of pairs at $\sigma = 2.5$ h_{70}^{-1}Mpc was compensated in plot (b) by using the perturbed redshift technique. Plot (c) suggested an off-centre peak in the cross-correlation at $\sigma = 1 - 1.5$ h_{70}^{-1}Mpc. This peak was not significant with an increase of only 1σ from $\sigma = 0 - 0.5$ h_{70}^{-1}Mpc. There was no significant change in $\Xi_{AG}(\sigma)$ out to $\sigma = 4$ h_{70}^{-1}Mpc.

Plots (a), (b) and (c) of Figure 5.23 show the number of real pairs, the number of random pairs and ξ_{AG} when binned along the projected separation with $\Delta\sigma = 0.5$ h_{70}^{-1}Mpc. This was the projected correlation $\Xi(\sigma)$ where;

$$\Xi(\sigma) = \int_{0}^{\infty} \xi(\sigma,\pi)\, d\pi$$

In practice the upper limit was taken to be the maximum separation of the pairs when binned in the 2D plots, π = 20 h_{70}^{-1}Mpc. The absolute separation between the absorbers and galaxies was used, therefore the lower limit is zero.

By summing over the line of sight we also removed any peculiar velocity distortions since these only displaced pairs along the π axis.

The general shape of plot (a) in Figure 5.23 supported what has already been mentioned in Section 5.1.2. As the survey volume increased the number of galaxies, hence pairs increased also. This peaked between 1-2 h_{70}^{-1}Mpc before decreasing because galaxies were now located at the edge of the chip. The fewer number of pairs at $\sigma = 2 - 2.5$ h_{70}^{-1}Mpc was real and can be deduced from the Pie-diagrams in Figures 5.7 and 5.8. In the field of view of HE 1122-1648 there were only 4 galaxies that were at a separation of $2 - 2.5$ h_{70}^{-1}Mpc. In the field of PKS 1127-145 there were 20, however only 4 of these were between the same redshift range as the absorbers.

The problem of this distribution was overcome in Figure 5.23 plot (b). By perturbing the redshifts we were able to simulate the same absorber-galaxy pair distribution for the random pairs.

Plot (c) of Figure 5.23 suggests that the peak in the absorber-galaxy cross-correlation was not at the smallest impact parameter. Rather the peak was off-set at $\sigma \gtrsim 1.5$ h_{70}^{-1}Mpc. However this peak was not significant with an increase of only 1σ from $\sigma = 0 - 0.5$ h_{70}^{-1}Mpc. The level of correlation then remained the same as there was no significant change in $\Xi_{AG}(\sigma)$ out to $\sigma = 4$ h_{70}^{-1}Mpc. The 2 bins above $\sigma = 4$ h_{70}^{-1}Mpc were ignored in this comparison because only 4 real pairs were found at this distance.

The plots in Figure 5.23 can be contrasted with the galaxy auto-correlation. A similar explanation for the distribution in plots (a) and (b) of Figure 5.24 can be applied here. One difference is that unlike plot (c) of Figure 5.23 the peak in Ξ_{GG} was at $\sigma \leq 0.5$ h_{70}^{-1}Mpc. Due to the substantial error and our peculiar galaxy distribution this was only 0.5σ greater than the value at $\sigma = 2.5 - 3$ h_{70}^{-1}Mpc.

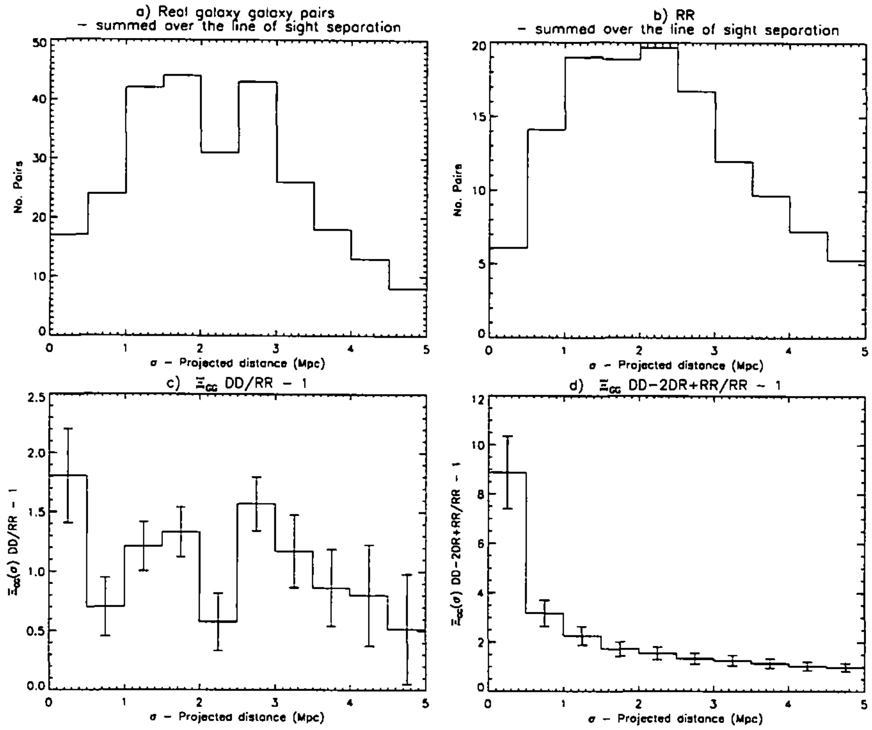


Figure 5.24 When binned along the projected separation the auto-correlation shown in plot (c) peaked as was expected at the shortest separation. However this was not even a 1σ peak when one compared to a similar value for $\Xi_{GG}(\sigma)$ at $\sigma = 2.5 - 3$ h_{70}^{-1}Mpc. Like the cross-correlation $\Xi_{GG}(\sigma)$ oscillated about 0 out to $\sigma = 4$ h_{70}^{-1}Mpc. Once again edge effects and the sample volume explain the number of DD and RR pairs in plots (a) and (b). Had the galaxy survey been more complete, plot (c) would look similar to $\Xi_{GG}(\sigma)$ from the Hawkins et al. (2003) data that has been plotted in (d).

$\Xi_{\rm AG}$, even though weaker at small separation, became comparable to $\Xi_{\rm GG}$ at $\sigma \geq 1$ h_{70}^{-1}Mpc. Plot (d) of Figure 5.24 is the results from Hawkins et al. (2003) re-binned to $\Delta\sigma = 0.5$ h_{70}^{-1}Mpc and summed over the line of sight. Using this we could see a definite peak below $\sigma = 0.5$ h_{70}^{-1}Mpc with $\Xi_{\rm GG} = 8.97 \pm 1.45$.

Summed along the projected separation

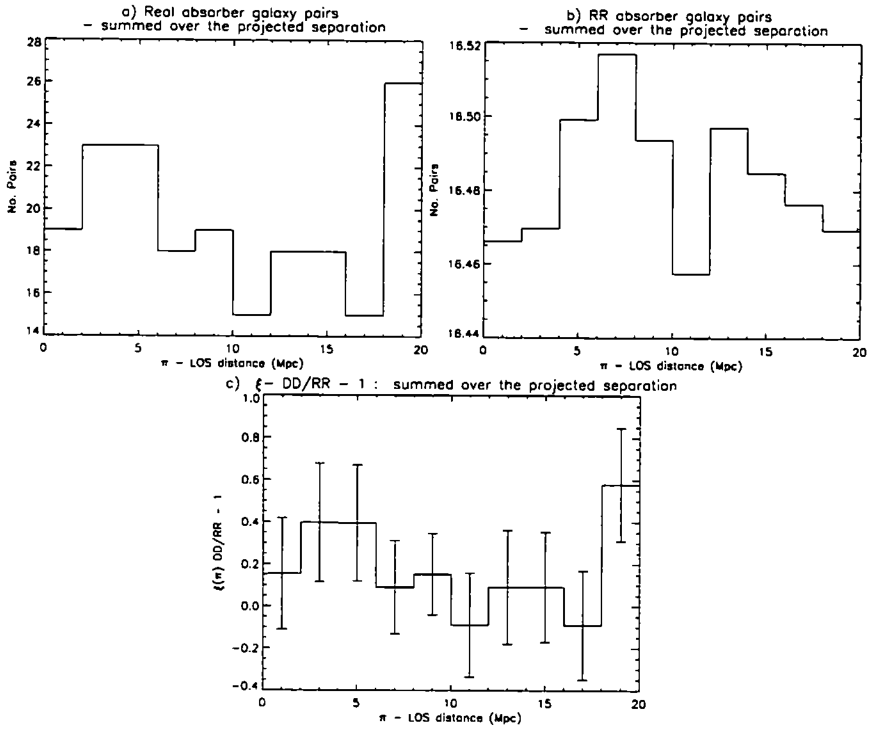


Figure 5.25 Plot (c) shows the cross-correlation, $\xi_{\rm AG}(\pi)$ as it is binned along the line of sight in steps of $\Delta\pi = 2$ h_{70}^{-1}Mpc. At these significance levels there was no evidence for any correlation between the galaxies and low column density absorbers. What is important is this lack of correlation was constant along the line of sight with a variance of less that 1σ as π increases from 0-20 h_{70}^{-1}Mpc.

When binned using $\Delta\pi = 2$ h_{70}^{-1}Mpc along the line of sight, the number of real absorber galaxy pairs, shown in Figure 5.25 plot (a) fluctuated within a Poisson variance about a mean value of 19.4. Now that each cylindrical bin along the line of sight occupied the same volume the number of random pairs observed in plot (b) of Figure 5.25 was also almost constant. The correlation function $\xi_{\rm AG}(\pi)$ in plot (c) of Figure 5.25 did not have any significant change in value as π increased from $0 - 20$ h_{70}^{-1}Mpc. There was no correlation, with a peak of $\xi_{\rm AG}(\pi) = 0.40 \pm 0.28$ and all other values remained within 1σ of this value.

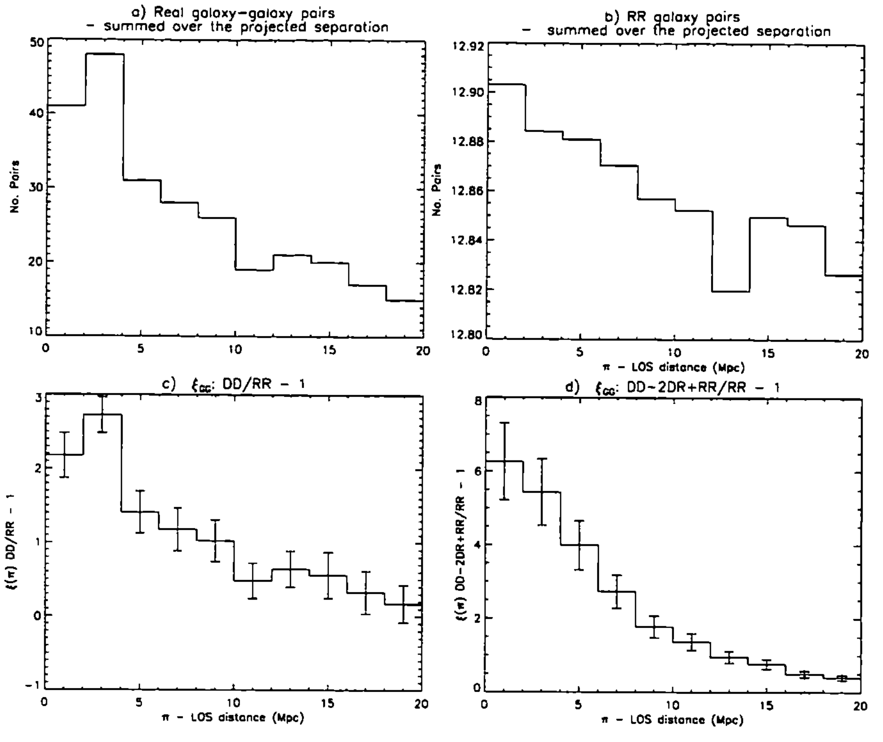


Figure 5.26 Unlike the cross-correlation there was a definite decrease in $\xi_{GG}(\pi)$ with distance. As π increased from 0-20 h_{70}^{-1}Mpc $\xi_{GG}(\pi)$ decreased 7σ from 2.73 ± 0.24 to 0.17 ± 0.25. The extent of the galaxy clustering along π was apparent when comparing the range in the number of DD and RR pairs in plots (a) and (b). The shape of (b) would have been flat had the redshifts been random. Since they are perturbed from their real positions a slight decrease in the number of RR was observed. This shows that beyond $\pi = 10$ h_{70}^{-1}Mpc the level of auto-correlation is very low, and when compared with Figure 5.25, comparable with the level of cross-correlation at similar distances. Plot (d) is of $\xi_{GG}(\pi)$ from the Hawkins et al. (2003) data.

Although there was no significant change in ξ_{AG} and the value was still a non-detection, conclusions could still be drawn when we compared this to the galaxy auto-correlation function that was shown in plot (c) of Figure 5.26.

Here we observed a drop in the auto-correlation. The value decreased from a peak of $\xi_{GG}(\pi) = 2.73 \pm 0.24$ down to 0.17 ± 0.25, as π increased from 2-20 h_{70}^{-1}Mpc, a decrease of 7σ. That this peak was not at $\sigma = 0 - 2$ h_{70}^{-1}Mpc was not significant as the two values are equivalent. The value of $\xi_{GG}(\pi)$ at small separation was less than what is usually observed in the literature. The result from Hawkins et al. (2003) is in plot (d) of Figure 5.26. This shows a 5.6σ decrease as π increased from 0-20

h_{70}^{-1}Mpc.

Therefore whether we consider our data or that from the 2dFGRS, it is evident that while there was a significant decrease in ξ_{GG} as π increased from 0-20 h_{70}^{-1}Mpc ξ_{AG} showed no evidence for any correlation across the same distance.

5.4 A Summary of Our Results to the 2 Point Correlation Function

We have measured the 2D 2-point correlation function between Lyα absorbers of a column density $\log_{10}(N_{HI}\,(cm^{-2}))= 13 - 17.4$ and galaxies using various estimators and bin sizes, and have compared this result to the galaxy auto-correlation.

The conclusions that we came to were:

- When binned over both the projected separation and line of sight out to 5 and 20 h_{70}^{-1}Mpc respectively with $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{70}^{-1}Mpc, there was no sign at a 3σ level for any correlation. The peak value at $\sigma = 3 - 4$, $\pi = 8 - 10$ h_{70}^{-1}Mpc gave a 3σ upper-limit of $\xi_{AG}\,(\sigma\pi) = 2.19$.

  The difference in cross-correlation between corresponding bins that used the formulae DD/DR-1 or DD/RR-1 were within $\pm 1\sigma$. Therefore $\xi_{AG}\,(\sigma\pi)$ remained the same regardless of whether a data-random or random-random catalogue of pairs was used in the calculation.

- We continued to get a null result if the method of generating random absorbers and galaxies was to perturb the real redshifts, so that the redshift distribution of both the random and real data were the same. The maximum value remained at the same location with a 3σ upper-limit of 1.83.

- The variance for most bins was comparable with the degree of correlation and the main reason for this was the low signal-to-noise caused by too few real pairs. Part of this problem originated from the small overlap between the redshift range of the galaxies and Lyα absorbers along each sightline. Therefore the level of binning was increased to $\Delta\sigma = \Delta\pi = 2$ h_{70}^{-1}Mpc. Binned to this extent our conclusions remained the same, there was no evidence for cross-correlation with a revised 3σ upper-limit of $\xi_{\mathrm{AG}}\,(\sigma\pi) = 1.08$.

- The cross-correlation was compared to the galaxy auto-correlation, for which we did see evidence of clustering. When binned $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{70}^{-1}Mpc the maximum value was $\xi_{\mathrm{GG}} = 3.89 \pm 0.65$. Earlier results did not have the peak at the central bin where it was expected. This offset was removed by generating the random galaxies by perturbing the redshifts of the real data.

  However our results were still lower at $\sigma = 0 - 1$, $\pi = 0 - 2$ h_{70}^{-1}Mpc than what was observed in both Hawkins et al. (2003) and Guzzo et al. (2008).

  An increase in the level of correlation at small distances is expected as we approach a redshift of zero because of gravitational collapse. Therefore it was not surprising when comparing Figures 5.16 and 5.19 that the central bin from the 2dFGRS survey was higher than that of the VVDS.

  The galaxy population that contributed to our value for ξ_{GG} came from redshifts between $\mathrm{z} = 0.2 \rightarrow 1.5$. So we expected a value that was comparable to both Hawkins et al. (2003) and Guzzo et al. (2008).

  Our value was too low because the galaxy survey was far from complete, and had too few real pairs particularly at small separation. A small number of pairs was meant to be compensated for by the random catalogue of galaxies. However our random catalogue, that was produced by perturbing the redshifts of the real galaxies, was insufficient to increase ξ_{GG} for the bins with $\sigma = 0 - 1$ h_{70}^{-1}Mpc.

  This situation was made worse by the difficultly in adequately sampling galaxies that were close when projected onto the mask because of overlapping slit spectra.

  A more complete analysis of ξ_{GG} such as Hawkins et al. (2003) compensated for these losses by using a weighted random catalogue when using the Landy and Szalay (1993) estimator.

- Despite this abnormal result in ξ_{GG}, when we compared the cross and auto-correlation we did find a minimum of a 4σ difference in the degree of correlation at the central bin for both levels of binning.

Were we to compare our value for ξ_{AG} in the central bin with the values of ξ_{GG} from Hawkins et al. (2003) and Guzzo et al. (2008), this difference in the level of auto and cross correlation exceeded 5σ. Therefore at a redshift $z \lesssim 1$ there was a significant difference between the auto-correlation of galaxies and the cross-correlation between galaxies and low-column density Lyα absorbers.

The same conclusions were reached when we collapsed ξ_{AG} and ξ_{GG} completely along the projected or line of sight separation.

- When summed over the line of sight the cross-correlation $\Xi_{\mathrm{AG}}\left(\sigma\right)$ had a 3σ upper-limit of 1.26. The peak in the auto-correlation was at $\sigma = 0 - 0.5$ h_{70}^{-1}Mpc with $\Xi_{\mathrm{GG}}\left(\sigma\right) = 1.81 \pm 0.40$. However this histogram had the wrong characteristic shape because of our poor number of galaxy pairs. A more reliable value for the auto-correlation which showed that $\Xi_{\mathrm{GG}}\left(\sigma\right) > \Xi_{\mathrm{AG}}\left(\sigma\right)$ at this separation was derived from the results in Hawkins et al. (2003), $\Xi_{\mathrm{GG}}\left(\sigma\right) = 8.87 \pm 1.48$.

- The effect of the poor number of galaxy pairs along the projected separation was diminished when the number of pairs were summed over this axis. Now the auto-correlation at $\pi = 0 - 2$ h_{70}^{-1}Mpc was $\xi_{\mathrm{GG}}\left(\pi\right) = 2.18 \pm 0.31$. This was 4.5σ greater than the 3σ upper-limit of $\xi_{\mathrm{AG}}\left(\pi\right) = 0.78$.

  The significance of this difference increased to 5σ when we compared this 3σ upper-limit with the result of $\xi_{\mathrm{GG}}\left(\pi\right)$ from Hawkins et al. (2003).

Chapter 6

The Carbon and Oxygen Metallicity of the Lyα Absorbers

6.1 The Carbon IV and Oxygen VI Absorptions in QSO Spectra

Using the method that was described in Section 4.4 absorption lines of both C IV and O VI were identified in the STIS and UVES spectra of quasars HE 1122-1648 and PKS 1127-145.

The metal lines in the UVES spectra had already been identified in earlier line lists (Carswell et al., 2002b, Ryabinkov et al., 2003). We looked for metal absorption at the corresponding wavelength of the Lyα lines when searching STIS spectra, starting with those Lyα absorptions that had the highest column density. O VI lines proved the most difficult to identify because these lines were the most contaminated by the Lyman series forest. No O VI lines were found in the STIS spectrum of PKS 1127-145 because the maximum possible wavelength of any O VI line (this is set by the redshift of the quasar at $z = 1.187$) was lower than the minimum wavelength in this spectrum.

Table 6.1 contains a summary of the numbers of those lines found. The minimum metal column density detected in each spectrograph for both species have also been included to allow an upper-limit in the metallicity to be calculated.

In order to calculate the metallicity the first step was to pair each metal absorber with the Lyα absorber that was closest in redshift. To do this the recessional velocity

Table 6.1. The number of carbon IV and oxygen VI lines in the quasar spectra

Instrument	no. of C IV	no. of O VI	C IV Upper-Limit $\log_{10}(N_{HI}\,(cm^{-2}))$	O VI Upper-Limit $\log_{10}(N_{HI}\,(cm^{-2}))$
		PKS 1127-145		
STIS[a]	5	0	12.94	n/a
UVES	10	n/a	11.86	n/a
		HE 1122-1648		
STIS[a]	17	15	12.91	13.24
UVES	41	18	11.86	12.77

[a] In all of the STIS spectra none of the metal lines found had a variance weighted significance greater than the minimum value of 3, where we have used the same minimum significance as the Lyα absorption lines. This was the minimum used in Janknecht et al. (2006).

of all lines was found using equation (6.1) (Hogg, 1999).

$$v = c\frac{(z^2+2z)}{(z+1)^2+1} \tag{6.1}$$

This is only an approximate value for the velocities, as both the Lyα and metal absorbers will also have a peculiar velocity and random motion on top of the Hubble flow. An additional velocity difference of ±200 km s^{-1} was permitted when finding the nearest Lyα line. When using this margin of error between the recessional velocities all apart from 3 of the STIS C IV absorbers in the spectrum of HE 1122-1648 (which all had an EW significance $\sigma \leq 3$, and so assumed an upper-limit in the metallicity) could be paired.

6.1.1 The Ion Fraction of H I, C IV and O VI

To calculate the metallicity of each absorber the column density of both the metal and paired Lyα absorber, and the ratio of the metal and H I ion fraction were used in equation (2.1). The metallicities were expressed in solar units by subtracting the solar abundances of carbon and oxygen. These are $[C/H]_\odot = -3.61$ and $[O/H]_\odot = -3.34$ (Asplund et al., 2005).

The main problem is that the ion fractions vary by many orders of magnitude depending on the temperature and density of the absorption clouds. The temperature and density also need not be constant in a multi-phase IGM, as was discussed in Sections 1.3.3 and 2.3.1, where evidence was listed which suggests there is a multi-phase IGM. Therefore the error associated with the metallicity is far greater than when solely considering the error derived from the ratio of the column densities.

Values for the ionisation fraction of C IV, O VI and H I had been calculated at various redshift, densities and temperatures using the CLOUDY software (version 07.02) (Ferland et al., 1998). These simulations are the same ones used for the GIMIC simulation in Chapter 7. Additional details for the ionisation tables can be found in Wiersma et al. (2009). The photoionisation rate $\Gamma\,(z)$ and UV background when using CLOUDY were the same as those described in Section 1.3.2. The ionisation fraction had been sampled at intervals in redshift up to $z = 9$ in steps that did not exceed $\Delta z = 0.5$ (up to a redshift of 2.5 this did not exceed $\Delta z = 0.27$). The simulations were run with a variation in temperature that went from $T = 10^2 \rightarrow 10^{8.5}$ K and variations in density from $\log_{10}(n_H\ \mathrm{cm}^{-3}) = -8.0 \rightarrow -0.025$.

Contour plots of the ion fractions as a function of temperature and density of hydrogen, at redshifts of 0.5, 1, 1.5 and 2 for C IV, O VI and H I, are plotted in Figures 6.1, 6.2 and 6.3 respectively.

Despite the large range in the possible ion fraction of C IV, O VI and H I that had been calculated, we were interested in the possible values that would be expected in the diffuse IGM. This is where temperatures range from $T \sim 10^4 \rightarrow 10^6$ K with a minimum that is set by the UV background. The maximum temperature is that found in the WHIM.

The required range in the density of hydrogen corresponded to the typical density that caused the minimum and maximum Lyα column densities observed in the STIS and UVES spectra, specifically $11.7 \leq \log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \leq 17.4$. Metal lines are almost always identified in systems with $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \gtrsim 14.5$ (Schaye et al., 2003). However quasar spectroscopy has shown that both C IV and O VI are present in low density systems where the density contrast (defined in Section 1.3.4) can have values $\delta \leq 10$ (Aguirre et al., 2008).

Assuming that these absorbers are in hydrostatic equilibrium and can be modelled using the description in Section 1.3.4, then according to equation (1.18) the

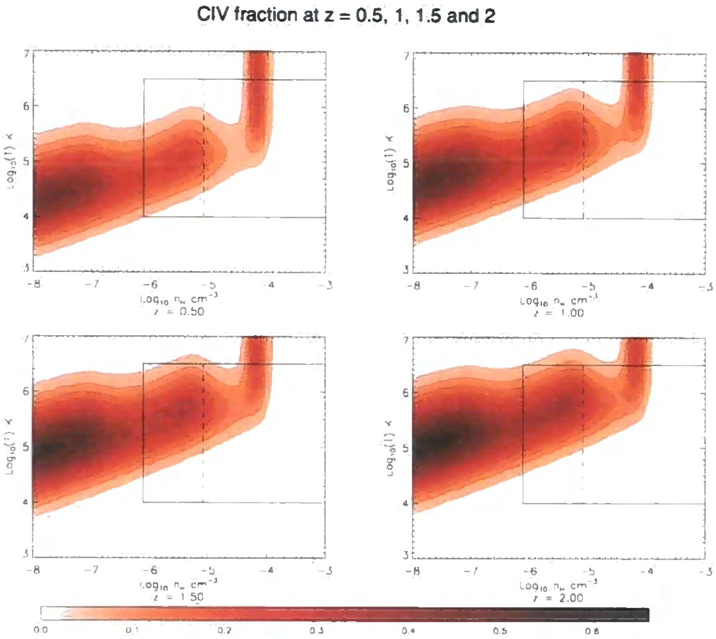

CIV fraction at z = 0.5, 1, 1.5 and 2

Figure 6.1 The ionisation fraction of C IV as a function of the n_H density and temperature and calculated using the CLOUDY program (Ferland et al., 1998). The values for the density and temperature were taken at 40 redshift intervals between z = 0 $\rightarrow$ 9. These plots show the ion fraction of C IV at redshifts z = 0.5, 1, 1.5 and 2. The tail at a temperature $T \gtrsim 10^5$ K and density $\log_{10}(n_H\,(\mathrm{cm}^{-3})) \sim 10^{-4}$ is where collisional ionisation has become significant. The enclosed rectangle marks the extreme temperatures and densities considered. The dashed line marks the density of a Lyα absorber of median column density ($\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 13.25$) assuming hydrostatic equilibrium.

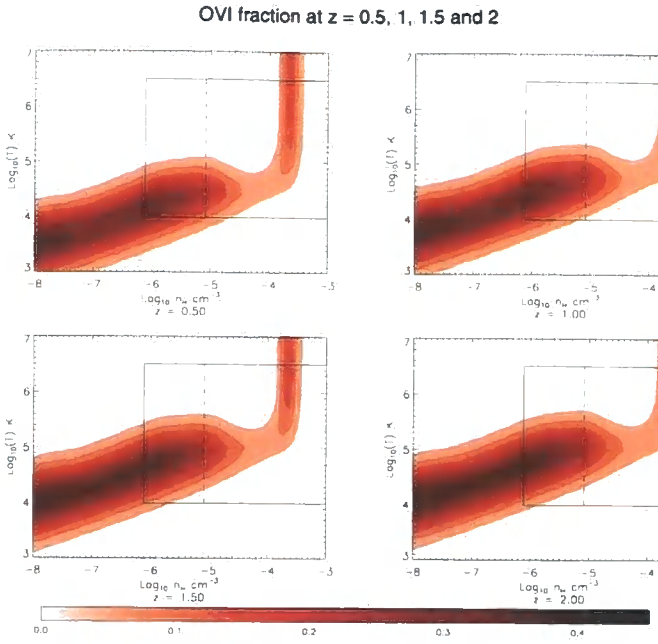


Figure 6.2 Calculated in the same manner as C IV in Figure 6.1 this shows the ion fraction of O VI at redshifts z = 0.5, 1, 1.5 and 2.

largest range in the possible density of hydrogen to use was $7.7\times10^{-7} \leq n_{H}\ (\mathrm{cm}^{-3}) \leq 4.8\times10^{-3}$. These limits in temperature and density are marked on the contour plots. This range in density represented the largest to consider because only 4 Lyα absorbers out of the total of 630 across both sightlines had a column density greater than $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq 16$ and only one had $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq 17$. A Ly$\alpha$ absorber of median column density that has $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 13.25$ has a density of $n_H = 8.3\times10^{-6}$ cm^{-3} at a redshift of 1; this is also marked on the contour plots.

6.2 Results

A table of all the significant C IV and O VI lines, together with the associated Lyα lines are listed in Appendix D.

6.2.1 The Carbon Metallicity

Of the 58 C IV absorbers found along the sightline towards HE 1122-1648 only 41 lines from the UVES spectrum had an equivalent width significance ≥ 3. Therefore the C IV absorbers identified in the STIS spectrum of HE 1122-1648 were used as

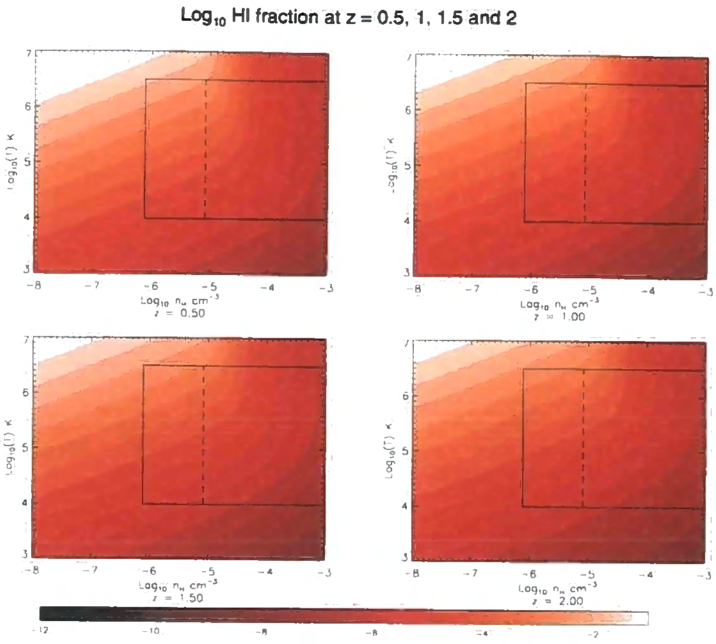


Figure 6.3 Log_{10} of the ion fraction of H I plotted as a function of redshift, temperature and density.

an estimate for an upper-limit in the metallicity only.

The number of data points for the C IV metallicity towards PKS 1127-145 was restricted because the 3 significant C IV lines that were paired were in the UVES spectrum. The Lyα emission line of PKS 1127-145 was at 2658.671 Å and so all of the Lyα lines were in the STIS spectrum. These three C IV lines were at redshifts of z = 1.058, 1.096 and 1.099.

An upper-limit in the carbon abundance was also found for all the Lyα lines from both sightlines that were not paired. In these cases the C IV column density took the value of the upper-limit in Table 6.1.

The metallicity calculations were split into 3 groups. First there were the significant metal lines found in the UVES spectrum with associated Lyα in either the UVES or STIS spectra. There were STIS C IV lines, these had an EW significance $\sigma \leq 3$ and so an upper-limit in the metallicity was calculated. Finally there were the Lyα lines in UVES and STIS that were not paired and assumed an upper-limit from Table 6.1. Figure 6.4 shows a scatter plot of the column densities from these 3 groups.

The maximum velocity separation between any paired C IV and H I absorber was

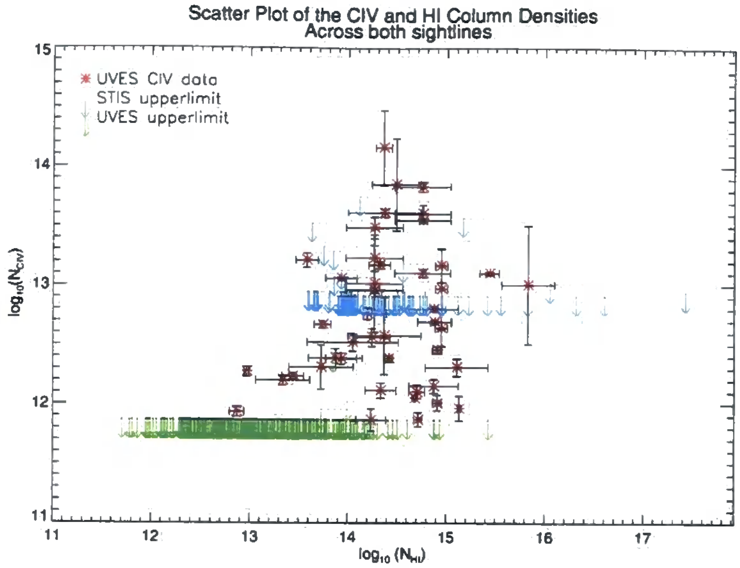


Figure 6.4 This scatter plot shows the column density of C IV compared to the matched H I column density. The 44 red data points are of the paired significant C IV lines found in the UVES spectra towards HE 1122-1648 and PKS 1127-145. The 15 paired marginal STIS C IV lines have been plotted as an upper-limit and are marked in blue. All of the unmatched STIS H I absorbers are also marked with a blue upper-limit. The 445 green upper-limits are for the unmatched H I lines from the UVES spectrum of HE 1122-1648.

$\Delta v = 121$ km s^{-1}. The minimum Lyα column density for an absorber from the STIS spectra that was paired with a significant C IV line was $\log_{10}(N_{HI}\,(cm^{-2})) = 13.58 \pm 0.11$. The minimum Ly$\alpha$ column density for an absorber from the UVES spectrum of HE 1122-1648 that was paired with a significant C IV line was $\log_{10}(N_{HI}\,(cm^{-2})) = 12.87 \pm 0.07$. The median column density was $\log_{10}(N_{HI}\,(cm^{-2})) = 14.37$. Figure 6.4 shows that C IV was common down to a H I column density of $\log_{10}(N_{HI}\,(cm^{-2})) \sim 14$. Only 10 Ly$\alpha$ absorbers with a column density lower than this had significant evidence for C IV absorption and all but one were from the UVES spectra. The minimum C IV column density observed was $\log_{10}(N_{CIV}\,(cm^{-2})) = 11.86 \pm 0.09$.

This demonstrates that in order to detect C IV absorbers with a small column density so that we can find the minimum metallicity of the IGM, a spectrum with a very high signal to noise ratio and low minimum detectable equivalent width is essential.

For the most accurate measurement of the metallicity we would require both the temperature and density of the H I and C IV absorbers, neither of which was known.

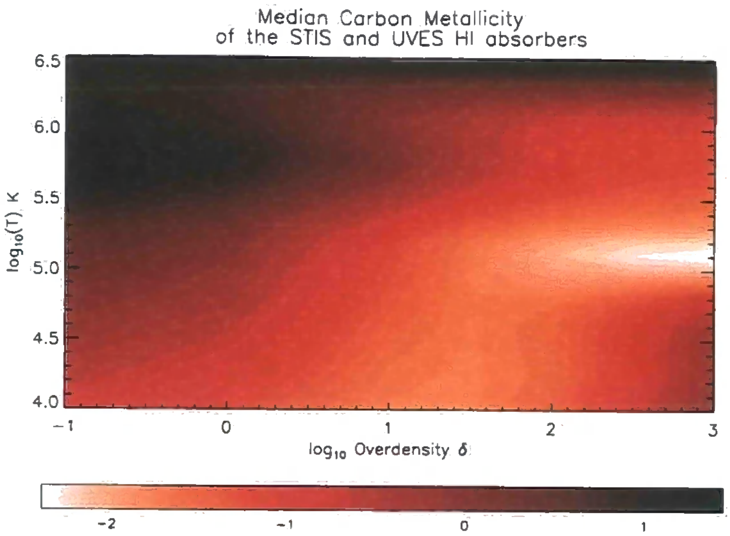


Figure 6.5 The median carbon metallicity is highly dependent on the temperature and density of hydrogen. This shows how the median carbon metallicity can vary from less than 10^{-2} to 10 times the solar value as the temperature varies from $T = 10^4 \rightarrow 10^{6.5}$ K and the overdensity δ ranges from $\delta = -0.9 \rightarrow 1000$.

The best we could do was place a limit on the temperature expected in the diffuse IGM, $T \sim 10^4 \rightarrow 10^6$ K. Constraints on the density could be derived by assuming that the Lyα absorbers remain optically thin ($\log_{10}(N_{HI}(cm^{-2})) \leq 17.3$).

The temperature was increased between $\log_{10}(T(K)) = 4 \rightarrow 6.5$ in steps of $\Delta\log_{10}(T) = 0.1$. The density is of course a function of redshift. Thus rather than compare values for the metallicity as a function of density, the metallicity was calculated as a function of the density contrast δ, defined in Section 1.3.4. This was done for values of $\log_{10}(\delta) = -1 \rightarrow 3$ in steps of $\Delta\log_{10}(\delta) = 0.5$. Figure 6.5 is a contour plot of the median [C/H] metallicity compared to the solar value, as a function of the temperature and density contrast δ within these limits. This shows that without a tighter constraint in the temperature and density the median carbon metallicity can vary by up to 4 orders in magnitude.

However according to Figure 6.5 as long as the temperature remained in the photoionised regime at $T \leq 10^{5.5}$ K and had a density of at least the mean cosmic density at that redshift then the maximum metallicity would not exceed the solar value.

We made the assumption that these absorbers are in hydrostatic equilibrium and are optically thin. This second assumption was reasonable since the maximum column density for those Lyα absorbers that had been paired with a significant C IV line was $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}}))= 15.83\pm0.27$. The values for the density could therefore be derived from the column density. To see how the Lyα column density varies with redshift, temperature and δ we refer the reader to equation (1.18) and Figure 1.2.

We also assumed a value for the temperature using a technique that was adopted in Danforth and Shull (2008). An upper-limit for the temperature can be obtained from the b-parameter of the C IV absorber. If we assume that the absorption line is dominated by Doppler broadening, and we neglect turbulence, then the temperature can be derived from equation (1.30) with the mass of the particle set to $12\mathrm{m_H}$. By doing this an upper-limit for the temperature was derived for the C IV absorber with the highest b-parameter of $\mathrm{T} = 1.1 \times 10^6$ K. The minimum b-parameter of 3.5 km $\mathrm{s^{-1}}$ gave an upper-limit of $\mathrm{T} = 8.9 \times 10^3$ K, just below that expected for the diffuse background temperature of the IGM. The median velocity of 11.2 km $\mathrm{s^{-1}}$ predicted an upper-limit in temperature of $\mathrm{T} = 9.1 \times 10^4$ K.

Using these values for the density and temperature, estimates were made for the carbon abundance. The numerical uncertainty in the metallicity mentioned below only considered the errors in the column densities of the Lyα and C IV absorbers.

The results of the [C/H] metallicity as a function of redshift are plotted in Figure 6.6. The mean [C/H] metallicity between redshifts $z = 1.0 \rightarrow 2.35$ has the relationship $[\mathrm{C/H}] = (-0.05\pm0.34)\,z - 1.24 \pm 0.58$. The median metallicity was $-1.27^{+0.52}_{-0.79}$.

Figure 6.6 shows that there was no change in the mean [C/H] metallicity with redshift between the redshifts of $z = 1 \rightarrow 2.35$. However the scatter was high ($\sigma^2 = 1.1$) with 2 H I absorbers at $z \sim 2$ with a carbon abundance that exceeded the solar value. There were also absorbers at a similar epoch with an abundance that was comparable with the floor carbon metallicity of $[\mathrm{C/H}] = -3.5$, predicted by Stocke et al. (2007) and mentioned in Section 2.3.3.

The histogram in Figure 6.7 compares the metallicity of those absorbers for which there was a C IV line with an upper-limit in the metallicity for those Lyα absorbers for which no significant metal line was present.

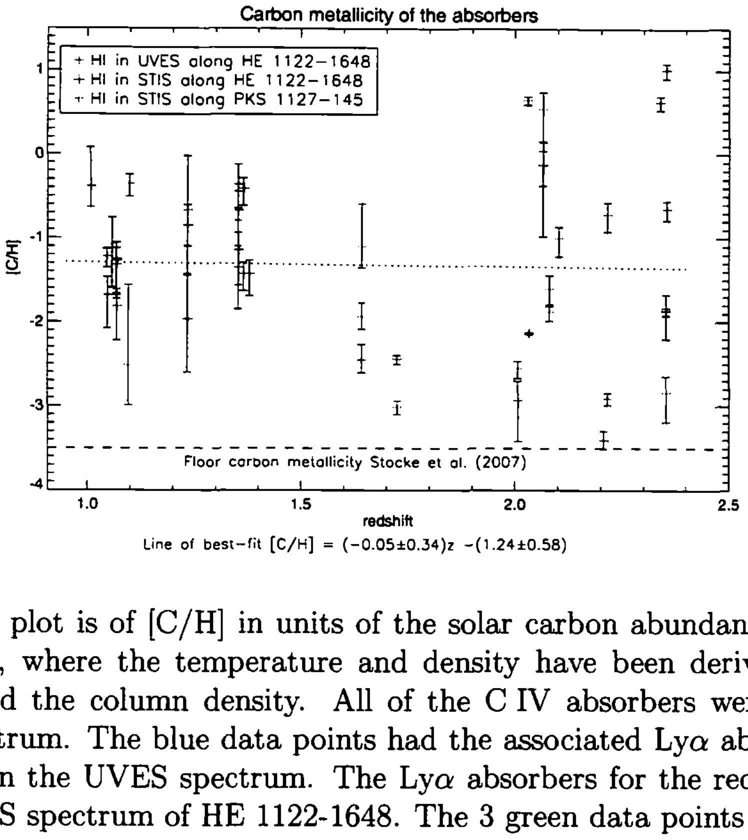


Figure 6.6 This plot is of [C/H] in units of the solar carbon abundance as a function of redshift, where the temperature and density have been derived from the b-parameter and the column density. All of the C IV absorbers were located in the UVES spectrum. The blue data points had the associated Lyα absorbers from HE 1122-1648 in the UVES spectrum. The Lyα absorbers for the red data points were in the STIS spectrum of HE 1122-1648. The 3 green data points are from the sightline towards PKS 1127-145. We find no evidence for the evolution of carbon abundance with redshift between z $\sim$ 1.0 $\rightarrow$ 2.35. The floor [C/H] metallicity at [C/H]$= -3.5$ predicted by Stocke et al. (2007) is shown.

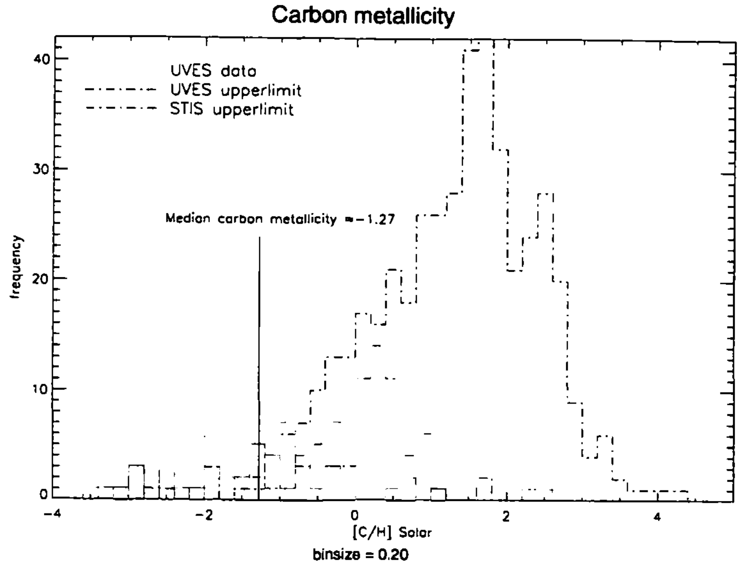


Figure 6.7 This histogram binned Δ[C/H]= 0.2 compares the metallicity of those absorbers with a carbon abundance with an upper-limit in the non-detections in the STIS and UVES spectra. Most upper-limits have been over-estimated because the minimum C IV column density was too high when compared to the weak Lyα absorber column densities.

These upper-limits in the carbon abundance are very high even when the errors associated with the ion fraction are considered. The excessive upper-limits are a consequence of dividing the minimum observed C IV column density by Lyα column densities that are too small. Clearly such a high contrast between the minimum observed metal and H I column densities was not appropriate.

What Figure 6.7 and this exercise help to illustrate are the questions posed earlier in Section 2.3.3: what is the minimum H I column density at which it is possible to observe a metal line, and what is the minimum column density of the metal absorbers?

When trying to find a lower constraint for the upper-limit in [C/H] either the appropriate C IV equivalent width was too low, and could not be detected using either STIS or UVES (the spectra for each instrument had a minimum detectable equivalent width of 0.13 and 0.011 Å respectively). Or we had passed the minimum C IV column density that exists in the IGM and these regions remain uncontaminated.

The best we could do to reduce the upper-limits was to remove the weak Lyα absorbers from the calculation. More data is required before we can impose a universal minimum H I column density below which metals are no longer seen. Until then we can only assume that the C IV absorbers that could be associated with these Lyα lines remain undetected.

When we increased the minimum required Lyα column density for those absorbers used in the upper-limits to the minimum column density for which an associated metal line had been observed (depending on the spectrograph), then the median upper-limit for the UVES data decreased from [C/H]= 1.44 to 0.75. The STIS data had a decrease in the median upper-limit of [C/H]= 0.29 to 0.13.

Using the UVES spectra we found an upper-limit on the minimum Lyα column density for which C IV could be observed of $\log_{10}(N_{HI}\,(cm^{-2}))= 12.87 \pm 0.07$ and a minimum C IV column density of $\log_{10}(N_{CIV}\,(cm^{-2})) = 11.86 \pm 0.41$. The equivalent width of this limiting C IV line was 0.018 ± 0.002 Å.

6.2.2 The Oxygen Metallicity

The [O/H] metallicity of HE 1122-1648 was investigated in the same way as for carbon. This time, because the significant O VI absorbers were all located in the UVES spectrum, the associated H I lines were located here as well. That meant the redshift range of the 18 H I-O VI pairs that were found was only $z = 2.0 \rightarrow 2.4$. The maximum separation between a Lyα and O VI absorber was 50.5 $km\,s^{-1}$. The minimum Lyα column density with an associated O VI absorber was $\log_{10}(N_{HI}\,(cm^{-2}))= 12.87 \pm 0.07$ and the minimum significant O VI column density, $\log_{10}(O_{VI}\,(cm^{-2})) = 12.77 \pm 0.13$.

When estimating the upper-limits in temperature we used equation (1.30); however because a greater fraction of O VI than C IV ions are believed to trace the WHIM, the mean temperature involved was higher. The maximum b-parameter of 60.1 $km\,s^{-1}$ implied a temperature of $T = 3.5 \times 10^6$ K. The median b-parameter was 16.7 $km\,s^{-1}$ and this gave a temperature of 2.7×10^5 K. While the smallest b-parameter of 6.2 $km\,s^{-1}$ has a temperature of 3.7×10^4 K. This difference in the abundance of oxygen at high temperature when compared to carbon can be seen in Figure 6.8 where we have plotted the median [O/H] as a function of temperature and δ. So long as the temperature of the absorber remained above $T = 10^{5.5}$ K if the

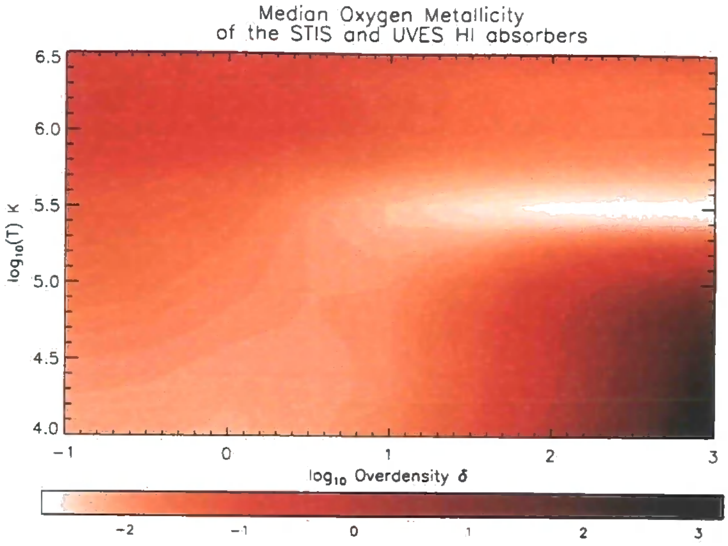


Figure 6.8 The median [O/H] metallicity can be compared with that of [C/H] that is shown in Figure 6.5. The increase in the O VI ion fraction at high temperatures, the property that makes O VI a tracer of WHIM gas, is the cause of these difference in the the two plots. At high density and T$\leq 10^5$ K the O VI ion fraction quickly falls.

density contrast exceeded 100, then the median oxygen metallicity did not exceed the solar value.

Figure 6.9 shows the evolution of oxygen abundance using the derived temperatures and densities. We see no evidence for a change in [O/H] with redshift between z = 2.0 → 2.4 and like the carbon result there is a large scatter ($\sigma^2 = 0.77$).

A flat evolution with the mean value [O/H] = $(1.01 \pm 1.54)\,z - (3.94 \pm 3.35)$ was found. The median oxygen metallicity was $-1.68^{+0.46}_{-1.01}$.

Most of the upper-limits were once again overestimated because the unmatched Lyα column densities were too low. By limiting the minimum column density to that of the lowest Lyα absorber for which a significant O VI line was detected ($\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))= 12.87 \pm 0.07$), the median UVES upper-limit decreased from -0.05 to -0.76.

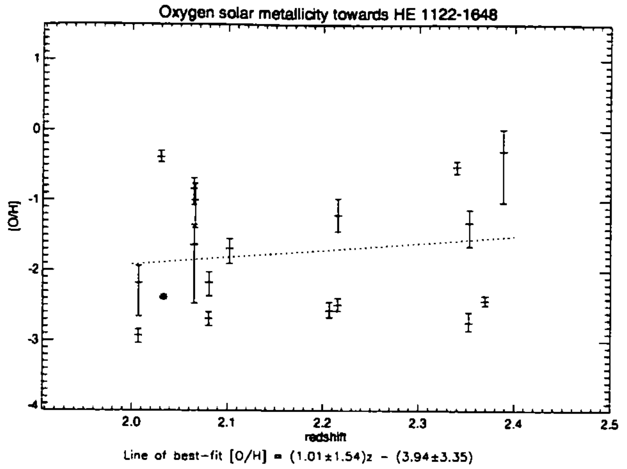


Figure 6.9 Between the redshifts of z = 2.0 → 2.4 towards HE 1122-1648 we detect no evolution in [O/H] but there is a large scatter with a variance of $\sigma^2 = 0.77$.

6.3 Discussion and Conclusions

Along both sightlines and for both species we saw no evidence for changes in the metallicity with redshift, with a mean carbon metallicity of $[C/H] = (-0.05 \pm 0.34)\,z - 1.24 \pm 0.58$ between redshifts $1 \leq z \leq 2.35$ and a mean oxygen metallicity of $[O/H] = (1.01 \pm 1.54)\,z - 3.94 \pm 3.35$ between $2 \leq z \leq 2.4$.

In an expanding universe this abundance is kept constant by the competing effects of changes in the rate of enrichment, which is set by the frequency of starbursts and feedback, changes in the density of the IGM, and the ionising background. However with so many different environments under which these metals are detected, their proximity to galaxies and critical changes in the ion fraction depending on the temperature and density, it is not surprising that we see a large amount of scatter.

Since all these processes are dependent on one another further attempts to explain these results would be speculation. It is also naive to rush to a conclusion that there is little evolution in the carbon and oxygen metallicity based on only two sightlines, and over short redshift path lengths. However more extensive investigations from the literature (Aguirre et al., 2008, Scannapieco et al., 2006, Schaye et al., 2003) have also found little variation in the carbon and oxygen abundance in the IGM, and have detected large amounts of scatter. Their results, together with ours will be discussed below.

6.3.1 Carbon

Scannapieco et al. (2006) traced 619 C IV systems in high resolution echelle spectra between the redshifts $1.5 \leq z \leq 3.1$ and found no dependence on redshift.

In a series of papers that studied the metallicity of 19 UVES and HIRES quasar spectra including HE 1122-1648, Schaye et al. (2003) also found a flat evolution in [C/H] between redshifts $1.8 \leq z \leq 4.1$. It was the density contrast that was found most strongly to affect the carbon abundance the most, with a median of $[C/H] = -3.47^{+0.07}_{-0.06} + 0.08^{+0.09}_{-0.1} (z - 3) + 0.65^{+0.1}_{-0.14} (\log\delta - 0.5)$.

Schaye et al. (2003) found dependence on the temperature to be weak as long as it remained within the range $10^4 \leq T \leq 10^{4.5}$ K, with densities low enough that the ion fraction was dominated by photoionisation only ($n_H \sim 10^{-5} \rightarrow 10^{-4}$ cm^{-3}). This can be compared with the result in Figure 6.5, where the C IV ion fraction remained high and hence the median carbon metallicity remained below the solar value as long as the temperature and density remained within these limits. Their results were backed up with simulated spectra that assumed temperatures of $T \sim 10^4$ K. However Schaye et al. (2003) also state that carbon may exist at temperatures $T \ll 10^4$ and $T \gg 10^5$ K, although this remains undetected as it does not produce significant absorption.

The measured mean carbon abundance from Schaye et al. (2003) was $[C/H] = -2.80 \pm 0.13$ which is less than 3σ lower than our median value of $-1.27^{+0.52}_{-0.79}$.

With this evidence supporting claims that there has been little change in the abundance of carbon, we can conclude that periods of contamination have occurred since a very high redshift. This lends support to the idea proposed by Stocke et al. (2007) that shortly after reionisation there was a sustained period of contamination of the universe that produced a lot of the metals we now observe. This supplied the floor in the metallicity that is marked in Figure 6.6. It is believed that this level of carbon abundance should be observed in the diffuse regions of the IGM and even in more rarefied regions. The data point in Figure 6.6 with a metallicity close to this universal minimum $[C\ IV/H] = -3.41^{+0.12}_{-0.09}$ was this low because of a relatively high Lyα column density of $\log_{10} (N_{HI} (\mathrm{cm}^{-2})) = 15.83 \pm 0.27$, and not because we were detecting the minimum C IV equivalent width possible.

The most interesting case is observing metal absorptions that are associated with Lyα absorbers that have the lowest measured column density. Schaye et al. (2003) found evidence at a 2.4σ level for carbon contamination in these under-dense regions where $-0.5 \leq \delta \leq 0$. More conclusive proof is therefore required, together with the positions of the surrounding galaxies, before we can say that carbon can be found in the low density gas in the voids of the IGM.

6.3.2 Oxygen

When measured over the short redshift range $2 \leq z \leq 2.4$ we found no change in the mean oxygen metallicity, but we did see a great deal of scatter in Figure 6.9. This is for the same reasons as the carbon abundance; rather than depending on the redshift, the metallicity depended on the temperature and the density.

Figure 6.8 shows that the estimation of the median oxygen metallicity changed considerably through many orders of magnitude as the temperature rose from $10^4 \leq T \leq 10^{6.5}$ K. This sudden change at high density can also be seen in Figure 6.2. At densities of $\delta \geq 100$ and below $T = 10^5$ K the O VI ion fraction drops by a significant amount which is why the median metallicity is many times the solar value. However when the temperature rises above 10^5 K, collisional ionisation becomes significant and so the abundance of observed oxygen increases.

Thus if we are to avoid metallicities that are many times the solar value (no data point in Figure 6.9 had a value $[\mathrm{O/H}] \geq 0$), according to Figure 6.8, any H I absorber with a density contrast $\delta \gtrsim 100$ must also have had a temperature $T \gtrsim 10^5$ K. At a redshift of $z = 2$ and assuming hydrostatic equilibrium this would correspond to a column density of $\log_{10}(N_{\mathrm{HI}}\,(\mathrm{cm}^{-2})) = 15.9$.

To estimate the expected oxygen metallicity of the IGM using the same 19 QSO spectra used in Schaye et al. (2003), Aguirre et al. (2008) first found the value of [O/C] using the carbon abundance that had been measured in Schaye et al. (2003).

Aguirre et al. (2008) then calculated the median oxygen metallicity across all the spectra for values of $\log\delta = -0.5 \rightarrow 2$. It was found that values for the median metallicity that had $\log_{10}(\delta) \leq 1.5$ were close to the expected value predicted by the comparison with carbon. However the prediction was an over-estimate for the median metallicity calculations that had $\log_{10}\delta \geq 1.5$. Aguirre et al. (2008) attribute this error to the collisional ionisation that occurs at these densities. Hence Aguirre et al. (2008), like in our investigations, also found a critical change in the ionisation fraction when $\delta \gtrsim 10$.

This is why Aguirre et al. (2008) split their data and worked out the metallicity of those absorbers with $\delta \leq 10$ separately from those with $\delta > 10$. The low density absorbers were found to have no evolution in metallicity with redshift over the range $1.9 \leq z \leq 3.6$, but were found to have a strong δ dependence. Those absorbers with $\delta \geq 10$ were found to have a metallicity that was less dependent on the density. Instead there was a large ion correction that was caused by collisional ionisation.

A final complication arises for both ion species because the upper-limit in the temperature that was derived from the b-parameter of the metal absorber often bore little relation to the upper-limit in the temperature of the associated Lyα absorber. For example the upper-limit in the temperature of the highest column density Lyα absorber was 5.3×10^4 K. The upper-limit in the temperature of the associated O VI absorber was 2.4×10^5 K. Therefore rather than an absorber where the temperature is homogeneous, the situation is of a multi-phase system where the metal and H I in the absorber are not in a state of equilibrium.

All of these uncertainties mean that this exercise, (rather than being a study where conservative constraints can be placed on the extent of metals in the IGM), has shown the problems that are faced in this subject and the assumptions that can be made to overcome them.

Chapter 7

Simulating the 2D 2-Point Correlation Function

7.1 The Galaxies-Intergalactic Medium Interaction Calculation

The use of high performance computing in astrophysics has allowed increasingly accurate simulations of the universe to be made. One of the largest that contained dark matter only in recent times was the Millennium Simulation. This was designed and run by the Virgo consortium (Springel et al., 2005). This simulation traced the evolution of 2160^3 particles in a cube of co-moving length 500 h^{-1}Mpc from a redshift of $z = 127$ to the present.

This was done using the parallel N-body code GADGET-2 (Springel, 2005b). Only dark matter was considered therefore the particles, that each had a mass $8.6 \times 10^8\ h^{-1}M_\odot$, behaved as part of a collision-less fluid that acted under gravitational forces only. By using this simulation Springel et al. (2005) were able to study the merger history of galaxy haloes and the creation of large scale structure, including the dark matter clusters and filaments that are thought to make up the cosmic web.

On large scales studying the creation of structure using dark matter only is sufficient, because it is only gravity that is known to have effects that span these distances. A problem arises if we want to study the universe on much smaller scales, including investigations such as the Lyα-galaxy cross-correlation. In order to do this we need to include baryons and gas physics. With so many further complications the computational requirements that would be needed to simulate the hydrodynamics on the scale of the Millennium Simulation would not be feasible.

With this in mind 5 regions that make up 0.13% of the Millenium Simulation were re-simulated using both dark matter and baryonic particles. This was the Galaxy-Intergalactic Medium Interaction Calculation (GIMIC) that is summarised in greater detail in Crain et al. (2009) (hereafter C09).

7.1.1 The Five GIMIC Regions

The five regions that were chosen for re-simulation were selected so that a wide dynamic range in the conditions and densities that exist in the Millennium Simulation could be studied in detail. A snapshot of the Millennium simulation at a redshift of $z = 1.5$ was analysed and spheres with a radius $\gtrsim$ 18 h^{-1}Mpc were found that were near to the mean density of the Millennium Simulation (1.7×10^{15} $h^{-1}M_{\odot}$ (C09)). Other spheres were then identified that were $-2, -1, +1$ and $+2\sigma$ of this mean value.

It was the wish of the Virgo group that the $+2\sigma$ region contain a dense cluster in the centre at $z = 0$. At a redshift of $z = 1.5$ there were no 18 h^{-1}Mpc spheres that contained a mass that was 2σ greater than the mean value. Thus the radius of the $+2\sigma$ region was extended to 25 h^{-1}Mpc and a short list of 12 possible spheres were produced (C09).

Two more criteria were specified when selecting the 5 regions. To avoid edge effects the outer periphery of all spheres had to be $\geq$ 50 h^{-1}Mpc from the edge of the Millennium volume. The regions also needed to be independent, so the centre of each sphere was $\geq$ 200 h^{-1}Mpc from another. A suitable sphere for the $+2\sigma$ was chosen, and random spheres for the other four regions were selected from the list of suitable locations. The five -2, -1, 0, +1 and $+2\sigma$ spheres were named regions 1, 2, 3, 4 and 5 respectively, and will be referred to by these names during the rest of this chapter.

Two realisations of the GIMIC simulation were carried out, these were identified as having 'high' and 'intermediate' resolution. The low resolution simulation was the Millennium run.

The restrictions on the computational power available meant that region 5 was not run at 'high' resolution, and regions 2, 3 and 4 were stopped at $z = 2$. We only

Table 7.1. Parameters of the 5 GIMIC regions at 'intermediate' resolution, data is from Crain et al. (in preparation)

Sphere	Co-moving radius h^{-1}Mpc	M_{gas} $\times 10^7$ $h^{-1}M_\odot$	$M_{dark\ matter}$ $\times 10^7$ $h^{-1}M_\odot$	No. of particles $\times 10^7$	Summation weight
1: -2σ	18	1.16	5.28	2.23	1/12
2: -1σ	18	1.16	5.28	2.80	1/6
3: 0σ	18	1.16	5.28	3.44	1/2
4: $+1\sigma$	18	1.16	5.28	4.30	1/6
5: $+2\sigma$	25	1.16	5.28	12.4	1/12

used data from the 'intermediate' resolution simulation because of this restriction in the regions simulated and the minimum redshift. The 'intermediate' resolution gas particle mass was still an improvement on those used in Wilman et al. (2007) and Pierleoni et al. (2008) at 2.72×10^8 and 4.3×10^7 $h^{-1}M_\odot$. Table 7.1 gives the parameters and details for each region at 'intermediate' resolution. The dark matter particles had a mass that was a factor $(\Omega_m\text{-}\Omega_b)/\Omega_b$ greater than those that represented baryons.

It was our aim to investigate the cross-correlation function in each region separately to see what effect the difference in the mean density would have. The result from each was then combined by a weighted mean. The appropriate weighting for regions 1-5 are also listed in Table 7.1 and are from C09.

7.1.2 The Cosmology

For reasons of comparison, up till now the cosmology we had adopted was chosen to be the same as that from Wilman et al. (2007). The cosmology that was used in the GIMIC simulation was the same as that used during the Millennium run. This had $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_b = 0.045$ and $H_0 = 100$ h km $s^{-1}Mpc^{-1}$ and h = 0.73. A higher value for Ω_Λ and lower value of Ω_m will mean the function E(z) in equation (1.7) was smaller, and this implies a larger universe. This lead to a larger co-moving distance to both the galaxies and the absorbers (equation (1.9)) and a larger separation between pairs. However this was off-set by the slight increase in the h parameter. Thus the relative separation between pairs increased, but with the coarse binning along the line of sight and projected separation on a scale of Mpc, we do not expect the distribution of ξ_{AG} to have changed significantly.

7.1.3 Details of the Gas Physics

The code used for the GIMIC simulation was GADGET-3. On top of the code that described the gravitational behaviour of dark matter and baryonic particles, additional modules had been added that dealt with the gas physics. These included; radiative cooling and heating by the background UV radiation (Wiersma et al., 2009), star formation (Schaye and Dalla Vecchia, 2008), galactic winds (Dalla Vecchia and Schaye, 2008) and the chemical enrichment of the IGM (C09). The steps that were taken to simulate these processes are summarised below. In many cases a semi-analytic solution was used. This is where the parameters in certain models were constrained and scaling factors included, so that the simulation would match the observations.

Radiative cooling and photo-ionisation

The gas was heated by the redshift dependent UV background. This was based on the model that was made by Haardt and Madau (2001) and was described in Section 1.3.2.

The ionisation front of the impacting UV photons was not simulated. Instead the background was assumed to be uniform and pervade the entire volume. This assumption caused a problem because the pockets of concentrated H I could not be accurately simulated. In large clouds of optically thick neutral gas self-shielding occurs. What this implied for the damped and sub-damped Lyα absorbers that have a column density $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \geq 17.3$ will be described in Section 7.1.4 below.

In order to implement star formation when sufficient gas particles had come together GADGET-3 had to simulate radiative cooling. This dominates in regions of high temperature and density, as opposed to cooling in the diffuse IGM which is mainly achieved by adiabatic expansion.

In the model radiative cooling came into effect after re-ionisation, which in the simulation occurred for hydrogen at $z = 9$ and at $z = 3.5$ for helium. The cooling rate was then dependent on redshift (because of the redshift dependent photo-ionisation), the density, temperature and the amount of metals present because of cooling by line emission. The density, temperature and ion fraction of the particles were determined using the CLOUDY ionisation tables (Ferland et al., 1998), that were described in Section 6.1.1.

The model assumed that the gas was optically thin and in a state of collisional ionisation equilibrium. Again these assumptions can be called into question, particularly when the density of the gas increased above that expected for optically thin media in sub-damped Lyα absorbers.

Star formation

A code that accurately predicted the cooling rate was needed in order that gas particles could collect, become gravitationally bound, and form star particles when fusion had assumed to commence. Too much heating in the inter-stellar medium (ISM) would cause the gas particles to disperse.

It was not possible to recreate the many phase changes that occur during the cooling and collapsing of gas clouds from which stars are made. Instead, when the gas particles had a density that exceeded $n_H = 0.1\ cm^{-3}$ they were assumed to be gravitationally unstable and were classed as star particles. Each star particle continued to have a mass of $1.16 \times 10^7\ h^{-1}M_\odot$, therefore each particle, having a mass that is more identifiable as a cluster of stars, represented an entire stellar population.

When they became stars these particles adopted a new equation of state. The parameters in this equation were fixed so that the rate of star formation would be the same that was observed in Kennicutt (1998).

Galactic winds and chemical feedback

The lifetime and evolutionary track of a star particle depended on the metallicity of the former gas particle and the stellar initial-mass-function (Chabrier, 2003). Any star particle representing a population of stars that each had a mass $6 \leq M_\odot \leq 100$ then underwent a supernova explosion and distributed heated gas and metals into the ISM.

These gas particles then remained in the ISM for a series of time steps that equated to 3×10^7 years, during this time the gas became enriched with metals. The ejected gas particle was then given a burst of kinetic energy in a random direction with a velocity of 600 km s^{-1}, regardless of the galaxy halo mass or environment. To simulate the mass loading the ejected particle was coupled with 4 adjacent gas particle neighbours. It was assumed that 80% of the heat ejected by the supernova was removed from the ISM by this escaping wind. The combined ejecta from all the supernovae created a galaxy wide outflow.

A more accurate treatment would have considered dust and metals being swept up by the escaping wind and added to the load, particularly because the escaping gas loses momentum as it escapes the galaxy potential. The load would then be deposited in the IGM or would fall back onto the galactic disk. However constraints on the usual size for a mass load are unknown, as is the efficiency of the wind to remove energy and material from the ISM. The initial outflow velocity also varies because of the deceleration caused by mass loading, the size of the supernova and the mass of the galaxy. With so many unknown parameters the load, wind speed and wind efficiency were fixed. This way contaminants of C, N, O, Ne, Mg, Si, S, Ca, and Fe were added to the IGM.

7.1.4 Limitations with the Simulation

Apart from the semi-analytic equations and fixed parameters that were added in order that the results from the simulation would recreate observations (such as the galaxy luminosity), there were several drawbacks with the simulation that could have a direct effect on our result and what we were able to reliably investigate.

Edge effects

During the re-simulation all particles that were outside the GIMIC regions were classed as composite collision-less dark matter/baryonic particles. The particles in what were dubbed 'high resolution' cells were within the GIMIC regions and could either be dark matter or baryons. Thus no gas existed beyond the GIMIC regions.

This created an edge effect as the absence of gas beyond the $\sim$ 18 h^{-1}Mpc boundary meant there was no pressure and no gas could flow into the sphere, however gas particles could leave. This effected the accretion rate and density of gas around the edge. For this reason no data was used outside what was deemed the safe 18 h^{-1}Mpc radius of regions 1-4, and beyond 25 h^{-1}Mpc of the centre of region 5.

No self-shielding and optically thick absorbers

When simulating radiative processes it was assumed that there was no self-shielding and all absorbers were optically thin. All of the gas was also in a state of collisionally and photo-ionised equilibrium. These assumptions are reasonable for the diffuse gas in the IGM that constitutes the low column density Lyα absorbers that have $\log_{10}(N_{HI}(\mathrm{cm}^{-2})) \leq 17.3$. It is those absorbers that had $\log_{10}(N_{HI}(\mathrm{cm}^{-2})) \geq 17.3$ that were modelled inaccurately, and the measured value for the column density and b-parameter were only approximate.

The line-profiles for these high column density absorbers from the model were also wrong. Figure 1.5 shows the Lorentzian shape that is expected when the column density reaches these levels. SPECWIZARD, the software that generated the artificial quasar spectra, did not simulate Lorentzian profiles, hence no absorber had damping wings. Instead any absorber that was recorded to have $\log_{10}(N_{HI}(\mathrm{cm}^{-2})) \geq 17.3$ would have a saturated Gaussian profile with a large FWHM. To compensate for these setbacks, when we binned our results by column density, any absorber that had a column density of $\log_{10}(N_{HI}(\mathrm{cm}^{-2})) \geq 17$ was placed in a separate bin. There were no damped Lyα absorbers in any of the line-lists, with a maximum column density recorded in region 5 of $\log_{10}(N_{HI}(\mathrm{cm}^{-2})) = 18.89 \pm 0.07$.

The $+2\sigma$ Region

The cluster at the centre of the $+2\sigma$ region offers the opportunity to study high density systems in a lot of detail. However, no such cluster existed along our lines of sight towards HE 1122-1648 or PKS 1127-145 at the redshifts studied. So that our observations could be more realistically simulated the cross-correlation function was also measured with the $+2\sigma$ region removed.

7.2 Creating a Lyα Absorber Linelist

Our plan when using GIMIC to investigate the Lyα absorber-galaxy cross-correlation function was to fire sightlines through all 5 of the GIMIC regions. Then using the exact same techniques and fitting parameters that were employed in Chapters 4 and 5, fit Voigt profiles to the generated Lyα absorbers, and correlate these with the galaxies in each region that were of a similar magnitude limited sample as our real data.

In order to see whether or not ξ_{AG} has a redshift dependency, and to compare to our observations, the GIMIC snapshots that were used had redshifts z = 1 and 0.5.

7.2.1 Generating the Spectra

The spectra were created using the program SPECWIZARD (version 080121) [1], this was written to extract sightline data specifically from GIMIC snapshots.

The SPECWIZARD parameters

First of all we had decide on the value for the parameters that would be used when running SPECWIZARD. Those parameters not mentioned below, such as the metallicity in the IGM of certain metal species, or the scaling of the Haardt and Madau (2001) UV background, were set to their default values.

- *Long or short spectra?* The GIMIC regions only had a diameter of $\sim$ 36 h^{-1}Mpc. At a redshift of z = 1, and using the GIMIC cosmology this equates to a velocity difference of only 2049 km s^{-1} ($\sim$ 26 Å). Therefore the simulated spectra were very short. The alternative was to have composite long spectra, where several short spectra from consecutive snapshots have been patched together. Analysing long spectra did not appear to have any advantage, and the short spectra were also better suited to our needs because we were only

---

[1] SPECWIZARD was written by J. Schaye, C. Booth, Leiden Observatory, PO Box 9513, 2300 RA Leiden, The Netherlands and T. Theuns, Institute for Computational Cosmology, Dept. of Physics, University of Durham, UK

going to study the correlation function at redshifts z = 1 and 0.5, and VPFIT would work much more quickly on a smaller number of pixels.

- *The ions to include.* SPECWIZARD calculated the optical depth caused by each ion species individually. Hence it was possible to only include metals when creating spectra. To avoid contamination by metals and absorbers in the Lyman Series above Lyα, only absorptions caused by the Lyα transition were selected to appear in the spectra.

- *The number of sightlines.* In order to substantially decrease the size of the error in the cross-correlation function when compared to our observed measurement, a lot of sightlines needed to be used. We fired 3 arrays of 15×15 evenly spaced sightlines through the centre of each GIMIC region along the 3 Cartesian axes. Hence 675 sightlines were used in each GIMIC snapshot.

- *The location and orientation of the sightlines.* All 675 sightlines were all contained within a cube of length 20 h^{-1}Mpc, this shared a common centre with the 36 h^{-1}Mpc safe region of the GIMIC sphere in order to avoid edge effects. The spacing between each sightline along a consecutive row or column was 1.43 h^{-1}Mpc. This way the outermost sightlines ran along the vertices of the cube.

  When choosing the trajectory of a sightline we specified the central [x,y,z] Cartesian coordinate through which the sightline would pass, and the spherical coordinates that controlled the axial and zenith angles $[\phi,\theta]$. All sightlines along the z-axis had $[\phi,\theta]=[0,0]$, along the y-axis, $[\phi,\theta]=[\pi/2,\pi/2]$ and along the x-axis, $[\phi,\theta]=[0,\pi/2]$.

A criticism with sampling the small GIMIC regions with this many sightlines was that we may have over-sampled the regions. An example of this problem would be a prominent feature, such as a filament, that existed along a particular axis. If too many sightlines passed along or through this then the number of close pairs would be unrealistically high.

Steps were taken to minimise this problem. First of all, if a prominent feature had existed in a region along a particular axis, then the effect was diminished by firing sightlines along all three axes.

Secondly all of the sightlines were 1.43 h^{-1}Mpc from each adjacent line. The level of binning used along the projected separation mimicked that from the observations, $\Delta\pi = 2$, $\Delta\sigma = 1$ h_{73}^{-1}Mpc. Thus no two absorbers that were in adjacent sightlines, that were paired with the same galaxy, were contained in the same bin.

A third precaution was selection of the redshift step. By sampling the same region across snapshots that were close in redshift there was the possibility that we would sample the same absorber-galaxy pair twice. The time difference between $z = 1 \rightarrow 0.5$ using equation (1.10) is 8.5×10^{16} s. If we assume a galaxy pair separation velocity of 412 km s^{-1} (Guzzo et al., 2008), then we can expect a galaxy within the GIMIC regions to have moved $\gtrsim 1.1$ h^{-1}Mpc. Hence it is possible that the same absorber-galaxy pair could have been sampled twice, particularly if the gas or galaxy remain reasonably static over this decrease in redshift. It is unlikely however that these pairs would occupy the same corresponding bin in the plots that show ξ_{AG} at $z = 0.5$ and 1.

A second criticism is one which concerns all simulations. How can such a small volume be assumed to be representative of the entire universe? Sampling the regions with many sightlines can also mean that structures or traits that are common in a particular realisation, but not necessarily the universe, will appear more significant than they should be.

Convolving, re-binning and adding noise to the spectra

For each sightline SPECWIZARD produced an array of the normalised optical depth as a function of the Hubble velocity from the centre of the GIMIC sphere. A program was written that transformed these arrays so that they mimicked the real STIS data.

First an absorption spectrum with an infinite signal-to-noise (S/N) was made by using the optical depth τ, in flux $= e^{-\tau}$.

We then convolved this spectrum with the Gaussian point-spread-function (PSF) of the STIS E230M instrument. The value of σ for this Gaussian was derived using equation (7.1), and the STIS FWHM of 12 km s^{-1}.

$$\sigma_{\mathrm{Gauss}} = \frac{\mathrm{STIS\ FWHM}}{\sqrt{2\ln 2}} \tag{7.1}$$

The spectrum and velocity arrays were then rebinned so that the velocity step for each pixel was $\Delta v = 4.8$ km s^{-1}, this was the same as the mean velocity step per pixel for the STIS E230M spectra.

A wavelength scale was then established using the rebinned velocity array, v, in equation (7.2). There was no evolution in redshift over a GIMIC snapshot.

$$\lambda = \lambda_0 (1 + z)\left(1 + \frac{v}{c}\right) \tag{7.2}$$

The final stage in creating a spectrum was to add random noise that had a Poisson distribution, with a mean value in the S/N that was derived from the real data.

In order to do this we needed to know the S/N of the STIS spectra as a function of wavelength and flux. We linearly interpolated over all the absorption and emission lines that existed in the STIS E230M spectra, and over the equivalent pixels that described the variance. This way any drop in the variance spectra that had been caused by a decrease in the flux because of Poisson statistics was removed. We then smoothed the remaining flux and variance spectra using a box-car algorithm of 1000 pixels. This removed the variation in the level of flux and variance that had been caused by the edges of each echelle order. We then divided the smoothed flux by the smoothed variance to get a function of the S/N per pixel as a function of wavelength. The results from all 3 spectra were then combined, to form one continuous function of the S/N for the STIS instrument between 2050 → 3050 Å. In order to recreate spectra that contained Lyα absorbers at a redshift $z = 0.5$, this function was extrapolated to 1820 Å. The variation in S/N as a function of wavelength was small, so extrapolating over a bandpass of 230 Å would not have caused a large error. In regions where the wavelength between the spectra of PKS 1127-145 and HE 1122-1648 overlapped, a variance weighted mean value for the S/N was used.

Now that the spectra had the characteristics of our STIS E230M data, the Lyα absorbers could be identified and analysed using VPFIT.

7.2.2 Fitting the Lyα Absorbers with VPFIT

We had fired 675 sightlines at 2 different redshifts and in all 5 regions. This meant there were 6750 spectra to analyse using VPFIT. A pipeline was written that automatically fitted Voigt profiles for these spectra one at a time, and produced an overall Lyα absorber linelist for each snapshot.

When automatically identifying and fitting Lyα absorbers we used the same parameters that are described in Section 4.4. Each spectrum was sent to RDGEN, which would find the wavelength of the absorption lines, measure the EW, error in the EW and significance. All of the absorptions could only have been caused by the Lyα transition, and all those that had an EW significance ≥ 3 were sent to VPFIT.

Manual tests of this procedure on simulated spectra, that had a fiducial S/N across all pixels of 50, showed that Lyα absorbers that had $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))< 13$ had an EW significance ≥ 3. These absorbers were the ones that were removed by the significance cut-off when the S/N was the same as the real STIS data ($\gtrsim 2$). Hence like the observed data, the simulated Lyα absorbers were required to have a minimum observable column density of $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 13$.

When fitting the lines VPFIT would automatically drop any line that had a b-parameter less than 10 km s^{-1}, where we once again have used the minimum possible velocity expected by Janknecht et al. (2006). If VPFIT was unable to fit a line without the 1σ error in the column density exceeding $\sigma(\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2})))= 0.5$, then the b-parameter was fixed at 26.5 km s^{-1}. This was the median velocity of the real Lyα absorbers.

Once VPFIT had successfully fitted all the significant lines in the spectrum (a successful fit was deemed to have occurred when the value of the reduced χ^2 for the fit was $0.85 \leq \chi^2 \leq 1.15$), the lines were sent to a program that used equations (4.17) and (4.19) to measure a variance weighted EW and variance in the EW. SPECWIZARD always produced normalised spectra, so we did not consider any error in the level of the continuum. All lines that had a variance weighted EW significance ≥ 3 were then added to the linelist for that snapshot.

The problem with automatic fitting, particularly in the case of the $+2\sigma$ region, was that sometimes the fit would be so bad that even when the b-parameters were fixed the errors were still too large or the value of χ^2 too high. This occurred in no more than 10% of all the sightlines that contained absorption lines for each snapshot. A plot of those sightlines that had failed when compared to the location of galaxies showed that failure was common in those areas where the density of galaxies was very high. Thus it was important that these spectra were re-analysed, so that the Lyα absorbers could be included in the final linelist.

The spectra for those sightlines that had failed were re-made and automatically fitted again, only this time the S/N per pixel had been increased to 5. This was a sufficient level that all the errors in the column density, with or without a fixed b-parameter, decreased to $\sigma(\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2})))\leq 0.5$. The increased S/N was not set too high, so that that Lyα absorbers that had $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))< 13$ still had a variance weighted EW significance < 3.

Figure 7.1 shows the separate GIMIC regions at a redshift z = 1. These are plots of the sightlines along the z-axis that contained significant absorption lines. These have been superimposed on the galaxies from each snapshot. The plots are in the x-y plane, with the direction of the sightline pointing into the page. Figure 7.1 illustrates the large difference in the number of sightlines that contained absorbers, especially between the -2σ and $+2\sigma$ regions. The number of sightlines that contained significant absorption did not change significantly when comparing lines that were from the same snapshot but fired along a different direction.

Tables 7.2 and 7.3 show the number of sightlines for each snapshot along which an absorber was detected, (each is out of a total of 675), and the number of Lyα absorbers above a specified column density.

It was always expected that the denser GIMIC regions would have more sightlines that had at least 1 Lyα absorber, and that the $+2\sigma$ sphere would have the most absorbers that had $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\geq 17$. What was not expected was the large decrease with redshift in the number of spectra that contained absorption lines. Figure 1.7 showed a flat evolution in the line-density from z = 1 → 0.5. This

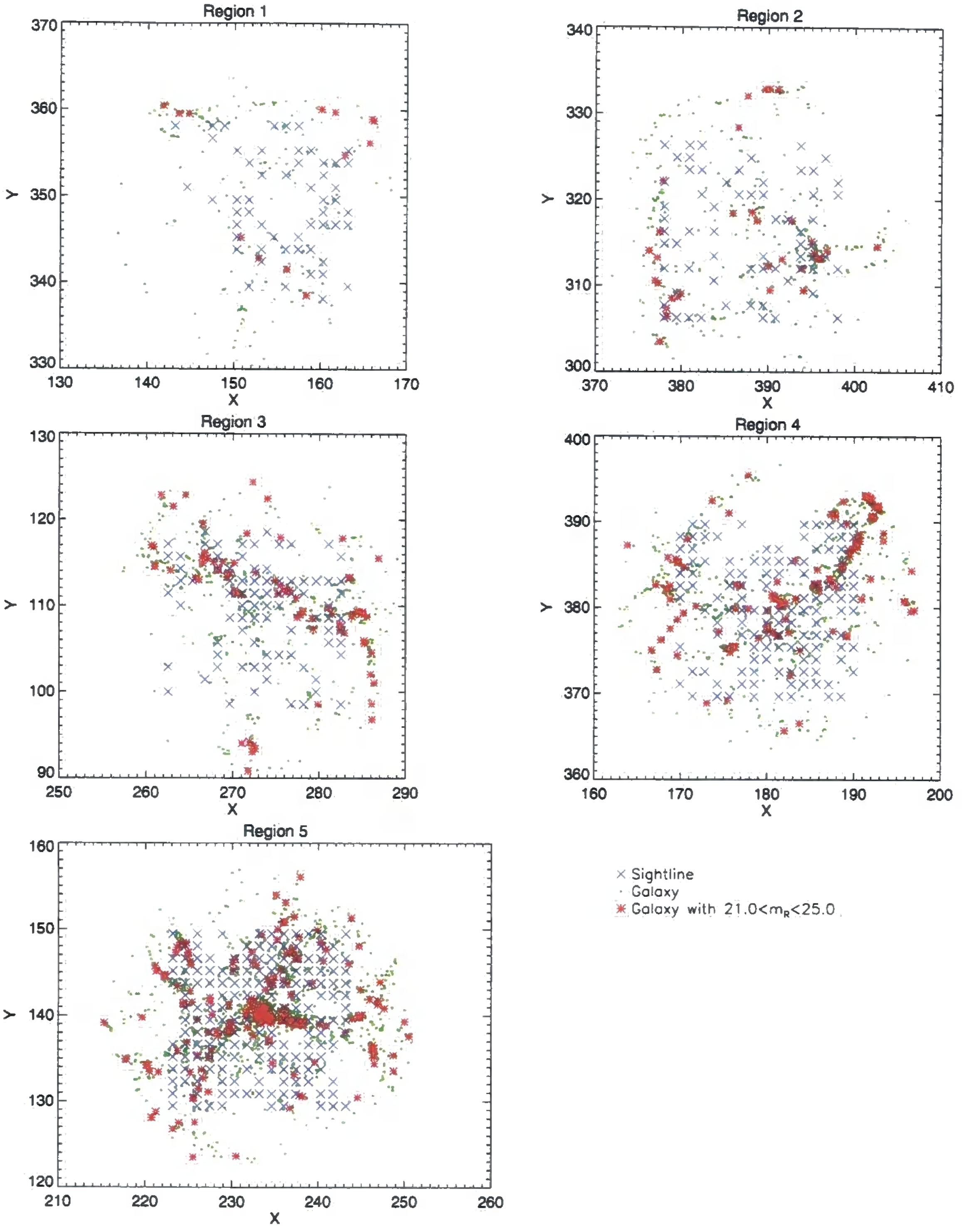


Figure 7.1 These plots show the location in the x-y plane of the galaxies (green and red asterisks) in the 5 GIMIC spheres at a redshift of z = 1. The sightlines that contained Lyα absorbers have then been marked with blue crosses and are directed into the page.

Table 7.2. The number of sightlines in each snapshot that contained a significant absorber with EW significance ≥ 3.

Redshift	snapshot				
	-2σ	-1σ	0σ	$+1\sigma$	$+2\sigma$
0.5	128	102	142	214	327
1.0	141	181	261	350	447

Table 7.3. The number of absorbers $\geq$ a given column density for each snapshot.

Snapshot	Column density cut $\log_{10}(N_{HI}\,(cm^{-2}))$				
	≥ 13	≥ 14	≥ 15	≥ 16	≥ 17
-2σ z = 0.5	140	6	2	0	0
-2σ z = 1.0	158	15	4	2	1
-1σ z = 0.5	122	6	1	1	0
-1σ z = 1.0	231	43	9	5	4
0σ z = 0.5	176	19	4	1	1
0σ z = 1.0	386	119	34	20	15
$+1\sigma$ z = 0.5	294	50	14	9	9
$+1\sigma$ z = 1.0	524	197	62	50	38
$+2\sigma$ z = 0.5	490	101	35	18	16
$+2\sigma$ z = 1.0	809	308	115	87	77

discrepancy arose because the size of the UV background at $z = 0.5$ was set too high.

7.3 Creating the Galaxy Catalogues

7.3.1 Identifying the Galaxies

In order to identify galaxies in a simulation that is made up of particles we had to make a decision as to what constituted a dark matter halo, and whether or not virialised particles within this structure could be defined as a galaxy.

The particles in each simulation were grouped based on the Friends-of-Friends (FoF) algorithm (Davis et al., 1985, Lacey and Cole, 1994). A dark matter particle was deemed to be a member of a group of particles if its distance from the rest was $\leq 0.2\times$ the mean inter-particle separation.

If at least 20 dark matter particles were identified in a group (C09), and this group contained a bound sub-structure of dark matter or baryonic particles, then this was deemed to be a dark matter halo.

The program SUBFIND then identified the closest dark matter particle from each baryonic particle. If this dark matter particle was part of a FoF group, or sub-group, then the baryonic particle also became part of this structure.

The total energy of each baryonic particle in the same halo, or sub-halo was then measured, and all those with a negative total energy were assumed to belong to the same bound system. The particles that represented a galaxy were the bound baryonic particles in these groups that had been assigned 'star' status (Dolag et al., 2008)[2].

7.3.2 Creating the Galaxy Magnitude Limited Sample

To create a galaxy catalogue for each snapshot the absolute R_{AB} magnitude was estimated for all of the galaxies identified in the snapshot that had a minimum

---

[2] All of the galaxy code and calculations, including the FoF, SUBFIND and magnitude calculations were either written and/or carried out by R. Crain, Swinburne University of Technology, PO Box 218, Hawthorn, Victoria 3122, Australia.

stellar mass of 1×10^9 $h^{-1}M_\odot$, and were contained within the 18 h^{-1}Mpc safe region. This meant there was a minimum number of 87 star particles for each galaxy.

The evolutionary track for each star particle, with values for metallicity and age, was calculated using the population software written by Bruzual and Charlot (2003). This returned a SED between the wavelengths of 91 Å to 160 μm for this particle, which was assumed to represent an entire stellar population.

This SED was then convolved with a Cousins R filter that was the nearest filter in bandpass and shape to the ESO R Bessel filter that was used during the FORS2 observations. The total absolute galaxy luminosity was equal to the sum of each of the constituent star particles. Using the redshift of the GIMIC region this was then converted to an apparent magnitude. At the time of this investigation the code did not contain dimming caused by dust. Those galaxies with an apparent magnitude $21.0 \leq R_{AB} \leq 25.0$ were then included in the catalogue.

Figure 7.2 shows that a linear relation exists between mass and luminosity for the galaxies that were to end up in our catalogue.

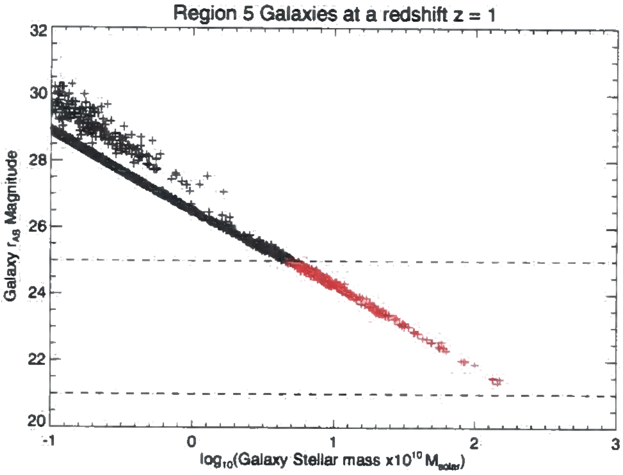


Figure 7.2 This figure shows a tight linear relation for the GIMIC galaxies, from the $+2\sigma$ region at a redshift of z = 1, between mass and luminosity for those galaxies that had $21.0 \leq R_{AB} \leq 25.0$. So had we known a reliable value to use for the mass-cut, this would have resulted in the same galaxy population to use in the correlation function. Those galaxies marked with a red cross and bound within the dashed lines were the ones that were included in the galaxy catalogue.

All of the galaxies that had $21.0 \leq R_{AB} \leq 25.0$ were added to a galaxy catalogue for each snapshot. Table 7.4 shows the total number of galaxies in each snapshot

Table 7.4. The number of galaxies in each GIMIC snapshot at redshifts z = 0.5 and 1

Region	z = 0.5 No. galaxies	 No. $21 \leq R \leq 25$	z = 1.0 No. galaxies	 No. $21 \leq R \leq 25$
1: -2σ	200	37	226	13
2: -1σ	366	74	369	33
3: 0σ	610	154	627	76
4: +1σ	753	167	847	121
5: +2σ	1587	458	1642	226

that had a stellar mass $\geq 10^9$ $h^{-1}M_\odot$ and the number of these that were included in the magnitude limited sample. This shows the large difference in the number of galaxies between the snapshots used in the correlation function, and helps to explain why there were so few absorber-galaxy pairs for the -2σ region.

7.4 Calculating the Cross and Auto-correlation Functions

The correlation function was calculated separately for each snapshot, where the methods and algorithms used for the simulated data were similar to those used for the real observations.

7.4.1 Identifying the Real Pairs

Both the final Lyα absorber linelist and galaxy table for each snapshot were sent to the pairing code, every galaxy within the 18 h^{-1}Mpc radius was then paired with every Lyα absorber, regardless of which direction the sightline had been fired. Thus the calculations that determined the π and σ separations depended on the values of ϕ and θ for the sightline. That is why the notation for an arbitrary direction for a line of sight is denoted i in the equations below. The j and k directions are then perpendicular to this axis.

The galaxy catalogue gave details of the position (q) of each galaxy in Millennium Simulation coordinates, as well as the peculiar velocity (v_p) along the x, y and z directions. These were used to find the co-moving distance to a galaxy along the line of sight for an observer at a redshift of 0.

First the separation Δd^i was found from the central coordinate of the sphere, O, to each galaxy. The redshift z_O at O was defined to be that of the GIMIC region.

$$\Delta d^i = q^i - O^i \tag{7.3}$$

This was then used to derive the difference in the velocity Δv_H^i caused by Hubble flow, relative to the centre of the sphere.

$$\Delta v_H{}^i = \frac{\Delta d^i H(z_O)}{1+z_O} \tag{7.4}$$

This was then added to the galaxy peculiar velocity along direction i. $\Delta v_{tot}{}^i = \Delta v_H{}^i + v_p{}^i$. The difference in redshift, between this galaxy along direction i and the centre was.

$$\Delta z^i = \frac{\Delta v_{tot}{}^i}{c}(1+z_O) \tag{7.5}$$

The redshift of a galaxy z_g, along direction i, for an observer at $z = 0$ was $z_g{}^i = z_O + \Delta z^i$.

The redshift of the absorber was the output from VPFIT. Equation (1.9) was then used to find the line of sight separations, π, between all Lyα absorbers and galaxies.

The projected separation σ between pairs, for an absorber at coordinate p, was given by equation (7.6)

$$\sigma = \sqrt{(p^j - q^j)^2 + (p^k - q^k)^2} \tag{7.6}$$

The coordinates $[p^j, p^k]$ were those chosen when the sightlines were created.

7.4.2 Generating Random Galaxies and Random Absorbers

The coordinates of a random galaxy were made by selecting 3 random numbers, using a uniform distribution, between 0 and 18 for coordinate shifts Δx, Δy and Δz from the centre of the sphere. The only restriction was $\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \leq 18$ h_{73}^{-1}Mpc.

The peculiar velocity for a random galaxy was the same as one from the GIMIC catalogue. This is because we used the same procedure as the real data, where each galaxy spawned the same number of randoms. These new coordinates and the velocity were then used to find the co-moving distance along the line of sight to each random galaxy.

The redshift of a random absorber was made by generating a random distance between ±10 h_{73}^{-1}Mpc (the distance along the line of sight from the centre of the cube within which the sightlines were contained), and using adapted versions of equations (7.3), (7.4) and (7.5) to convert this into a redshift.

A problem when creating random absorbers and galaxies was to decide on the ratio of how many random targets would be required for each 'real' one, so that the random distribution of binned pairs would not contain significant noise, and no bins would contain zero random pairs. In the real observations a value of 200 was chosen. With so many pairs in the GIMIC simulation, particularly in the $+2\sigma$ region, 200 random galaxies and absorbers for each 'real' target would take too long to compute and subsequently bin. An adequate solution, found by trials, was that if the number of galaxies with $21.0 \leq R_{AB} \leq 25.0$ in the snapshot was ≤ 100 then the ratio would be 20, ≥ 100 and the ratio was 10. If there were more than 300 galaxies then the ratio was equal to 1.

The line of sight and projected separation was then found between all real absorbers and random galaxies for the correlation estimator that uses equation (2.2a), and between all random absorbers and random galaxies to use in equation (2.2b).

The pairs from each GIMIC snapshot were then binned separately, using the same algorithm and level of binning that was used for the observations. A weighted mean using the factors from Table 7.1 was used to compute the overall correlation function and error.

Figure 7.3 shows the real and random pair distributions from region 3 at a redshift z = 1. Figure 7.3 illustrates the non-obvious distribution of pairs that arises when correlating targets that are all within a sphere of radius 18 h_{73}^{-1}Mpc, where the absorbers are contained within a central cube of length 20 h_{73}^{-1}Mpc. This should be compared with Figures 5.9 and 5.10 from the observations, where all the Lyα absorbers and galaxies were contained within a cylinder, 5 co-moving h_{70}^{-1}Mpc in diameter, yet $\gg$ 1000 co-moving h_{70}^{-1}Mpc long.

The pairs along the line of sight in Figure 7.3 that are $\geq$ 36 h_{73}^{-1}Mpc apart are caused by the apparent increase in separation because of the galaxy peculiar velocity.

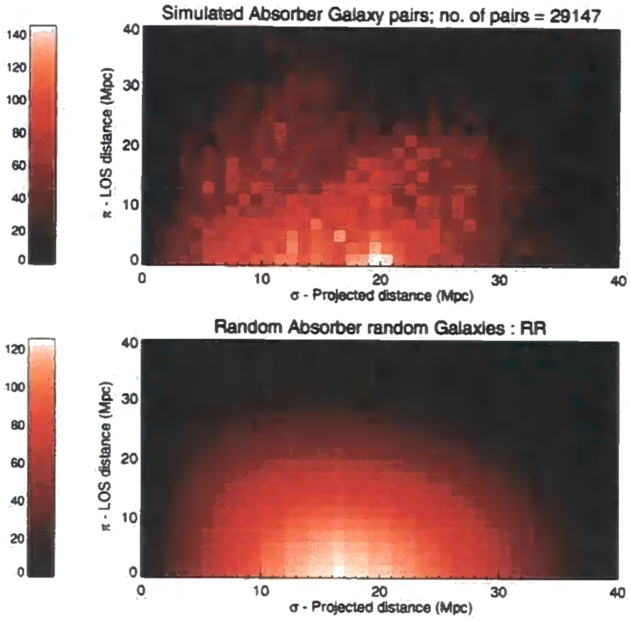


Figure 7.3 The real Lyα absorber-galaxy pairs and the random-random pairs in the 0σ region, at z = 1, binned $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{73}^{-1}Mpc.

For the subsequent binning and plots the correlation function was only computed out to a maximum distance of $\sigma = 5$, $\pi = 20$ h_{73}^{-1}Mpc. This was so edge effects would be avoided and to directly compare with the observations.

7.4.3 The Galaxy Auto-Correlation Function

As a means to compare ξ_{AG} with ξ_{GG}, and to make sure our methods were correct, we computed the galaxy auto-correlation for each snapshot. The technique for finding the line of sight and projected separation between galaxy pairs and generating random positions was the same as that described above.

The one difference was that, in order to avoid having the same pair of galaxies sampled twice in the same snapshot when calculating ξ_{GG}, only one axis was chosen for the line of sight. The initial results for ξ_{GG} were made with the line of sight along the z-axis. ξ_{GG} was also calculated along the x and y axes as well.

7.4.4 Calculating the Errors for the ξ_{AG} and ξ_{GG}

With numerous lines of sight we could now use the same methods to calculate an error in ξ_{AG} as those in Wilman et al. (2007) and Ryan-Weber (2006). We used jackknife resampling, the formula for which is given in equation (5.4). All of the pairs were removed that contained an absorber from each consecutive sightline, and the correlation function was re-calculated.

The error in the auto-correlation, $\sigma^2_{\xi_{\mathrm{GG}}}$, presented a harder problem. In real observations the population of galaxy pairs to remove from the sample when jackknifing can be chosen by removing those that surround a particular quasar sightline, or by RA. This was how $\sigma^2_{\xi_{\mathrm{GG}}}$ was calculated in Wilman et al. (2007) and Ryan-Weber (2006). All of our galaxies were contained within the same sphere of radius 18 h^{-1}Mpc, therefore we couldn't use this method.

Instead $\sigma^2_{\xi_{\mathrm{GG}}}$ was estimated by removing a certain percentage of the galaxies, then re-calculating the auto-correlation function. Consider a field of view that contains N galaxies that surround a quasar sightline. Assuming that cosmic variance in different galaxy fields would cause a difference of $\sim \sqrt{\mathrm{N}}$ in the number of galaxies, we removed $\sqrt{\mathrm{N}}$ randomly selected galaxies from our catalogue. This was done $\sqrt{\mathrm{N}}$ times, randomly selecting different galaxies each time so that no galaxy was removed more than once, then jackknifing these results to determine an error.

7.5 Results and Discussion

Even with measurements from just 2 epochs there were many different plots and comparisons that could be made. In order that we can answer those questions posed in Section 2.1, the correlation function was studied for the following topics.

1. Is there a difference in ξ_{AG} when using the GIMIC simulation between the two different correlation estimators, equations (2.2a) and (2.2b)?

2. How does ξ_{AG} vary, depending on the GIMIC region, as a function of redshift, or as a function of the column density?
3. How does the cross-correlation compare with the galaxy auto-correlation function, and what does this say about the distribution of matter in the IGM?
4. When the column density of the Lyα absorbers are restricted to the same range as the observations, with the $+2\sigma$ pairs removed, how do the simulations compare with the results from Chapter 5?
5. Now that there are sufficient pairs, how does the correlation function change if we decrease the size of the bins?

Each of these questions will be tackled separately.

7.5.1 The Correlation Estimator

The same result was seen that was observed in Section 5.2.5. There was no significant difference in the value of ξ_{AG} if either equation (2.2a), where the denominator in the estimator is the number of real-random pairs, or equation (2.2b), where the denominator is the number of random-random pairs was used.

In each case the values for ξ_{AG} agreed within the error margin. For example, ξ_{AG} at a redshift z = 1, from all 5 regions, with column densities $\log_{10}(N_{HI}\,(cm^{-2}))\geq 13$, in the bin $\Delta\sigma = 0 - 1$, $\Delta\pi = 0 - 2$ h_{73}^{-1}Mpc had a value of $\xi_{AG} = 2.66 \pm 0.78$ when using the RR estimator. ξ_{AG} was equal to 2.79 ± 0.80 when the DR estimator was used. Thus all results to ξ_{AG} that are quoted below were made using equation (2.2b).

Figure 7.4 shows the small differences in the distribution of DR and RR pairs from region 3 at a redshift z = 1.

7.5.2 How does ξ_{AG} Vary Depending on the GIMIC Region?

As the GIMIC regions represent significant differences in the matter density, and in the environment, this provided a unique opportunity to investigate to what extent these factors have on the value of ξ_{AG}. The expectation was that ξ_{AG} would be highest in the $+2\sigma$ region. Table 7.3 did show that this region has the highest mean column density, with the most absorbers with $\log_{10}(N_{HI}\,(cm^{-2}))\geq 17$. This, accord-

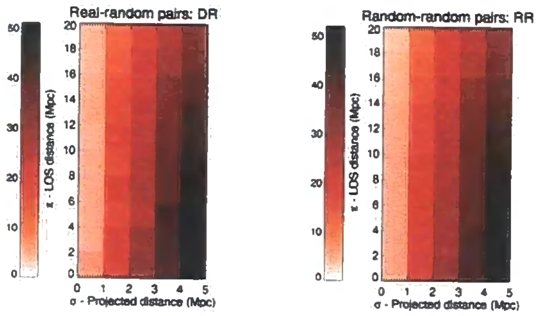


Figure 7.4 A comparison of the data-random, random-random pairs from the 0σ sphere at a redshift of z = 1. Minor differences in the number of pairs per bin created $\ll 1\sigma$ variations in $\xi_{\rm AG}$ found using either the data-random or random-random estimator.

Table 7.5. Values in the central bin for $\xi_{\rm AG}$ for each GIMIC snapshot at a redshift z = 1 and 0.5.

Sphere	$\xi_{\rm AG}$ at z = 0.5	No. of DD,RR pairs	$\xi_{\rm AG}$ at z = 1.0	No. of DD,RR pairs
1: -2σ	2.37 ± 1.92	3, 0.89	3.55 ± 3.18	2, 0.44
2: -1σ	1.11 ± 1.64	4, 1.90	5.64 ± 2.43	10, 1.51
3: 0σ	6.06 ± 2.32	37, 5.24	1.90 ± 1.19	15, 5.18
4: +1σ	1.53 ± 0.77	25, 9.87	1.82 ± 0.73	35, 12.4
5: +2σ	3.69 ± 0.71	136, 29.0	2.05 ± 0.41	100, 32.8
6: combined	3.96 ± 1.21	-	2.65 ± 0.78	-
7: +2σ removed	3.37 ± 1.23	-	2.45 ± 0.77	-

ing to Wilman et al. (2007), is a critical column density at which $\xi_{\rm AG}$ significantly increases in value in the bin of smallest separation.

Figure 7.5 shows the variation at a redshift z = 1 in the number of Lyα absorber-galaxy pairs, $\xi_{\rm AG}$ and $\sigma_{\xi_{\rm AG}}$ for all 5 regions with $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2})) \geq 13$. Table 7.5 gives the values in the central bin for these figures. The additional columns give the number of real (DD) and normalised random (RR) pairs for this bin, and the results for the central bin at a redshift z = 0.5. The bottom two rows are the combined result, where $\xi_{\rm AG}$ and $\sigma_{\xi_{\rm AG}}$ have been added in a weighted mean using the factors in Table 7.1, and the combined result where region 5 has been removed.

Figure 7.5 shows that $\xi_{\rm AG}$ peaked in all regions in the central bin, and that these values were different by $\leq 2\sigma$. Region 5 did not, as opposed to what was expected,

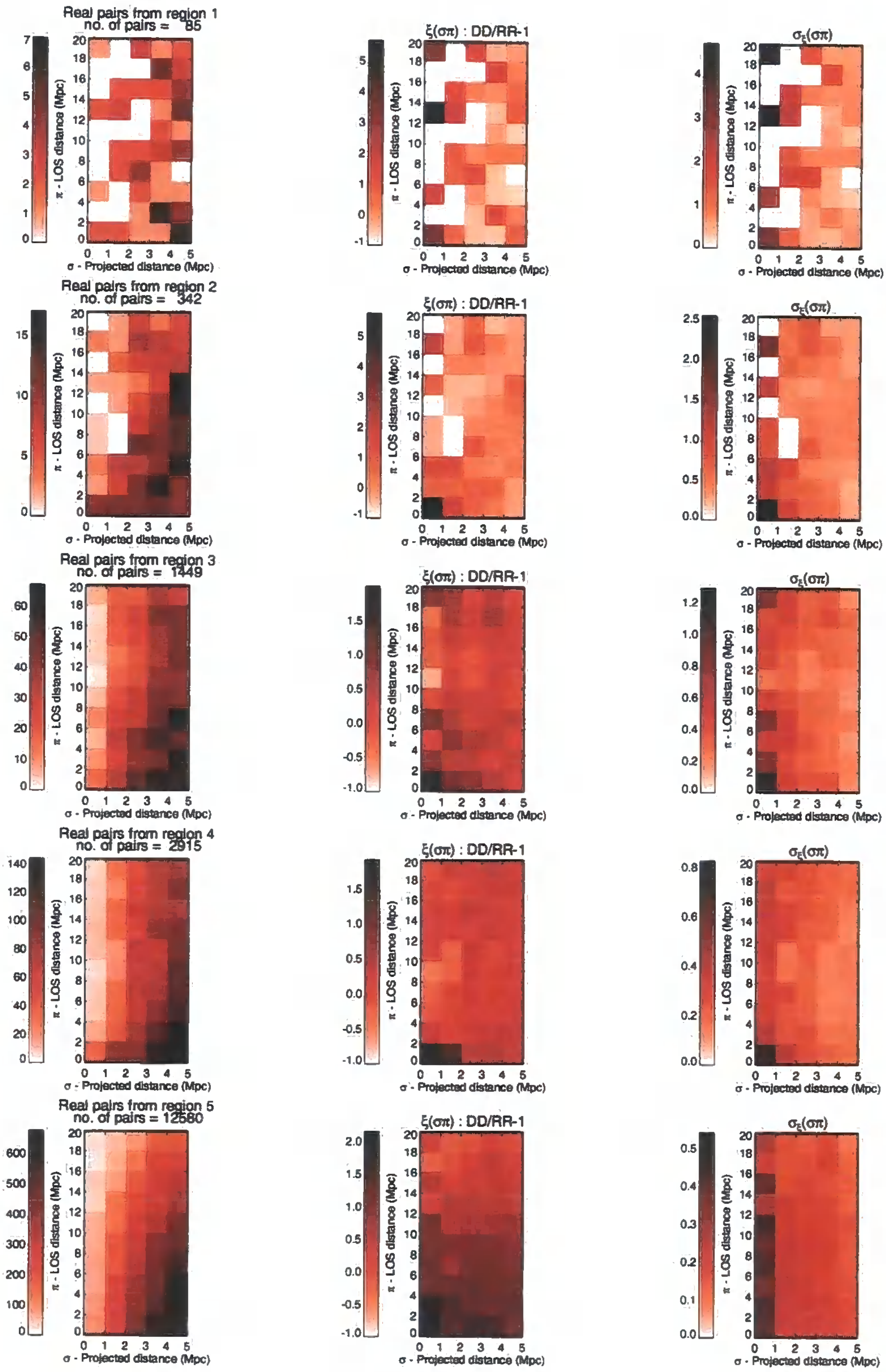


Figure 7.5 These plots show the real pairs, ξ_{AG} and $\sigma_{\xi_{AG}}$ for the 5 GIMIC regions, at a redshift of z = 1 and binned $\Delta\pi = 2$, $\Delta\sigma = 1$ h_{73}^{-1}Mpc. $\sigma_{\xi_{AG}} = 0$ if there were 0 real pairs in that bin.

have a significantly higher value. A reason for this is that it is not the region that was important, rather it was the column density of those absorbers that were paired. Despite possessing the most absorbers with $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \geq 17$, region 5 also had the most (501 according to Table 7.3) that had $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) = 13 - 14$. Thus because all of the pairs were binned, regardless of the column density, the level of correlation was the same.

The plots in Figure 7.5 for regions 1 and 2 show the consequences of there not being enough pairs in a sample. Even though the value for the correlation in the central bin agrees with that in regions 3, 4 and 5, the data is dominated by noise.

Only marginal evidence for velocity dispersion along the line of sight was detected, and even then only in the $+2\sigma$ sphere. A strong "finger-of-god" had been expected, particularly in region 5, where because of the high density cluster virialised galaxies have a faster peculiar velocity. Hence either the sightlines that were incident in these virialised regions did not detect any neutral gas, or the gas is not moving at a sufficient velocity relative to the galaxies to enhance the "finger-of-god". A third possibility is that any elongation is being smoothed out by the large binning along the projected separation. This was investigated when a smaller bin size was used, with results that are summarised in Section 7.5.7 below.

7.5.3 How does ξ_{AG} Vary as a Function of Redshift

The expectation was that ξ_{AG} would increase in the central bin as the redshift decreased. This is because gas collapses over time, causing more Lyα absorbers to collect in the vicinity of the galaxies in our sample.

Figure 7.6 shows ξ_{AG} and $\sigma_{\xi_{\mathrm{AG}}}$ for the combined GIMIC data at redshifts $z = 0.5$ and 1.

The value in the central bin for the combined GIMIC regions only increased from $\xi_{\mathrm{AG}} = 2.65 \pm 0.78$ to 3.96 ± 1.21. This means there was only tentative evidence, with an increase of 0.9σ, for an increase in the level of correlation with redshift.

That the correlation should stay the same is surprising because there were significant changes in the number of absorbers and galaxies present in each region. The values in Table 7.4 show that the number of galaxies in each region with a mass

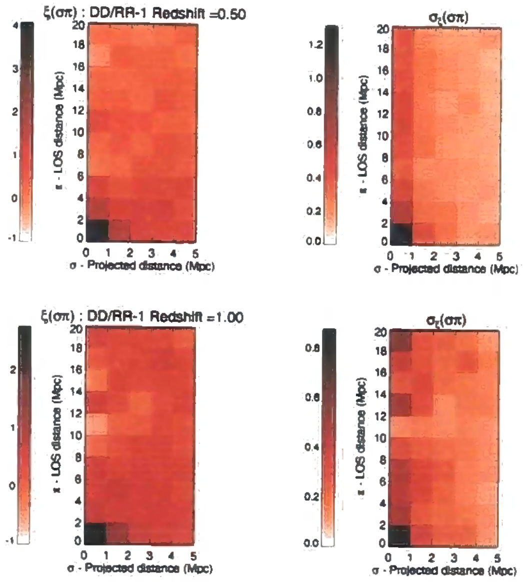


Figure 7.6 Only marginal evidence is found for the evolution with redshift in the level of ξ_{AG} as z decreases from 1 to 0.5.

$\geq 1 \times 10^9$ $\mathrm{h^{-1}M_\odot}$ decreased, (albeit by a small fraction), as the redshift decreased. Meanwhile the number that had $21.0 \leq R_{\mathrm{AB}} \leq 25.0$ increased by a factor of ~ 2 in all but the $+1\sigma$ sphere, where the increase went from 121 to 167 galaxies.

Despite there being more galaxies with which to correlate at z = 0.5, the number of pairs did not increase significantly because of the decrease in the number of Lyα absorbers. In regions 3, 4 and 5 the number of absorbers decreased by a factor of 54%, 44% and 40% respectively. Such a large drop is surprising because observations show the number density of Lyα absorbers is flat over the redshift separation $z = 1 \rightarrow 0.5$ (Janknecht et al., 2006). A solution to this would have been to scale the UV background in the GIMIC simulation, so that the line-density and mean optical depth as a function of redshift would match those values from Janknecht et al. (2006) and Kirkman et al. (2007) in Figures 1.7 and 1.6.

We also did not detect a significant "finger-of-god" at z = 0.5. The data only showed marginal evidence for a stronger correlation at bins away from the centre along the line of sight than at z = 1. For example the bin at $\Delta\sigma = 0-1$, $\Delta\pi = 2-4$ $\mathrm{h_{73}^{-1}Mpc}$ at a redshift of z = 1 has a value of $\xi_{\mathrm{AG}} = 0.57 \pm 0.44$. By z = 0.5 this had increased to 1.53 ± 0.65.

We could argue that a stronger "finger-of-god" should exist at lower redshifts, because of an acceleration in the peculiar velocity of gas and galaxies as the density of the collapsed structures increased. However a change in redshift of only $\Delta z = 0.5$ may be too small to detect a significant difference.

7.5.4 How does ξ_{AG} Vary as a Function of the Column Density?

The number of absorbers in Table 7.3 is too small for it to be possible to do a separate study for each region with cuts in the column density.

The combined data was analysed instead, with measurements for the correlation taken with column densities that varied from; $\log_{10}(N_{HI}\,(cm^{-2}))= 13-14$, $14-15$, $15-17$ and $\log_{10}(N_{HI}\,(cm^{-2}))\geq 17$. The larger cut that spanned two decades was necessary to avoid shot noise caused by too few pairs. Those absorbers with $\log_{10}(N_{HI}\,(cm^{-2}))\geq 17$ were binned together because of the limitations in GIMIC when trying to simulate Lyman-limit, sub-damped and damped absorbers that do not have Lorentzian profiles.

Figure 7.7 shows ξ_{AG} binned according to the column density at a redshift of $z = 1$.

Regardless of the range in the column density the peak in ξ_{AG} remained in the central bin. This is except for those that had $\log_{10}(N_{HI}\,(cm^{-2}))= 15-17$, where noise dominated and there was only a 0.6σ difference in ξ_{AG} between the bins at $\sigma = 0-1$, $\pi = 0-2$ and $\sigma = 1-2$, $\pi = 0-2$ h_{73}^{-1}Mpc.

We do not re-produce an anti-correlation in the central bin for systems with a low-column density, $\log_{10}(N_{HI}\,(cm^{-2}))\leq 14$, for which there was only marginal evidence in the observations. Thus when binned to this level the simulation predicts that these Lyα absorbers are not expelled from the galactic halo by winds.

ξ_{AG} for the low column density absorbers remained very low, with a central peak of only $\xi_{AG}= 1.87\pm0.63$ and a value ~ 0 for most of the other separations. With no sign of a significant correlation with galaxies this adds further evidence to suggest that it is Lyα absorbers with a column density $\log_{10}(N_{HI}\,(cm^{-2}))\leq 14$ that make up the diffuse gas in the IGM.

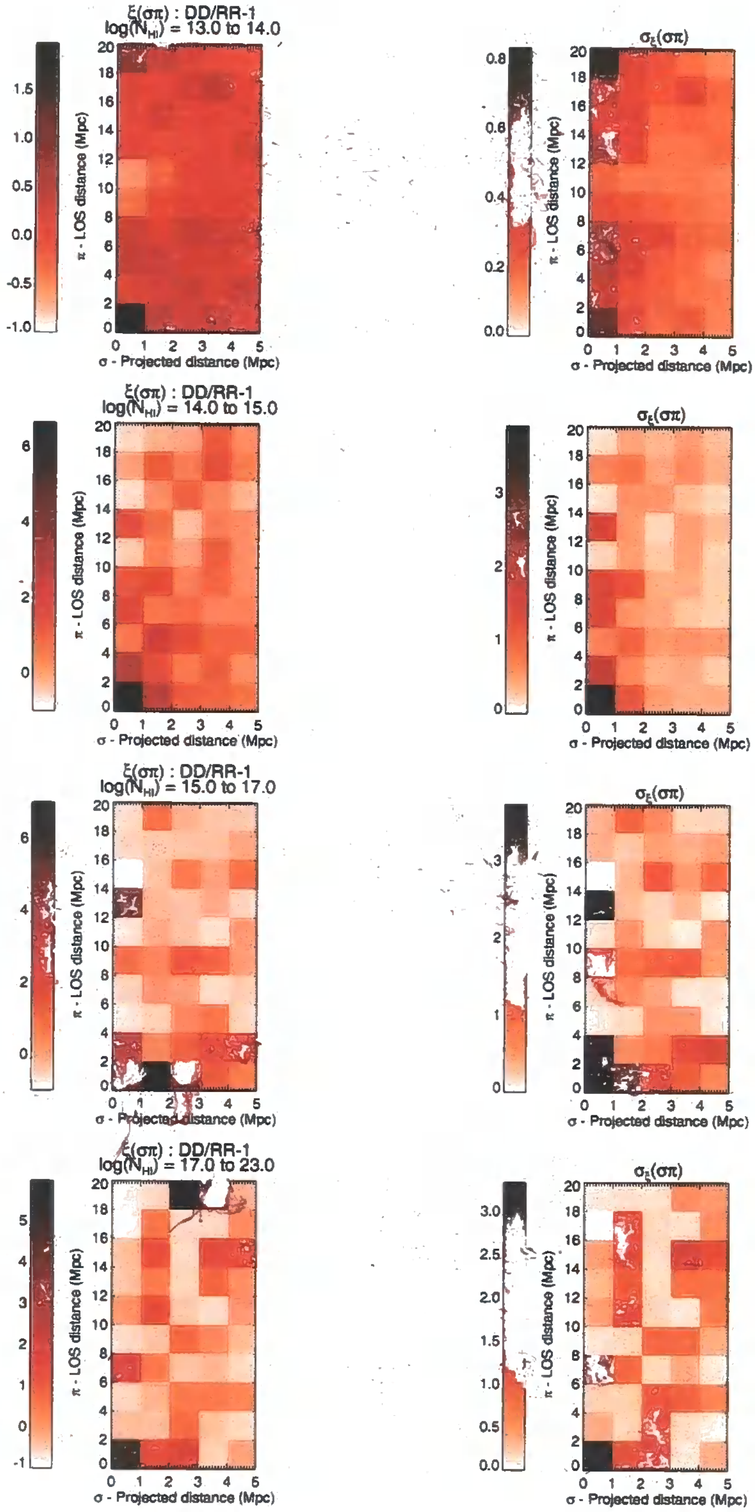


Figure 7.7 The combined GIMIC data, binned according to the column density. At a redshift z = 1, only a slight increase in the correlation is observed as the column density of the Lyα absorbers exceeds $\log_{10}(N_{HI}\,(cm^{-2})) \geq 14$.

The correlation of absorbers with $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 14-15$, while containing a lot of noise, did show a marginal 1.2σ increase in the central bin to $\xi_{AG}= 6.54\pm3.80$. ξ_{AG} then remained $\leq 1\sigma$ of this value in the other central bins, with values of $\xi_{AG}= 4.33\pm3.38$ for $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 15-17$, and $\xi_{AG}= 5.03\pm3.23$ for absorbers with $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq 17$.

There were few constraints that could be made because of the level of noise, however what we can observe from Figure 7.7 was that, apart from the slight increase in ξ_{AG} as the column density exceeded $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))= 14$, there was no evidence for a significant increase in the cross-correlation at a particular column density at a redshift $z = 1$. At bins away from the centre the correlation between galaxies and high column density systems ($\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq 15$) remained low with $\xi_{AG}\equiv 0$. Any bin that exhibited correlation also had a high error margin.

Variations in the column density also do not have any impact on elongation along the line of sight, there being no significant "finger-of-god" in any of the plots in Figure 7.7.

7.5.5 How do the Lyα Absorber-Galaxy Cross-Correlation and Galaxy Auto-Correlation Compare?

Figure 7.8 shows the galaxy auto-correlation when observed along the z-axis of GIMIC regions 3, 4, 5 and for the combined data at $z = 1$. Even though the pair statistics contributed to the weighted mean, ξ_{GG} for regions 1 and 2 have not been included in Figure 7.8 because they contained 8 and 52 pairs respectively within $\Delta\sigma = 0-5$, $\Delta\pi = 0-20\ h_{73}^{-1}$Mpc and so were dominated by noise.

Comparing ξ_{GG} just between the GIMIC spheres we saw no significant difference in the auto-correlation for the central bin. The value for the combined data set was $\xi_{GG}= 27.05\pm4.06$.

The difference between the GIMIC spheres was the large "finger-of-god" that was present in regions 3, 4 and 5 which became more significant with each region. This reached a peak in the $+2\sigma$ region because most of these galaxies were virialised and belonged to the central dense cluster. Hence all had large peculiar velocities.

When comparing the galaxy auto-correlation function from Hawkins et al. (2003)

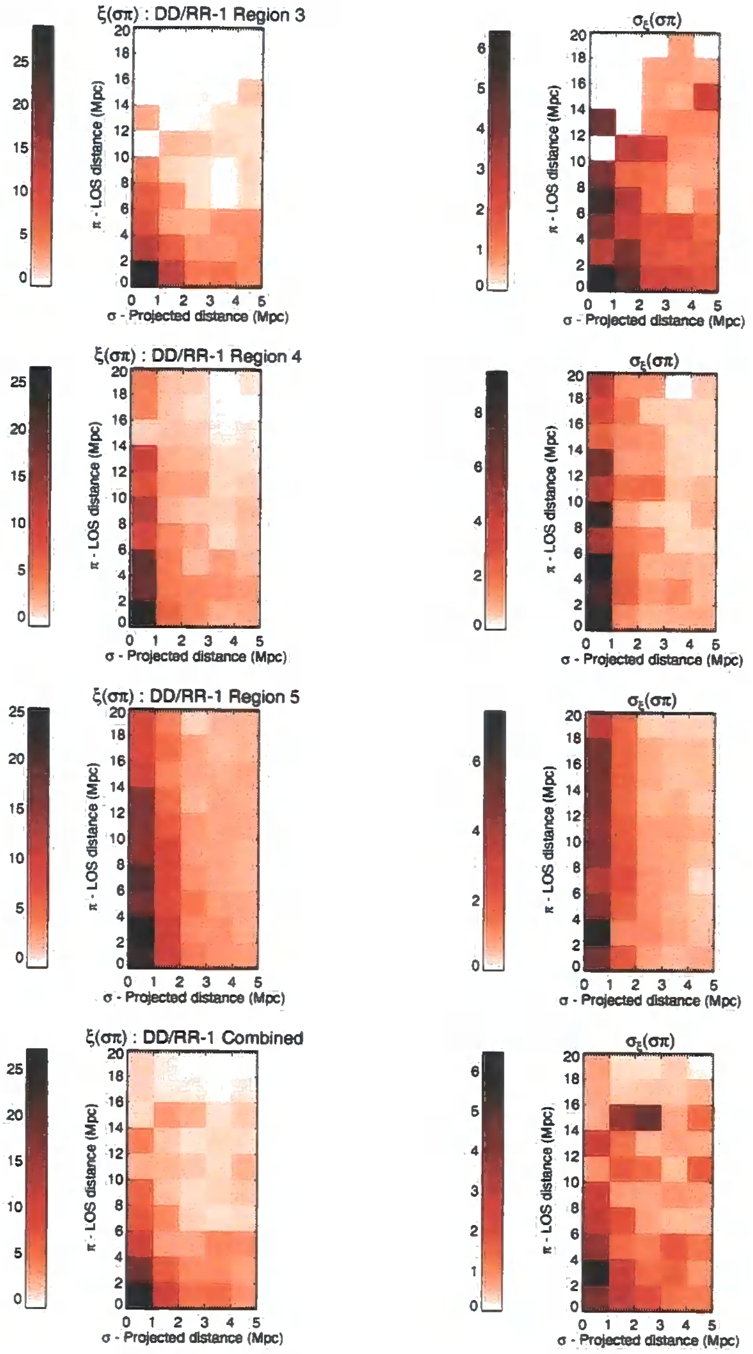


Figure 7.8 The galaxy auto-correlation ξ_{GG} at a redshift z = 1 for the GIMIC regions 3, 4, 5 and for the combined data. The value of ξ_{GG} in the central bin, while much greater than ξ_{AG}, is not seen to differ between the spheres. A significant "finger-of-god" is seen with increasing prominence in regions 3, 4 and 5.

and Pollo et al. (2005) with the upper limit in $\xi_{\rm AG}$ for the observations, there was a $> 5\sigma$ difference between $\xi_{\rm GG}$ and $\xi_{\rm AG}$ in the central bin. At a redshift $z = 1$ the difference in the central bin between $\xi_{\rm GG}$ and $\xi_{\rm AG}$ for the GIMIC simulation was 5.9σ. Thus when the pairs were plotted in redshift space, we continued to see a much greater value in the galaxy auto-correlation than the cross-correlation at small separations.

That we did not observe a significant "finger-of-god" in the cross-correlation functions at $z = 0.5$ and 1 in Figure 7.6, while in the auto-correlation function of Figure 7.8 we do, is also important. What this implies is that while galaxies have a high velocity with respect to one another, the gas is more quiescent, and is perhaps being accreted along filaments and into galaxy haloes at a slower rate.

There was no significant change in the auto-correlation when the galaxies at $z = 0.5$ were compared. The central bin for the combined data had a value $\xi_{\rm GG} = 24.94 \pm 4.82$. An increase in $\xi_{\rm GG}$ as the redshift decreased (such as the results from Hawkins et al. (2003) and Pollo et al. (2005), where $\xi_{\rm GG}$ increased from $\sim 10 \rightarrow \sim 17$), was expected. However, we were still hampered correlating galaxies in the 18 h_{73}^{-1}Mpc spheres. Thus neither $\xi_{\rm AG}$ or $\xi_{\rm GG}$ showed signs of significant evolution as the redshift decreased from $z = 1 \rightarrow 0.5$.

Uncertainty in $\xi_{\rm GG}$ will have arisen when sampling the GIMIC spheres because small scale variations in the structure along a particular axis could enhance or diminish the auto-correlation function. This was confirmed in Figure 7.9, where $\xi_{\rm GG}$ has been recalculated along the x and y direction. Values for the auto-correlation in the central bin changed by $\sim 1\sigma$ depending on the line of sight. We continued to see significant "fingers-of-god" however, regardless of the direction and $\xi_{\rm GG}$ still remained significantly larger than $\xi_{\rm AG}$ in the central bin. Thus we can rule out differences in structure along different lines-of-sight as having a large impact on the result of $\xi_{\rm GG}$, and can claim the GIMIC regions were reasonably isotropic.

7.5.6 How does $\xi_{\rm AG}$ Compare with the Observational Data?

When simulating the 2D 2-point correlation function an effort had been made to only collect data using the same constraints and the same methods as the observations.

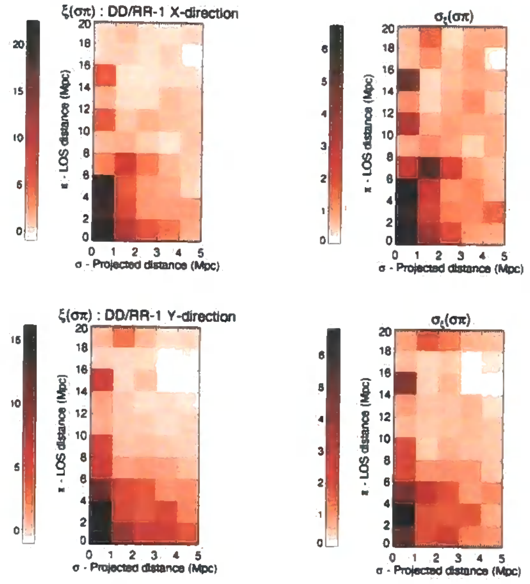


Figure 7.9 Small scale variations in the distribution of galaxies in the GIMIC spheres caused ξ_{GG} to change by $\sim \pm 1\sigma$ when observed along the x, y and z-axis. These variations did not change the fact that ξ_{GG} was still much larger than ξ_{AG} in the central bin.

Thus in order that the observations could be recreated we had to further limit the GIMIC sample. In the STIS linelist only 11 out the 141 Lyα absorbers that were correlated (7.8%) had a column density $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \geq 15$, and only 1 had $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) \geq 17$. It was also highly unlikely that either of the sightlines towards HE 1122-1648 or PKS 1127-145 had passed through or near a galaxy cluster (that was represented by the $+2\sigma$ region), within the redshift range of the E230M spectra. All of our absorptions had Gaussian dominated profiles. This was why ξ_{AG} was recalculated, only with the $+2\sigma$ region removed and a column density range of $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2})) = 13 - 15$.

With the removal of region 5 two alternative weighting schemes were considered. One of these, method 1, was where the weight previously counted towards region 5 was split equally between the remaining 4 regions. This meant that the factors were now equal to those weights listed in Table 7.1 plus $\frac{1}{48}$.

It was the highest density fraction that was removed from the sample, and this represented $\frac{1}{12}$ of the Millennium Simulation. So one could argue that the remaining high density region needed a stronger weighting. In method 2 we weighted region 4

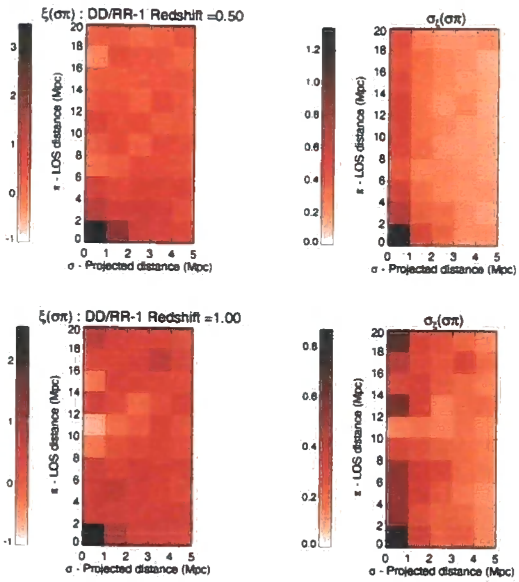


Figure 7.10 In order to more accurately simulate our observations, $\xi_{\rm AG}$ was recalculated with region 5 removed and the Lyα column densities restricted to $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))= 13-15$. These plots show the combined result from regions 1-4 of $\xi_{\rm AG}$ and $\sigma_{\xi_{\rm AG}}$ at z = 0.5 and 1.

by $\frac{1}{4}$ rather than $\frac{1}{6}$.

Figure 7.10 shows the weighted $\xi_{\rm AG}$ at redshifts z = 0.5 and 1 for this sample. Method 1 has been used to calculate the weighted mean.

No significant difference arose between using the two different weighting strategies. Values in the central bin at z = 1 were $\xi_{\rm AG}= 2.45 \pm 0.77$ and 2.34 ± 0.71 for methods 1 and 2 respectively.

Having removed the $+2\sigma$ region and having restricted the absorbers that were correlated to those with $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))= 13-15$, the result in the central bin was only marginally less than the measurement of $\xi_{\rm AG}$ in Figure 7.6. At a redshift of z = 0.5 the central bin was $\xi_{\rm AG}= 3.37 \pm 1.23$. With the removal of the clustered galaxies and high column density absorbers the expectation was that the correlation would have had a more significant decrease. Thus the contribution to the correlation signal in the central bin, made by those absorbers that had $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))\geq 15$, was not as strong as was expected.

This result also has implications when compared to the cuts in column density in Figure 7.7. The peak in the cross-correlation when the column densities were restricted to $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}})) \leq 14$ was $\xi_{\mathrm{AG}} = 1.83 \pm 0.63$, only marginally less than that in Figure 7.10.

What we can conclude from this is that an increase in column density that spans orders of magnitude only causes a gradual increase in the level of correlation. This is the same result that was observed by Wilman et al. (2007). Above $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}})) = 17$ they found a jump in ξ_{AG} to 15 ± 12. With such a large error this does not contradict our value of $\xi_{\mathrm{AG}} = 5.03 \pm 3.23$.

Our result does not show a critical column density at which a significant ($\geq 3\sigma$) jump in the correlation is observed. We should note however that it was the Lyα absorbers with $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm^{-2}})) \geq 17$ that we had problems modelling, and there are countless examples in the literature where DLAs are only observed within or near to galaxy haloes (Monier et al., 1998, Rao et al., 2003a, Smette et al., 1995, Zwaan et al., 2008). We could speculate that if a significant change in the level of correlation does occur it is only likely for the DLAs that have a column density $\mathrm{N_{HI}} \gg 10^{17.3}$ cm^{-2}.

In our observations we detected an upper-limit in ξ_{AG} of 1.83 and this was not at the centre. That this peak wasn't in the central bin but at a distance of $\sigma = 3-4$, $\pi = 8-10$ $\mathrm{h_{70}^{-1}}$Mpc has been attributed to the large error when using only 194 pairs. Thus the level of correlation when binned $\Delta\sigma = 1$, $\Delta\pi = 2$ $\mathrm{h_{73}^{-1}}$Mpc, at redshifts $z = 0.5$ and 1 is in relatively good agreement with this upper-limit, with differences of 1.3, and 0.8σ respectively. When we compare the value in the central bin from Figure 5.20 we get differences of 2.4 and 2.3σ.

7.5.7 Reducing the Bin Size to $\Delta\sigma = 0.4$, $\Delta\pi = 1$ $\mathrm{h_{73}^{-1}}$Mpc

Figures 7.5 and 7.6 had shown tentative evidence for "fingers-of-god" along the line of sight, particularly for the high σ GIMIC regions and at redshift $z = 0.5$. The question as to whether or not these features are significant would be easier to answer with a finer bin size. It is possible that any elongation along the line of sight was diminished by having to average out over 1 $\mathrm{h_{73}^{-1}}$Mpc along the projected separation.

Figure 7.11 shows plots of ξ_{AG} from the GIMIC regions 3, 4 and 5 at a redshift $z = 1$ and the combined data at redshifts $z = 0.5$ and 1, binned $\Delta\sigma = 0.4$, $\Delta\pi = 1$

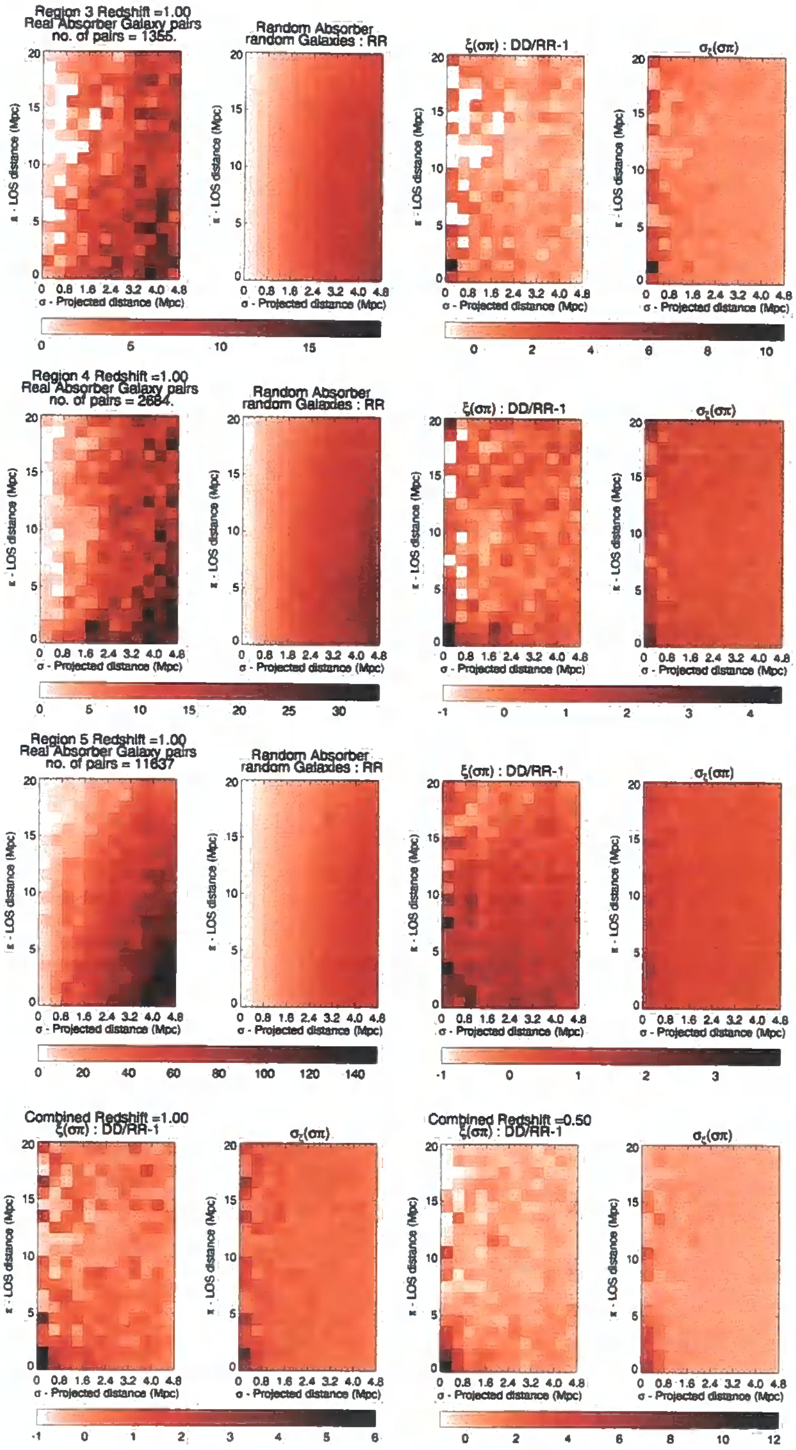


Figure 7.11 The Lyα absorber-galaxy pairs, random pairs, ξ_{AG} and $\sigma_{\xi_{\mathrm{AG}}}$ from regions 3, 4, 5 and the combined data at z = 1, and for the combined data at z = 0.5. The pairs were binned $\Delta\sigma = 0.4$, $\Delta\pi = 1$ h_{73}^{-1}Mpc.

h_{73}^{-1}Mpc.

With this finer level of binning we can see that the correlation function no longer always peaked in the central bin and is higher than what was previously observed at small separation, but with a larger jackknife error. The correlation was seen to drop to a value comparable with that at larger separations by $\sigma = 2$ h_{73}^{-1}Mpc. This finer binning also shows that each of the peaks along the line of sight, that once suggested the presence of a "finger-of-god", may just have been noise as the associated $\sigma_{\xi_{AG}}$ is very large.

Despite dealing with a different galaxy population, column density range and redshift, what we are seeing may be the difference that was observed between the results of Pierleoni et al. (2008) and Ryan-Weber (2006). When there are enough pairs at small projected separation, all fluctuations along the line of sight that suggest there may be a "finger-of-god" disappear.

In Ryan-Weber (2006) evidence for a strong "finger-of-god", and a ξ_{AG} that was comparable to ξ_{GG}, was found by Pierleoni et al. (2008) to be caused by the uncertainty when only 27 lines of sight were used. No "finger-of-god" was observed when Pierleoni et al. (2008) used 999 sightlines along the 3 Cartesian axes of their simulation. This difference was illustrated in Figures 2.2 and 2.3.

Pierleoni et al. (2008) found that elongation along the line of sight could arise if a significant number, of a small sample of sightlines, passed through virialised regions. This is probably why we observed tentative evidence for elongation along the line of sight from the the $+2\sigma$ region.

7.6 A Summary of these Results

Summarising the results from the simulated 2D 2-point correlation function we found that:

- When binned $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{73}^{-1}Mpc, there was no significant difference in ξ_{AG} between the separate regions. Region 5 showed tentative evidence for elongation along the line-of-sight, this could be due to an increase in the number of virialised systems.

- $\xi_{\rm AG}$ did not significantly change as the redshift decreased from z = 1 → 0.5, with an increase in the central bin of $\xi_{\rm AG} = 2.65 \pm 0.78 \rightarrow 3.96 \pm 1.21$.
- Only a mild increase ($< 3\sigma$) in $\xi_{\rm AG}$ was observed when the Lyα column density was increased by orders of magnitude from $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2})) = 13 - 17$.
- Finer binning, with $\Delta\sigma = 0.4$, $\Delta\pi = 1$ showed that the tentative evidence for elongation along the line-of-sight could just be a consequence of increased noise in the column of smallest projected separation.

We compared the Lyα absorber cross-correlation to the galaxy auto-correlation and found that:

- Regardless of along which direction $\xi_{\rm GG}$ was measured, $\xi_{\rm GG}$, for those galaxies with $21.0 \leq R_{\rm AB} \leq 25.0$, in the central bin when plotted in redshift space was significantly larger than $\xi_{\rm AG}$.
- $\xi_{\rm GG}$ did not significantly change depending on the GIMIC region, and was not seen to evolve as the redshift decreased from z = 1 → 0.5.

We also compared $\xi_{\rm AG}$ from the simulation with that from the observational data. This was done by limiting the Lyα absorber column densities to $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2})) = 13 - 15$ and excluding region 5. The upper-limit of $\xi_{\rm AG}$=1.83 from the observed data was 1.3 and 0.8σ lower than the simulated result at redshifts z = 0.5 and 1 respectively. This implies that there was good agreement between the observations and simulations in the low level of Lyα absorber-galaxy correlation.

Chapter 8

The Discussion, Further Work and a Final Word

8.1 Discussion and Conclusions to the 2D 2-Point Cross-Correlation Function

8.1.1 The Initial Question

The original question posed in Section 2.1 was to what extent are galaxies correlated with the neutral gas in the IGM, depending on the separation, redshift, column density and the local environment? Using the results of the 2D 2-point correlation function from the observations, the GIMIC simulation and the literature these dependencies will be discussed in turn.

8.1.2 ξ_{AG} as a Function of the Separation and Column Density

The two straw-man models mentioned in Sections 1.3.3 and 2.1 were either that the gas follows the galaxy distribution and collects in galaxy haloes, or it forms part of a diffuse network, located near but not necessarily correlated with galaxies. Such a black or white picture is clearly naive, because even though ξ_{AG} was observed to peak within 1 Mpc of galaxies in Ryan-Weber (2006), Pierleoni et al. (2008) and the GIMIC simulations, in most cases this peak was not significantly different from ξ_{AG} at further separations along the line of sight and at a larger impact, parameter, or zero.

There were insufficient pairs collected during observations for Figure 5.20 to give a conclusive detection of the ξ_{AG}, however for galaxies and Lyα absorbers that have 13.18 $\leq \log_{10}(N_{HI}(cm^{-2})) \leq$ 17.42, we calculated a 3σ upper-limit of $\xi_{AG}=1.83$ when binned $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{70}^{-1}Mpc. As only 194 absorber-galaxy pairs existed within $\Delta\sigma = 0 - 5$, $\Delta\pi = 0 - 20$ h_{70}^{-1}Mpc no cuts in the column density could be made without further increasing an already large error. Figure 5.5 however showed that the overwhelming majority of Lyα absorbers from the STIS spectra had $\log_{10}(N_{HI}(cm^{-2})) \leq 15$.

The decrease in the 3σ upper-limit to 1.08 in Figure 5.21, caused by the increase in binsize to $\Delta\sigma = \Delta\pi = 2$ h_{70}^{-1}Mpc is further evidence that ξ_{AG} is near to 0 when averaged over these distances. This result was supported by the 1D plots of $\xi_{AG}(\pi)$ and $\Xi_{AG}(\sigma)$ along the line of sight and projected separation in Figures 5.25 and 5.23.

Figures 7.7 and 7.10 that showed ξ_{AG} with $\log_{10}(N_{HI}(cm^{-2})) = 13 - 14$ and $\log_{10}(N_{HI}(cm^{-2})) = 13 - 15$ from the GIMIC simulation were in agreement with this observed upper-limit. From these figures it was apparent that as long as $\log_{10}(N_{HI}(cm^{-2})) \leq 15$ there was a $\approx 3\sigma$ significant correlation between gas and galaxies, with values in the central bin of $\xi_{AG}=1.87\pm0.63$ for $\log_{10}(N_{HI}(cm^{-2})) \leq 14$ and $\xi_{AG}=2.45 \pm 0.77$ for $\log_{10}(N_{HI}(cm^{-2})) = 13 - 15$. The correlation was highest in the central bin and ξ_{AG} was consistent with zero at all separations, except for the bins at a distance $\sigma \leq 1$, $\pi = 8 - 12$ h_{73}^{-1}Mpc in Figure 7.7.

At this separation there were enough pairs that had $\log_{10}(N_{HI}(cm^{-2})) = 13 - 14$ that the anti-correlation in these bins was statistically significant. The bin at $\sigma = 0 - 1$, $\pi = 10 - 12$ h_{73}^{-1}Mpc had a value $\xi_{AG}= -0.59 \pm 0.18$, a decrease of 3.8σ from the peak at the centre. We note that this drop in correlation was also present in our observations, albeit with a much larger error margin, in Figure 5.20, and Figure 2.5 that plots ξ_{AG} for absorbers with $\log_{10}(N_{HI}(cm^{-2})) = 13 - 15$ from Wilman et al. (2007). A drop in correlation at similar separations and for absorbers with the same column density was also seen in Figure 3 of Ryan-Weber (2006) that is shown in Figure 2.2 and Figure 20 from Adelberger et al. (2003). There were insufficient pairs at higher column densities to confirm whether or not this anti-correlation is confined to just low column density Lyα absorbers. It remains unknown whether this is a geometric effect, a consequence of the binned 2D 2-point correlation, or a genuine property of the distribution of Lyα absorbers.

ξ_{AG} from the GIMIC simulation increased by only 0.72σ as the column density rose from $\log_{10}(N_{HI}\,(cm^{-2}))= 13-14$ to $\log_{10}(N_{HI}\,(cm^{-2}))= 15-17$. It is doubtful that such a change in ξ_{AG} in the real data would have been detected had a similar cut in the column density been investigated.

The only difference in the plots of Figure 7.7 as the column density cut increased from $\log_{10}(N_{HI}\,(cm^{-2}))= 14-15$ to $15-17$ was a slight decrease in ξ_{AG} beyond ~ 4 h_{73}^{-1}Mpc along π and σ. This was accompanied by additional shot-noise. There were still too few pairs to adequately rule out an anti-correlation at high separation. At best we can say that the correlation in the central bin increases with the column density, albeit by $< 3\sigma$, and at distances beyond ~ 4 h_{73}^{-1}Mpc, $\xi_{AG}= 0$. What this means is that, even though Lyα absorbers with $\log_{10}(N_{HI}\,(cm^{-2}))\sim 15-17$ are more common near to a galaxy halo, we cannot rule out their existence at further separations, and they may even lie in the voids of the IGM.

This is the same conclusion that was reached by Wilman et al. (2007). In this study the cross-correlation remained low up to $\log_{10}(N_{HI}\,(cm^{-2}))= 17$. For column densities $\log_{10}(N_{HI}\,(cm^{-2}))= 17-19$ at a separation of $\sigma < 0.4$, $\pi < 2$ h_{70}^{-1}Mpc, ξ_{AG} jumped to 15 ± 12. The GIMIC result did not detect a large increase in ξ_{AG}. In the bin $\sigma < 1$, $\pi < 2$ h_{73}^{-1}Mpc, at a redshift of $z = 0.5$, for Lyα absorbers that have $\log_{10}(N_{HI}\,(cm^{-2}))\geq 17$ $\xi_{AG}= 5.03 \pm 3.23$. Hence there is still insufficient evidence from the GIMIC simulation for a very large increase in ξ_{AG} as the column density increases above $\log_{10}(N_{HI}\,(cm^{-2}))= 17$. This may be because the model was unable to accurately simulate Lorentzian profiles and sub-DLA systems. These values are consistent within the margin of error, which is large because both Wilman et al. (2007) and the GIMIC data suffered from an insufficient number of pairs that had $\log_{10}(N_{HI}\,(cm^{-2}))\geq 18$, with only 16 and 17 Ly$\alpha$ absorbers respectively.

The column density at which we could potentially see a significant jump in the correlation could be found by studying Mg II systems. Churchill et al. (2005) reported that Mg II absorbers have been found that are associated with Lyα absorbers that have $\log_{10}(N_{HI}\,(cm^{-2}))= 15.5-20.5$. The weak Mg II absorbers that have

EW(2796) $\leq$ 0.3 Å and are seen in Lyα absorbers with 15.5 $\leq \log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\sim$ 19 have been detected with impact parameters of 30 – 100 kpc from luminous (L > 0.05L*) galaxies (Narayanan et al., 2006). Meanwhile stronger Mg II absorbers with 0.3 $\leq$ EW(2796) $\leq$ 0.6 Å, are nearly always observed with Lyα absorbers that have column densities $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq$ 19 and are located within 40 kpc. Thus Lyman-limit systems can be found close to galaxies, but the possible locations of sub-damped absorbers and DLAs are much closer. Hence we can predict that on a finer grid a very large jump in correlation, potentially much greater than the galaxy auto-correlation, would be observed for Lyα absorbers that had $\log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\geq$ 19.

Contrary to the result of a low cross-correlation, Ryan-Weber (2006) measured $\xi_{AG}\sim$ 11 $\pm$ 0.5 at the central bin at distances $\leq$ 100 kpc, even though the column density of the Lyα absorbers was restricted to 12.41 $\leq \log_{10}(N_{HI}\,(\mathrm{cm}^{-2}))\leq$ 14.81. This increase in ξ_{AG} may be due to a lower redshift and/or a different galaxy population and will be discussed in Section numbers 8.1.3 and 8.1.4 below. Alternatively the value in the central bin could be so high because of the smaller bin size of $\Delta\sigma = \Delta\pi = 0.1$ h_{100}^{-1}Mpc. We could argue that the absorbers that are at $\sigma < 0.1$ h_{100}^{-1}Mpc are within the galaxy halo, so ξ_{AG} is bound to substantially increase.

The different levels of binning used for the GIMIC simulation showed that this can have an effect on the correlation function. When binned $\Delta\sigma = 0.4$, $\Delta\pi = 1$ h_{73}^{-1}Mpc, the central bin in Figure 7.11 that shows ξ_{AG} for all the GIMIC regions at all column densities at a redshift of z = 0.5 had a value of ξ_{AG}= 12.07$\pm$5.99. An increase in the bin size to $\Delta\sigma = 1$, $\Delta\pi = 2$ h_{73}^{-1}Mpc caused ξ_{AG} to decrease 1.3σ to 3.96$\pm$1.21. Hence increases in ξ_{AG} can be expected because of a smaller bin size, but significantly more pairs would be required to decrease the large error that accompanies this finer grid. Pierleoni et al. (2008) having binned $\Delta\sigma = \Delta\pi = 0.4$ h_{72}^{-1}Mpc were able to achieve an error of only $\sigma_{\xi_{AG}}\sim$ 0.2, but they had fired 999 lines of sight and had a catalogue of 4980 galaxies and so were able to correlate many more pairs on a finer grid than what we were able to achieve with the GIMIC simulation.

The "finger-of-god" at small projected separation

The unusually high value in Figure 2.2 from Ryan-Weber (2006) and the large "finger-of-god" could have been caused by cosmic variance from using too few sightlines. The Pierleoni et al. (2008) data came from a single cube with a mean density that was equivalent to the 0σ GIMIC region. They used 999 sightlines and there was no "finger-of-god" in Figure 2.3. When the number of sightlines was reduced to 27, the same as the dataset from Ryan-Weber (2006), Pierleoni et al. (2008) claimed that the increased shot-noise created a higher value for $\xi_{\rm AG}$ in the central bin with evidence for elongation along the line of sight.

Only minor elongation along the line of sight was evident in the GIMIC plots of Figures 7.5 and 7.11. This elongation contained pairs that mostly came from region 5 where large velocity separations between galaxies are expected in the virialised cluster. Pierleoni et al. (2008) found that directing sightlines along a particular cosmic structure, such as a filament, did not produce a large "finger-of-god". Rather we suggest, using the tentative evidence from the GIMIC simulation, a "finger-of-god" can be expected if a sightline is directed towards a virialised region where the galaxies and Lyα absorbers are gravitationally bound. In this case there will also be a significant correlation in the central bin. There will however be a large RMS that is caused by jackknifed sightlines that do not share this virialised environment. That is why any figure that contained evidence for elongation along the line of sight also had a large error in the corresponding bin of the $\sigma_{\xi_{\rm AG}}$ plot. When averaged over many sightlines the variation in $\xi_{\rm AG}$ diminished and the "finger-of-god" disappeared.

No "finger-of-god" was observed in the STIS and FORS2 observations or Wilman et al. (2007), even though fewer sightlines than Ryan-Weber (2006) had been used. This could be because these sightlines did not penetrate highly virialised systems. Or because the large binsize and poor sampling contained insufficient pairs for a "finger-of-god" to be resolved.

8.1.3 ξ_{AG} as a Function of the Redshift

The pie-diagrams in Figures 5.7 and 5.8 show that Lyα absorber-galaxy pairs were concentrated at redshifts z $\sim$ 0.8 and z $\sim$ 1 for the sightlines towards HE 1122-1648 and PKS 1127-145 respectively. The majority of pairs in Wilman et al. (2007) were at z $\leq$ 0.5 (Morris and Jannuzi, 2006). When comparing the peak values in correlation between our observations and Wilman et al. (2007) we saw no significant change with redshift. In Figure 2.5 pairs that had $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 13 - 15$ had a cross-correlation $\xi_{\mathrm{AG}}= 3.0 \pm 2.0$, and the 3σ upperlimit from our observations was 1.83, both of these observations however were dominated by shot noise.

An equivalent comparison in ξ_{AG} was made between GIMIC snapshots at z = 1 and 0.5 with the same conclusion. Figure 7.6 that considered absorbers with a column density $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))\geq 13$ showed an increase in ξ_{AG} of only 0.9σ from 2.65 ± 0.78 to 3.96 ± 1.21 in the central bin. When the column densities were restricted to $\log_{10}(\mathrm{N_{HI}}\,(\mathrm{cm}^{-2}))= 13-15$ and region 5 was removed in order to better simulate the observations, Figure 7.10 showed an increase that was even less significant with values for ξ_{AG} of 2.45 ± 0.77 and 3.37 ± 1.23.

Figure 7.11 that compared ξ_{AG} at z = 1 and 0.5 with a smaller bin size of $\Delta\sigma = 0.4$, $\Delta\pi = 1$ h_{73}^{-1}Mpc also displayed no significant change in ξ_{AG} with redshift for the combined GIMIC data. Values in the central bin were 5.68 ± 3.61 and 12.07 ± 5.99. Further comparisons between the separate GIMIC regions also showed no significant evolution in ξ_{AG} with redshift.

Extrapolating from these findings we suspect that the unusually high correlation from Ryan-Weber (2006) is not because the data was taken at z $\sim$ 0. This suggestion is supported by Pierleoni et al. (2008), who after simulating the cross-correlation at z = 0 found low values for ξ_{AG} that are similar to our results.

The argument from Section 2.1 was that the cross-correlation function ought to increase at lower redshift as matter continues to collapse and gas drains along filaments and falls into galaxy haloes. Gravitational collapse however is not the only process that would have occurred as the redshift decreased and that could have had an impact on the correlation function.

Any evolution in ξ_{AG} with redshift could also be caused by the decrease in the density of neutral gas caused by Hubble expansion. The density of low column density Lyα absorbers would have been rarefied as space expanded, meaning that those absorbers that did have $\log_{10}(N_{HI}(cm^{-2}))\sim 13$ would no longer be detected. Meanwhile those of a high column density, that were Jeans unstable, may have collapsed further causing more discrete absorbers of a higher column density to form (Schaye, 2001).

Changes in the line-density caused by this structure formation would have been offset by the decreasing rate of re-ionisation by the UV background that is shown in Figure 1.1. The decrease in $\Gamma(z)$ meant that more H I gas should be present, implying an increase in the line density. The overall effect of these two processes is the almost flat evolution below $z \sim 1.7$ that was described by Figures 1.6, 1.7, 1.8, and by the data collected from the STIS and UVES Lyα absorbers in Figure 5.6.

Hence absorbers that originally had a very low density may have dispersed. Meanwhile absorbers of a sufficient density may have continued to acrete gas, increasing in column density, and caused ξ_{AG} to increase in the central bin as this gas gravitated towards the nearest galaxy. Despite these processes we found no evidence for a significant increase in ξ_{AG} for low column density absorbers when comparing data from separate epochs.

8.1.4 ξ_{AG} Dependence on the Galaxy Population

One fact highlighted by the plots in Figure 7.1 is that even when the nearest galaxy included in the correlation function was $> 1\ h_{73}^{-1}$Mpc away, significant absorption may still have been detected because it was near to a galaxy that did not make the magnitude cut imposed on the GIMIC sample. This is why a further condition must be made when defining ξ_{AG} because the correlation with galaxies can only be measured down to a minimum galaxy mass or luminosity.

The galaxies in the FORS2 survey were limited in magnitude to $21.5 \leq R_{Vega} \leq 24.5$, and the majority (148/200) were identified using strong [O II] emission. Thus we can only claim that ξ_{AG} is low with respect to these galaxies.

The galaxies that were correlated in the GIMIC simulation were selected on a similar basis to the FORS2 sample, with a magnitude cut $21.0 \leq R_{AB} \leq 25$. Figure

7.2 however shows that in the simulation selecting galaxies by their magnitude was equivalent to a stellar mass cut. It is galaxies with a lower mass ($M_{\rm halo} \leq 10^{11.4}$ $M_{\odot}$) that Kereš et al. (2005) claimed acrete gas directly from the IGM, rather than the classical modes of accretion discussed in Section 2.2.1. This difference could explain why the most pairs were observed in the central bin in all the plots of the simulation, while in both our observations and Figure 2.5 from Wilman et al. (2007) this peak was not significant.

Differences in $\xi_{\rm AG}$ because of the galaxy sample could also explain why Pierleoni et al. (2008) were unable to replicate the observations made by Ryan-Weber (2006). The galaxies used in Ryan-Weber (2006) were from the HIPASS catalogue (Zwaan et al., 2003), a blind survey that targeted gas rich galaxies of any mass by 21 cm radiation. The mean halo mass of the HIPASS sample was 10^{11} $M_{\odot}$. Pierleoni et al. (2008) had a minimum halo mass of 8×10^{10} $M_{\odot}$ that was set by the resolution of the simulation. So as well as having a possible lower mass limit, the galaxies in Ryan-Weber (2006) would have a very high cross-correlation at small separation (on scales of the halo diameter) because they were detected because they contained large quantities of neutral hydrogen. This localised gas may also have been part of a virialised system, or part of a swift outflow, and it is this that could lead to the large "finger-of-god" in Figure 2.2.

By testing different galaxy mass limits in their simulation Pierleoni et al. (2008) detected no significant change in $\xi_{\rm AG}$. This does not contradict our findings, rather Pierleoni et al. (2008), Wilman et al. (2007) and our investigations all agreed that $\xi_{\rm AG}$ is small, regardless of the mass cut. Ryan-Weber (2006) may have only found a higher value because of the selection effect caused by targeting gas rich galaxies.

8.1.5 $\xi_{\rm AG}$ as a Function of the Environment

Any ideas as to how the local conditions effect $\xi_{\rm AG}$ is largely based on previous investigations and tentative evidence. The three subjects considered are; the mean density, investigated by comparing the separate GIMIC regions, differences in $\xi_{\rm AG}$ that could arise by the UV background and temperature, and the effects of galactic winds.

The mean density

Figure 7.5 showed that even though there was a large difference in the number of pairs between regions 1 and 5 of the GIMIC simulation, there was no significant difference in the correlation of the central bin ($\xi_{\rm AG}= 3.55 \pm 3.18$ and 2.65 ± 0.78 for regions 1 and 5 respectively at a redshift $z = 1$). With the most Lyα absorbers that had $\log_{10}({\rm N_{HI}}\,({\rm cm}^{-2})) \geq 17$, a significantly larger correlation had been expected in region 5.

The only difference between the regions in Figure 7.5 was in the $+2\sigma$ sphere, where correlation along the line of sight was higher because of the virialised cluster and this caused a slight "finger-of-god". In the bin $\sigma = 0 - 1$, $\pi = 2 - 4$ h_{73}^{-1}Mpc at a redshift $z = 1$, $\xi_{\rm AG}$ in region 4 was 0.23 ± 0.39. While in region 5 $\xi_{\rm AG}$ was 3.2σ higher at 2.05 ± 0.41, a value equivalent to the correlation at the centre. Thus within a gravitationally bound system, like that described by region 5, the accretion of gas occurs at a slightly larger rate than the more empty regions of space. The "finger-of-god" is much less significant however than that seen in the galaxy auto-correlation function that is described below in Section 8.1.6.

The uniform background UV intensity and temperature

What was missing from the GIMIC simulation, which would show a significantly different result to from that found were the Mg II and Fe II bearing DLA systems that have $\log_{10}({\rm N_{HI}}\,({\rm cm}^{-2})) \geq 20.3$. These systems were absent because of the uniform UV background which did not allow optically thick absorbers.

Had absorptions with large Lorentzian wings existed in the spectra, the number of pairs would also have been lower, because the low column density systems that show no significant correlation would have remained undetected in the damping wings, particularly when using mock spectra that had a low signal-to-noise (S/N).

Figure 7.1 shows that no significant absorption was detected in lines of sight along the z direction in the centre of region 5, even though one might expect to find high column density absorbers in these areas. In fact along all three directions at redshifts $z = 1$ and 0.5 only 3 absorbers were detected in the 6 spectra that penetrated the centre of the cluster, and all these had a column density $\log_{10}({\rm N_{HI}}\,({\rm cm}^{-2})) <$

14. It was possible that the S/N was so low that many absorption systems were neglected, however manual inspection of these spectra at a S/N $\gg 2$ showed that this was not the case. Very few Lyα absorbers were present and definitely none that had $\log_{10}(N_{HI}\,(cm^{-2}))\geq 17$. Thus either GIMIC was unable to accurately simulate the high column density absorbers that are close to galaxies in region 5. Or the temperature was so great in the centre of the cluster that Lyα absorption was no longer significant. To detect gas in these regions would require O VI absorbers that trace the WHIM, or the x-ray transitions O VII and O VIII that occur when the temperature is $T \geq 10^7$ K. These conditions are expected in the intra-cluster medium when the in-falling gas is shock-heated to the virial temperature.

The effects of a galactic wind

Section 2.2.2 explained how winds play an important role in the processes that occur on the "boundary" between the IGM and a galactic halo. Thus in order to notice any effect winds have on the cross-correlation function and whether neutral gas is expelled, we would need a much finer level of binning than $\Delta\sigma = 1$, $\Delta\pi = 2$ Mpc. With only a single wind model in the GIMIC simulation these theories could not be tested, and the decrease in bin size to $\Delta\sigma = 0.4$, $\Delta\pi = 1$ h_{73}^{-1}Mpc did not produce a lower value in ξ_{AG}.

What evidence there is for the effects of winds came from Adelberger et al. (2005, 2003) and in simulations from Pierleoni et al. (2008). In Adelberger et al. (2005) the correlation function compared Lyman-break-galaxies that expel super-winds with Lyα absorbers at redshifts $z \sim 3$. Only marginal evidence was detected that winds have an effect on sub-Mpc scales, and even then this was in only 1/3 of the galaxies studied.

A similar non-result was detected in the simulation. Pierleoni et al. (2008) measured ξ_{AG} using 3 winds models, no winds, a strong wind and an extremely strong wind, where both winds travelled at 484 km s^{-1} and reached distances of 20 and 60 kpc from the galaxy respectively. No significant difference in ξ_{AG} was detected when comparing the wind models. There was only a minor decrease in the cross-correlation at distances ≤ 300 kpc for the extremely strong wind.

Thus the speculation that arose in Wilman et al. (2007) and our data in Figure 5.9, that the reason for the peak in cross-correlation being off-set for low-column density absorbers because of galactic winds is probably wrong. The off-set peak was a symptom of the shot-noise.

8.1.6 Comparisons of $\xi_{\rm AG}$ with the Galaxy Auto-Correlation Function

Using our observations, the GIMIC simulation and results from the literature we were able to show that within a distance $\sim$ 4 Mpc galaxies are more strongly clustered with one another, rather than with the low column density Lyα absorbers.

Comparing the central bin in Figure 5.17 with the 3σ upper-limit of the cross-correlation we found $\xi_{\rm AG}= 1.83$ and $\xi_{\rm GG}= 3.89\pm0.65$. $\xi_{\rm GG}$ was much lower than had been expected, especially when compared to the galaxy auto-correlation measured by Hawkins et al. (2003) and Guzzo et al. (2008), that had values in the central bin of $\xi_{\rm GG}= 17.4 \pm 3$ and $\xi_{\rm GG}\sim 10$ respectively. With a catalogue of only 200 galaxies that produced 266 galaxy pairs within $\sigma \leq 5$, $\pi \leq 20$ $\rm h_{70}^{-1}Mpc$ it is not surprising that we have not been able to reproduce these literature results. When we compared our value for the cross-correlation with $\xi_{\rm GG}$ from the literature we found a $\gg 5\sigma$ difference between $\xi_{\rm AG}$ and $\xi_{\rm GG}$.

In all figures that plotted $\xi_{\rm GG}$ beyond a distance of $\sim$ 4 Mpc from the centre, in bins not effected by the "finger-of-god", we found that $\xi_{\rm GG}$ was effectively 0 and thus consistent with $\xi_{\rm AG}$. Hence the galaxy auto-clustering was only dominant within small separations. This might be because the galaxies that were of a sufficient magnitude ($21.0 \leq \rm R \leq 25.0$) were concentrated within the filaments and nodes of the cosmic web. This was illustrated for those galaxies in the simulation by Figure 7.1. Beyond these dense regions resided galaxies that did not make the magnitude cut.

The GIMIC simulation and literature results of the auto-correlation also showed that the velocity difference between galaxies along the line of sight was much greater than the velocity difference between pairs of galaxies and absorbers. Figures 5.16, 5.19 and 7.8 all showed significant "fingers-of-god" along the line of sight. This suggests that within filaments, particularly in regions described by the $+2\sigma$ region, the extent of virialisation between galaxies is much greater than that between galaxies and low column density absorbers. This adds further evidence that low column density absorbers that constitute the filamentary cosmic web exist as a separate entity within which galaxies are embedded, and that this quiescent gas is dominated by Hubble flow rather than accretion into a particular galactic halo.

Despite this consensus, comparisons along different directions and epochs did show variations in $\xi_{\rm GG} \geq 1\sigma$. $\xi_{\rm GG}$ is expected to increase at lower redshifts because of gravitational collapse. This explains to some extent the difference between $\xi_{\rm GG}$ from Hawkins et al. (2003), that studied galaxies from the 2dFGRS at redshifts $z \leq 0.2$, and Guzzo et al. (2008) that used galaxies from the VVDS survey at $z \sim 0.8$. Further differences would also arise because of the cosmic variance between the regions surveyed and the galaxy populations. Galaxies obtained in the 2dFGRS had a magnitude limit $b_J = 19.45$. Those from the VVDS had a magnitude $B_{AB} \leq 25.7$ within a 3″ aperture.

$\xi_{\rm GG}$ did not increase in the GIMIC simulation as the redshift decreased from $z = 1 \rightarrow 0.5$, nor was there a significant difference in $\xi_{\rm GG}$ between the different GIMIC regions. This could be because of limitations with the data. Rather than an insufficient number of pairs the problem this time was the volume. Small scale variations in the galaxy distribution were significant because we correlated pairs of galaxies that were all contained within a sphere of radius 18 h_{73}^{-1}Mpc. This was evident by differences in the auto-correlation when $\xi_{\rm GG}$ was measured along different directions in Figure 7.9. Along the z-direction $\xi_{\rm GG}$ in the combined GIMIC regions was 27.05 ± 4.06 at a redshift of $z = 1$. This decreased to 22.44 ± 6.02 and 16.02 ± 5.09 when observed along the x and y directions respectively.

8.2 Further Work

8.2.1 The Three Archive STIS Sightlines

The main drawback with our observations was the shot noise that was caused by too few pairs that had distances along the projected separation and line of sight of

$\sigma \leq 5$ and $\pi \leq 20\ h_{70}^{-1}$Mpc. There were three problems that created this deficit.

Firstly, and this was unavoidable, was the poor weather conditions that meant spectra for the vast majority of galaxies were collected within 1 night. Hence we only had a catalogue of 200 galaxies between 2 sightlines.

Secondly, the archive STIS E230M spectra were of low S/N, therefore many low column density systems remained undetected and we only managed a minimum Lyα EW limit of 0.14 Å for absorbers that had an EW significance ≥ 3. For comparison the Lyα absorbers from the UVES spectrum of HE 1122-1648 had an EW limit of 0.01 Å.

Thirdly, the pie-diagrams in Figures 5.7 and 5.8 show how the problem of a small galaxy catalogue and Lyα line-list was compounded by the small overlap in the redshift range of galaxies that was covered by the redshifts of the absorbers. Although the method of targeting galaxies in the magnitude limited sample $21.5 \leq R_{\mathrm{Vega}} \leq 24.5$ did mean most of the galaxies were at a redshift $z \geq 0.68$, many MOS slits were wasted on foreground galaxies.

That is why further observations have been planned that will correlate galaxies collected in a VIMOS (LeFevre et al., 2003) survey in three fields of view, with Lyα absorbers from STIS E230M spectra from the STScI archive.

These three quasar spectra were selected because they had the highest S/N available (a minimum of 4), with reasonable coverage at wavelengths equivalent to Lyα absorption at a redshift $z \sim 1$, 2300-3100 Å. The quasars are:

- PKS 0232-042 ($z = 1.438$), with a spectrum that contains 128 Lyα absorbers between 2280-2941 Å (Janknecht et al., 2006).
- The Hubble Deep Field South (HDF-S) quasar J2233-6033 ($z = 2.238$) (Savaglio et al., 1999). The completed line-list, that runs up to optical wavelengths with additional lines from the UCLES spectrum from the Anglo-Australian Telescope (Outram et al., 1999), has 66 Lyα absorbers between 2675-3050 Å. A Lyman-limit system prevents further study below these wavelengths.
- HE 1104-1805A ($z = 2.319$) (HST Proposal 8471; Jenkins).

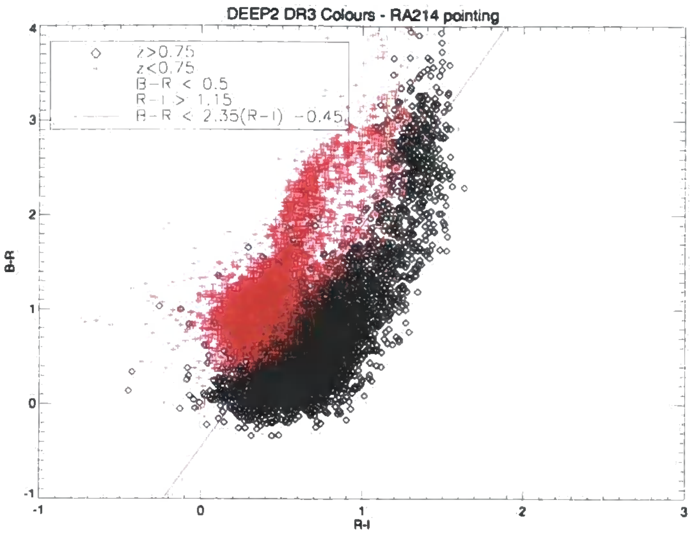


Figure 8.1 No colour cut was used for those galaxies from the Groth Strip in the DEEP2 Redshift Survey. Hence all galaxies, regardless of colour and redshift are present. This way the necessary colour cuts in order to target galaxies with z ≥ 0.75 could be calculated.

8.2.2 The VIMOS Proposal

So that we can maximise the number of galaxies within the required redshift range we requested BRI colour imaging in the surrounding $16' \times 14'$ field of view of each quasar. Potential galaxies for follow-up spectroscopy were then selected based on the colour cut that was adopted in the DEEP2 Redshift Survey (Coil et al., 2004). The DEEP2 survey targeted ~ 50000 galaxies at $0.75 < z < 1.4$ with the DEIMOS spectrograph (Davis et al., 2003b, 2007). These had been preselected using BRI photometry, and were from four $120' \times 30'$ fields studied down to $I \leq 24.5$. No colour cut was used in the first of these fields, the Groth Strip (14 17 00, +52 30 00), in order that the necessary colours could be decided.

Figure 8.1 shows the $B - R$, $R - I$ colours of all galaxies from the Groth Strip from the DEEP2 Data release 3[1]. By applying the following colour cuts those galaxies with a redshift $z \geq 0.75$ were selected.

- $B - R \leq 0.5$
- $R - I \geq 1.15$

---

[1] Data supplied courtesy of the DEEP2 collaboration at UC Berkeley, CA 94720, USA, UC Santa Cruz, CA 95064, USA and the University of Hawaii, HI 96822-1839, USA.

- $B - R \leq 2.35(R - I) - 0.45$

Image reduction

VIMOS BRI imaging[2] was requested in service mode in the fields of view of the three quasars, and reduction of these images was carried out using the instrument pipelines at Garching, and locally using IRAF and IDL programs.

The VIMOS instrument was selected because the low-resolution ($R = 210$) mode allows for multi-object-spectroscopy that can target hundreds of galaxies simultaneously, therefore vastly increasing the number of galaxies with which to correlate with absorbers. By using VIMOS we could also observe galaxies that are at a much wider field of view than the limit of FORS2 which limited us to $\sigma \sim 4\ h_{70}^{-1}$Mpc at redshifts of $z \geq 1$. This way more foreground galaxies can be included that are not restricted by the maximum possible impact parameter of the FORS2 instrument. This was illustrated in Figure 5.2.

Source Extractor was run on each R band image to identify those sources with a magnitude $R_{\mathrm{Vega}} \leq 25.0$. Using circular apertures with a diameter $2''$ the flux through each of these targets was then measured in the B, R and I images. Those sources that the neural network gave a CLASS STAR value ≥ 0.8 in the R band were assumed to be stars and were flagged so that they would not be passed to the list of candidate galaxies for spectroscopy. The necessary steps when reducing imaging data and using Source Extractor are described in Section 3.2.

The Colour Cut and Galaxy Catalogue Generation

$B - R$, $R - I$ colours for each of the 4 chips installed on the VIMOS instrument were then plotted in the same manner as Figure 8.1. All of the targets from chip 2 along the line of sight towards HE 1104-1805A are plotted in Figure 8.2. Those galaxies that were in the defined region were then sent to the software VMMPS that is used to design masks for the VIMOS instrument.

---

[2] Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. ESO programme number 381.A-0874.

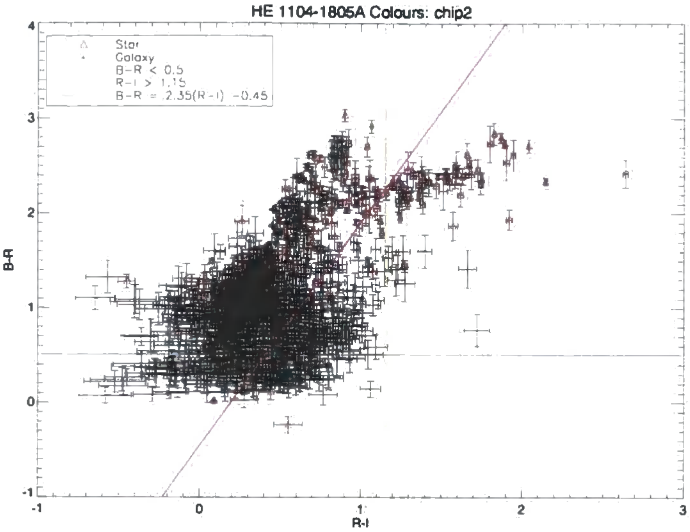


Figure 8.2 This shows the B − R, R − I colours for all of the targets from chip 2 in the field of HE 1104-1805A. Using the same colour cuts as those described in Coil et al. (2004), those galaxies with a predicted redshift $z \geq 0.75$ were selected for spectroscopy. The red triangles represent the main-sequence of stars, these are flagged so as not to be included in the final galaxy catalogue that is passed on to VMMPS to make masks.

Masks for all 4 VIMOS chips that contain a total of 668 and 527 slits for sightlines PKS 0232-042 and HE 1104-1805A respectively were then made. At the current time of writing spectroscopy has been carried out on these 2 fields[3], and spectroscopy is due to be carried out in service mode on those galaxies that surround J2233-6033 in June 2009.

8.3 A Final Word

Our investigations into the relationship between baryons in the IGM and galaxies using the 2D 2-point correlation function has shown that this relationship is far from simple.

While the cross-correlation may not depend on the epoch as $z = 1 \rightarrow 0.5$, ξ_{AG} did seem to depend on the column density, the distance from the galaxy when plotted in redshift space, and the local environment. We conclude:

---

[3] Observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. ESO programme numbers 082.A-0812 and 083.A-0940.

- $\xi_{\rm AG}$ regardless of the column density or redshift peaks at the smallest separation. Marginal evidence, when we reduced the size of the bin in our simulations, showed that $\xi_{\rm AG}$ could be significantly larger if the bin size were reduced to scales that are comparable in size to a galaxy halo, $\Delta\sigma < 0.4$ Mpc.
- Significant increases in $\xi_{\rm AG}$ are expected only if the column density is at least $\log_{10}(N_{\rm HI}\,({\rm cm}^{-2}))\gtrsim 18$. The column density in our data and simulations increased by orders of magnitude from $13 \leq \log_{10}(N_{\rm HI}\,({\rm cm}^{-2})) \leq 17$ yet no significant change in $\xi_{\rm AG}$ was seen.
- Elongation in $\xi_{\rm AG}$ along the line of sight that produced a "finger-of-god" only appeared when the quasar lines of sight passed through a heavily virialised region, such as a large galaxy cluster. Even then this elongation was only marginal and on much smaller scales than that observed in the galaxy autocorrelation.

These experiments showed that $\xi_{\rm AG}$ was very low, peaking in the central bin, but only marginally greater than bins at further separation which have a cross-correlation equivalent to 0. These results suggest that the low column density, diffuse gas is largely quiescent, and has a distribution in the IGM that is uniform on large scales, with perturbations that trace the cosmic web. This result should be compared to the galaxies, which in our sample were embedded in this filamentary structure (at least in the GIMIC simulation). These galaxies have a strong autocorrelation and exhibit large anisotropies on small scales.

Clearly more work and particularly more data is required before more stringent constraints on these limits can be made and questions can be answered. For example it remains unknown why in Chen et al. (2005) $\xi_{\rm AG}$ was greater when absorbers were correlated with galaxies that were dominated by emission lines. Could it be that the absorption dominated galaxies, such as luminous-red-galaxies, have a higher halo temperature, thus less neutral H I is present? The galactic winds that distribute metals and gas from emission line dominated, star-forming galaxies cannot be the full explanation, as so far winds have been seen to have a negligible effect.

The further study of hydrogen and metal absorption lines in quasar spectra is what is needed, in addition to more galaxy surveys in the fields of view of these quasars. These investigations would need to include lines such as; Mg II, that mark the periphery of a galactic halo, C IV, that is expelled by galactic winds, O VI, that traces the WHIM, and the O VII and O VIII X-ray transitions that occur at the high temperatures found in the intra-cluster medium and galactic haloes.

With the use of; ground-based telescopes, the repaired STIS and installed COS instruments on HST, the future James Webb Space telescope, the X-ray observatories XMM-Newton, Chandra and the planned XEUS satellite these questions will be answered as spectrographs probe the H I and metal absorbers of the inter-galactic medium.

Bibliography

K. L. Adelberger, A. E. Shapley, C. C. Steidel, M. Pettini, D. K. Erb, and N. A. Reddy. The Connection between Galaxies and Intergalactic Absorption Lines at Redshift $2\lesssim$ z $\lesssim 3$. **ApJ**, 629:636–653, Aug. 2005.

K. L. Adelberger, C. C. Steidel, A. E. Shapley, and M. Pettini. Galaxies and Intergalactic Matter at Redshift z˜3: Overview. **ApJ**, 584:45–75, Feb. 2003.

A. Aguirre, C. Dow-Hygelund, J. Schaye, and T. Theuns. Metallicity of the Intergalactic Medium using Pixel Statistics. IV. Oxygen. **ApJ**, 689:851–864, Dec. 2008.

A. Aguirre, J. Schaye, L. Hernquist, S. Kay, V. Springel, and T. Theuns. Confronting Cosmological Simulations with Observations of Intergalactic Metals. **ApJL**, 620: L13–L17, Feb. 2005.

I. Appenzeller, K. Fricke, W. Fürtig, W. Gässler, R. Häfner, R. Harke, H.-J. Hess, W. Hummel, P. Jürgens, R.-P. Kudritzki, K.-H. Mantel, W. Meisl, B. Muschielok, H. Nicklas, G. Rupprecht, W. Seifert, O. Stahl, T. Szeifert, and K. Tarantik. Successful commissioning of FORS1 - the first optical instrument on the VLT. *The Messenger*, 94:1–6, 1998.

M. Asplund, N. Grevesse, and A. J. Sauval. The Solar Chemical Composition. In T. G. Barnes, III and F. N. Bash, editors, *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, volume 336 of *Astronomical Society of the Pacific Conference Series*, page 25, Sept. 2005.

J. Bachall and E. Salpeter. On the Interaction of Radiation from Distant Sources with the Intervening Medium. **ApJ**, 142:1677–1680, 1965.

S. Bajtlik, R. C. Duncan, and J. P. Ostriker. Quasar ionization of Lyman-alpha clouds - The proximity effect, a probe of the ultraviolet background at high redshift. **ApJ**, 327:570–583, Apr. 1988.

R. Barkana and A. Loeb. The physics and early history of the intergalactic medium. *Reports on Progress in Physics*, 70:627–657, Apr. 2007.

J. Bechtold. The Lyman-Alpha forest near 34 quasi-stellar objects with Z greater than 2.6. **ApJS**, 91:1–78, Mar. 1994.

J. Bechtold, A. Dobrzycki, B. Wilden, M. Morita, J. Scott, D. Dobrzycka, K.-V. Tran, and T. L. Aldcroft. A Uniform Analysis of the Lyα Forest at z = 0-5. III. Hubble Space Telescope Faint Object Spectrograph Spectral Atlas. **ApJS**, 140: 143–238, June 2002.

E. Bertin. *SExtractor v2.5 User's Manual.* Institut d'Astrophysique & Observatoire de Paris, 2006.

E. Bertin and S. Arnouts. SExtractor: Software for source extraction. **A&AS**, 117: 393–404, June 1996.

S. Bertone and S. D. M. White. How do galactic winds affect the Lyα forest? **MNRAS**, 367:247–258, Mar. 2006.

P. Bevington and D. Robinson. *Data Reduction and Error Analysis for the Physical Sciences; 3rd Edition.* ©McGraw-Hill Science/Engineering/Math, 2002.

J. N. Bregman. The Search for the Missing Baryons at Low Redshift. **Annu. Rev. Astron. Astrophys**, 45:221–259, Sept. 2007.

V. Bromm and R. B. Larson. The First Stars. **Annu. Rev. Astron. Astrophys**, 42:79–118, Sept. 2004.

G. Bruzual and S. Charlot. Stellar population synthesis at the resolution of 2003. **MNRAS**, 344:1000–1028, Oct. 2003.

J. A. Burke. Mass flow from stellar systems-I. Radial flow from spherical systems. **MNRAS**, 140:241, 1968.

B. Carswell, J. Schaye, and T.-S. Kim. The Enrichment History of the Intergalactic Medium: O VI in Lyα Forest Systems at Redshift z$\sim$ 2. **ApJ**, 578:43–59, Oct. 2002a.

B. Carswell, J. Schaye, and T.-S. Kim. The Enrichment History of the Intergalactic Medium: O VI in Lyα Forest Systems at Redshift z$\sim$ 2. **ApJ**, 578:43–59, Oct. 2002b.

R. F. Carswell, K. M. Lanzetta, H. C. Parnell, and J. K. Webb. High-resolution spectroscopy of Q1100 - 264 again. **ApJ**, 371:36–48, Apr. 1991.

R. F. Carswell, D. C. Morton, M. G. Smith, A. N. Stockton, D. A. Turnshek, and R. J. Weymann. The absorption line profiles in Q1101-264. **ApJ**, 278:486–498, Mar. 1984.

R. Cen and J. P. Ostriker. Where Are the Baryons? **ApJ**, 514:1–6, Mar. 1999.

G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. **PASP**, 115: 763–795, July 2003.

H.-W. Chen, K. M. Lanzetta, J. K. Webb, and X. Barcons. The Gaseous Extent of Galaxies and the Origin of Ly alpha Absorption Systems. III. Hubble Space Telescope Imaging of Ly alpha -absorbing Galaxies at Z $<$ 1. **ApJ**, 498:77, May 1998.

H.-W. Chen, J. X. Prochaska, B. J. Weiner, J. S. Mulchaey, and G. M. Williger. Probing the Intergalactic Medium-Galaxy Connection toward PKS 0405-123. II. A Cross-Correlation Study of Lyα Absorbers and Galaxies at z<0.5. **ApJL**, 629: L25–L28, Aug. 2005.

C. Churchill. *QSO Absorption Lines Studies: Ultraviolet and Optical Spectroscopy.* ©Churchill 2007, to submit to Cambridge University Press, 2007.

C. Churchill, C. Steidel, and G. Kacprzak. Mg II Absorbing Galaxies: Morphologies and Kinematics. In R. Braun, editor, *Extra-Planar Gas*, volume 331 of *Astronomical Society of the Pacific Conference Series*, page 387, June 2005.

C. W. Churchill, R. R. Mellon, J. C. Charlton, B. T. Jannuzi, S. Kirhakos, C. C. Steidel, and D. P. Schneider. Low- and High-Ionization Absorption Properties of Mg II Absorption-selected Galaxies at Intermediate Redshifts. II. Taxonomy, Kinematics, and Galaxies. **ApJ**, 543:577–598, Nov. 2000.

J. G. Cohen, D. W. Hogg, R. Blandford, L. L. Cowie, E. Hu, A. Songaila, P. Shopbell, and K. Richberg. Caltech Faint Galaxy Redshift Survey. X. A Redshift Survey in the Region of the Hubble Deep Field North. **ApJ**, 538:29–52, July 2000.

A. L. Coil, B. F. Gerke, J. A. Newman, C.-P. Ma, R. Yan, M. C. Cooper, M. Davis, S. M. Faber, P. Guhathakurta, and D. C. Koo. The DEEP2 Galaxy Redshift Survey: Clustering of Groups and Group Galaxies at z$\sim$ 1. **ApJ**, 638:668–685, Feb. 2006.

A. L. Coil, J. F. Hennawi, J. A. Newman, M. C. Cooper, and M. Davis. The DEEP2 Galaxy Redshift Survey: Clustering of Quasars and Galaxies at z = 1. **ApJ**, 654: 115–124, Jan. 2007.

A. L. Coil, J. A. Newman, N. Kaiser, M. Davis, C.-P. Ma, D. D. Kocevski, and D. C. Koo. Evolution and Color Dependence of the Galaxy Angular Correlation Function: 350,000 Galaxies in 5 Square Degrees. **ApJ**, 617:765–781, Dec. 2004.

S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk. Hierarchical galaxy formation. **MNRAS**, 319:168–204, Nov. 2000.

M. Colless. First results from the 2dF Galaxy Redshift Survey. In G. Efstathiou and et al., editors, *Large-Scale Structure in the Universe*, page 105, 1999.

R. A. Crain, T. Theuns, C. Dalla Vecchia, V. R. Eke, C. S. Frenk, A. Jenkins, S. T. Kay, J. A. Peacock, F. R. Pearce, J. Schaye, V. Springel, P. A. Thomas, S. D. M. White, and R. P. C. Wiersma. Galaxies-Intergalactic Medium Interaction Calculation –I. Galaxy formation as a function of large-scale environment. *ArXiv e-prints*, June 2009.

D. Cristóbal-Hornillos, M. Balcells, M. Prieto, R. Guzmán, J. Gallego, N. Cardiel, Á. Serrano, and R. Pelló. K_s Number Counts in the Groth and Coppi Fields. **ApJ**, 595:71–84, Sept. 2003.

C. Dalla Vecchia and J. Schaye. Simulating galactic outflows with kinetic supernova feedback. **MNRAS**, 387:1431–1444, July 2008.

A. Dall'Aglio, L. Wisotzki, and G. Worseck. An unbiased measurement of the UV background and its evolution via the proximity effect in quasar spectra. **A&A**, 491:465–481, Nov. 2008.

C. W. Danforth and J. M. Shull. The Low-z Intergalactic Medium. III. H I and Metal Absorbers at z < 0.4. **ApJ**, 679:194–219, May 2008.

R. Davé, L. Hernquist, N. Katz, and D. H. Weinberg. The Low-Redshift Lyα Forest in Cold Dark Matter Cosmologies. **ApJ**, 511:521–545, Feb. 1999.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. **ApJ**, 292:371–394, May 1985.

M. Davis, S. M. Faber, J. Newman, A. C. Phillips, R. S. Ellis, C. C. Steidel, C. Conselice, A. L. Coil, D. P. Finkbeiner, D. C. Koo, P. Guhathakurta, B. Weiner, R. Schiavon, C. Willmer, N. Kaiser, G. A. Luppino, G. Wirth, A. Connolly, P. Eisenhardt, M. Cooper, and B. Gerke. Science Objectives and Early Results of the DEEP2 Redshift Survey. In P. Guhathakurta, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4834 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 161–172, Feb. 2003a.

M. Davis, S. M. Faber, J. Newman, A. C. Phillips, R. S. Ellis, C. C. Steidel, C. Conselice, A. L. Coil, D. P. Finkbeiner, D. C. Koo, P. Guhathakurta, B. Weiner, R. Schiavon, C. Willmer, N. Kaiser, G. A. Luppino, G. Wirth, A. Connolly, P. Eisenhardt, M. Cooper, and B. Gerke. Science Objectives and Early Results of the DEEP2 Redshift Survey. In P. Guhathakurta, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4834 of *Society of*

Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pages 161–172, Feb. 2003b.

M. Davis, P. Guhathakurta, N. P. Konidaris, J. A. Newman, M. L. N. Ashby, A. D. Biggs, P. Barmby, K. Bundy, S. C. Chapman, A. L. Coil, C. J. Conselice, M. C. Cooper, D. J. Croton, P. R. M. Eisenhardt, R. S. Ellis, S. M. Faber, T. Fang, G. G. Fazio, A. Georgakakis, B. F. Gerke, W. M. Goss, S. Gwyn, J. Harker, A. M. Hopkins, J.-S. Huang, R. J. Ivison, S. A. Kassin, E. N. Kirby, A. M. Koekemoer, D. C. Koo, E. S. Laird, E. Le Floc'h, L. Lin, J. M. Lotz, P. J. Marshall, D. C. Martin, A. J. Metevier, L. A. Moustakas, K. Nandra, K. G. Noeske, C. Papovich, A. C. Phillips, R. M. Rich, G. H. Rieke, D. Rigopoulou, S. Salim, D. Schiminovich, L. Simard, I. Smail, T. A. Small, B. J. Weiner, C. N. A. Willmer, S. P. Willner, G. Wilson, E. L. Wright, and R. Yan. The All-Wavelength Extended Groth Strip International Survey (AEGIS) Data Sets. **ApJL**, 660:L1–L6, May 2007.

M. Davis and P. J. E. Peebles. A survey of galaxy redshifts. V - The two-point position and velocity correlations. **ApJ**, 267:465–482, Apr. 1983.

J.-M. Deharveng, V. Buat, and J. Bergeron. Lyman α emission from galaxies: the case of absorption line selected galaxies at z$\sim 0.3-0.4$. **A&A**, 298:57, June 1995.

H. Dekker, S. D'Odorico, A. Kaufer, B. Delabre, and H. Kotzlowski. Design, construction, and performance of UVES, the echelle spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory. In M. Iye and A. F. Moorwood, editors, *Proc. SPIE Vol. 4008, p. 534-545, Optical and IR Telescope Instrumentation and Detectors, Masanori Iye; Alan F. Moorwood; Eds.*, volume 4008 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 534–545, Aug. 2000.

A. Dobrzycki, J. Bechtold, J. Scott, and M. Morita. A Uniform Analysis of the Lyα Forest at z = 0-5. IV. The Clustering and Evolution of Clouds at z$\leq$1.7. **ApJ**, 571:654–664, June 2002.

K. Dolag, S. Borgani, G. Murante, and V. Springel. Substructures in hydrodynamical cluster simulations. *ArXiv e-prints: 0808.3401*, Aug. 2008.

J. Dyson and D. Williams. *The Physics of the Inter Stellar Medium: 2nd Edition.* ©Institute of Physics Publishing, 1997.

S. M. Fall and G. Efstathiou. Formation and rotation of disc galaxies with haloes. **MNRAS**, 193:189–206, Oct. 1980.

T. Fang, K. R. Sembach, and C. R. Canizares. Chandra Detection of Local O VII Heα Absorption along the Sight Line toward 3C 273. **ApJL**, 586:L49–L52, Mar. 2003.

G. J. Ferland, K. T. Korista, D. A. Verner, J. W. Ferguson, J. B. Kingdon, and E. M. Verner. CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. **PASP**, 110:761–778, July 1998.

A. J. Fox, C. Ledoux, P. Petitjean, and R. Srianand. C IV absorption in damped and sub-damped Lyman-α systems. Correlations with metallicity and implications for galactic winds at z $\sim$ 2-3. **A&A**, 473:791–803, Oct. 2007.

A. Friedman. Über die Krümmung des Raumes. *Zeitschrift für Physik A Hadrons and Nuclei*, 10:377–386, Dec. 1922.

A. Goobar, S. Hannestad, E. Mörtsell, and H. Tu. The neutrino mass bound from WMAP 3 year data, the baryon acoustic peak, the SNLS supernovae and the Lyman-α forest. *Journal of Cosmology and Astro-Particle Physics*, 6:19, June 2006.

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

A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23(2):347–356, Jan 1981.

L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, O. Le Fèvre, C. Marinoni, B. Garilli, J. Blaizot, G. De Lucia, A. Pollo, H. J. McCracken, D. Bottini, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, L. Tresse, G. Vettolani, A. Zanichelli, C. Adami, S. Arnouts, S. Bardelli, M. Bolzonella,

A. Bongiorno, A. Cappi, S. Charlot, P. Ciliegi, T. Contini, O. Cucciati, S. de la Torre, K. Dolag, S. Foucaud, P. Franzetti, I. Gavignaud, O. Ilbert, A. Iovino, F. Lamareille, B. Marano, A. Mazure, P. Memeo, R. Merighi, L. Moscardini, S. Paltani, R. Pellò, E. Perez-Montero, L. Pozzetti, M. Radovich, D. Vergani, G. Zamorani, and E. Zucca. A test of the nature of cosmic acceleration using galaxy redshift distortions. **Nature**, 451:541–544, Jan. 2008.

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

F. Haardt and P. Madau. Modelling the UV/X-ray cosmic background with CUBA. In D. M. Neumann and J. T. V. Tran, editors, *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, 2001.

M. Hamuy, N. B. Suntzeff, S. R. Heathcote, A. R. Walker, P. Gigoux, and M. M. Phillips. Southern spectrophotometric standards, 2. **PASP**, 106:566–589, June 1994.

E. Hawkins, S. Maddox, S. Cole, O. Lahav, D. S. Madgwick, P. Norberg, J. A. Peacock, I. K. Baldry, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, B. Jones, I. Lewis, S. Lumsden, W. Percival, B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe. **MNRAS**, 346:78–96, Nov. 2003.

G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Data Processing, Sky Maps, and Basic Results. **ApJS**, 180:225–245, Feb. 2009.

D. W. Hogg. Distance measures in cosmology. *ArXiv Astrophysics e-prints; arXiv:astro-ph/9905116*, May 1999.

D. W. Hogg, M. A. Pahre, J. K. McCarthy, J. G. Cohen, R. Blandford, I. Smail, and B. T. Soifer. Counts and colours of faint galaxies in the U and R bands. **MNRAS**, 288:404–410, June 1997.

A. M. Hopkins. On the Evolution of Star-forming Galaxies. **ApJ**, 615:209–221, Nov. 2004.

K. Horne. An optimal extraction algorithm for CCD spectroscopy. **PASP**, 98: 609–617, June 1986.

E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. **PNAS**, 15:168–173, Mar. 1929.

E. Janknecht, D. Reimers, S. Lopez, and D. Tytler. The evolution of Lyman α absorbers in the redshift range $0.5 < z < 1.9$. **A&A**, 458:427–439, Nov. 2006.

E. Jehin and K. O'Brien. *FORS User Manual, Doc. No. VLT-MAN-ESO-13100-1543.* ESO, 2007.

C. L. Joseph, V. S. Argabright, J. Abraham, D. Dieball, S. Franka, M. Styonavich, C. van Houten, A. Danks, and B. E. Woodgate. Performance results of the STIS flight MAMA detectors. In C. B. Johnson and E. J. Fenyves, editors, *Proc. SPIE Vol. 2551, p. 248-259, Photoelectronic Detectors, Cameras, and Systems, C. Bruce Johnson; Ervin J. Fenyves; Eds.*, volume 2551 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 248–259, Sept. 1995.

G. G. Kacprzak, C. W. Churchill, C. C. Steidel, M. T. Murphy, and J. L. Evans. A Correlation between Galaxy Morphology and Mg II Halo Absorption Strength. **ApJ**, 662:909–922, June 2007.

A. Kaufer, S. D'Odorico, L. Kaper, C. Ledoux, and G. James. *UV-Visual Echelle User Manual, Doc. No. VLT-MAN-ESO-13200-1825.* ESO, 2007.

D. D. Kelson. Optimal Techniques in Two-dimensional Spectroscopy: Background Subtraction for the 21st Century. **PASP**, 115:688–699, June 2003.

R. C. Kennicutt, Jr. Star Formation in Galaxies Along the Hubble Sequence. **Annu. Rev. Astron. Astrophys**, 36:189–232, 1998.

D. Kereš, N. Katz, D. H. Weinberg, and R. Davé. How do galaxies get their gas? **MNRAS**, 363:2–28, Oct. 2005.

C. Keyes, A. Koratkar, D. M., J. Hayes, J. Christensen, and S. Martin. *Faint Object Spectrograph Instrument Handbook.* The STScI FOS Group: Baltimore STScI, 1995.

T.-S. Kim, J. S. Bolton, M. Viel, M. G. Haehnelt, and R. F. Carswell. An improved measurement of the flux distribution of the Lyα forest in QSO absorption spectra: the effect of continuum fitting, metal contamination and noise properties. **MNRAS**, 382:1657–1674, Dec. 2007.

T.-S. Kim, R. F. Carswell, S. Cristiani, S. D'Odorico, and E. Giallongo. The physical properties of the Lyα forest at $z > 1.5$. **MNRAS**, 335:555–573, Sept. 2002.

J. Kim Quijano et al. *STIS Instrument Handbook, Version 8.0.* Baltimore: STScI, 2007.

D. Kirkman, D. Tytler, D. Lubin, and J. Charlton. Continuous statistics of the Lyα forest at $0 \leq z \leq 1.6$: the mean flux, flux distribution and autocorrelation from HST FOS spectra. **MNRAS**, 376:1227–1237, Apr. 2007.

D. Kirkman, D. Tytler, N. Suzuki, C. Melis, S. Hollywood, K. James, G. So, D. Lubin, T. Jena, M. L. Norman, and P. Paschos. The HI opacity of the intergalactic medium at redshifts $1.6 \leq z \leq 3.2$. **MNRAS**, 360:1373–1380, July 2005.

R. G. Kron. Photometry of a complete sample of faint galaxies. **ApJS**, 43:305–325, June 1980.

M. J. Kurtz and D. J. Mink. RVSAO 2.0: Digital Redshifts and Radial Velocities. **PASP**, 110:934–977, Aug. 1998.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. **MNRAS**, 271:676, Dec. 1994.

S. D. Landy and A. S. Szalay. Bias and variance of angular correlation functions. **ApJ**, 412:64–71, July 1993.

K. M. Lanzetta, D. V. Bowen, D. Tytler, and J. K. Webb. The gaseous extent of galaxies and the origin of Lyman-alpha absorption systems: A survey of galaxies in the fields of Hubble Space Telescope spectroscopic target QSOs. **ApJ**, 442: 538–568, Apr. 1995.

V. N. Lebedenko, H. M. Araujo, E. J. Barnes, A. Bewick, R. Cashmore, V. Chepel, D. Davidge, J. Dawson, T. Durkin, B. Edwards, C. Ghag, V. Graffagnino, M. Horn, A. S. Howard, A. J. Hughes, W. G. Jones, M. Joshi, G. E. Kalmus, A. G. Kovalenko, A. Lindote, I. Liubarsky, M. I. Lopes, R. Luscher, P. Majewski, A. StJ. Murphy, F. Neves, J. Pinto da Cunha, R. Preece, J. J. Quenby, P. R. Scovell, C. Silva, V. N. Solovov, N. J. T. Smith, P. F. Smith, V. N. Stekhanov, T. J. Sumnery, C. Thorne, and R. J. Walker. Result from the First Science Run of the ZEPLIN-III Dark Matter Search Experiment. *ArXiv e-prints; 0812.1150*, Dec. 2008.

O. LeFevre, M. Saisse, D. Mancini, S. Brau-Nogue, O. Caputi, L. Castinel, S. D'Odorico, B. Garilli, M. Kissler-Patig, C. Lucuix, G. Mancini, A. Pauget, G. Sciarretta, M. Scodeggio, L. Tresse, and G. Vettolani. Commissioning and performances of the VLT-VIMOS instrument. In M. Iye and A. F. M. Moorwood, editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4841 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1670–1681, Mar. 2003.

A. Liddle. *An Introduction to Modern Cosmology.* ©John Wiley & Sons, Inc, 1999.

L. Lu, A. M. Wolfe, and D. A. Turnshek. The redshift distribution of Ly-alpha clouds and the proximity effect. **ApJ**, 367:19–36, Jan. 1991.

C. R. Lynds and A. R. Sandage. Evidence for an Explosion in the Center of the Galaxy M82. **ApJ**, 137:1005, May 1963.

W. G. Mathews and F. Brighenti. Hot Gas in and around Elliptical Galaxies. **Annu. Rev. Astron. Astrophys**, 41:191–239, 2003.

A. L. Melott. Quasar L-alpha absorbers - Are precise conclusions possible. **ApJ**, 241:889–893, Nov. 1980.

H. J. Mo and S. L. Morris. The Sources of Lyman-Alpha Absorption at Low Redshifts - Galaxies Haloes and Minihaloes. **MNRAS**, 269:52, July 1994.

E. M. Monier, D. A. Turnshek, and O. L. Lupie. Hubble Space Telescope Observations of the Gravitationally Lensed Cloverleaf Broad Absorption Line QSO H1413+1143: Spectroscopy of the Ly alpha Forest and Metal-Line Systems. **ApJ**, 496:177, Mar. 1998.

S. L. Morris and B. T. Jannuzi. The association between gas and galaxies - I. CFHT spectroscopy and pair analysis. **MNRAS**, 367:1261–1281, Apr. 2006.

S. L. Morris, R. J. Weymann, A. Dressler, P. J. McCarthy, B. A. Smith, R. J. Terrile, R. Giovanelli, and M. Irwin. The Environment of Lyman- alpha Absorbers in the Sight Line toward 3C 273. **ApJ**, 419:524, Dec. 1993.

D. C. Morton. Atomic Data for Resonance Absorption Lines. III. Wavelengths Longward of the Lyman Limit for the Elements Hydrogen to Gallium. **ApJS**, 149:205–238, Nov. 2003.

C. Murali and J. Dubinski. Determining the Galactic Mass Distribution Using Tidal Streams from Globular Clusters. **AJ**, 118:911–919, Aug. 1999.

M. T. Murphy, P. Tzanavaris, J. K. Webb, and C. Lovis. Selection of ThAr lines for wavelength calibration of echelle spectra and implications for variations in the fine-structure constant. **MNRAS**, 378:221–230, June 2007.

A. Narayanan, T. Misawa, J. C. Charlton, and T. Kim. The Evolution of weak MgII Absorbers from $0 < z < 2.4$. In *Bulletin of the American Astronomical Society*, volume 38 of *Bulletin of the American Astronomical Society*, page 997, Dec. 2006.

J. B. Oke. Faint spectrophotometric standard stars. **AJ**, 99:1621–1631, May 1990.

J. B. Oke and D. G. Korycansky. Absolute spectrophotometry of very large redshift quasars. **ApJ**, 255:11–19, Apr. 1982.

B. D. Oppenheimer and R. Davé. Cosmological simulations of intergalactic medium enrichment from galactic outflows. **MNRAS**, 373:1265–1292, Dec. 2006.

B. D. Oppenheimer and R. A. Davé. The Nature and Origin of Low-Redshift O VI Absorbers. *ArXiv e-prints; 0806.2866*, June 2008.

P. J. Outram, B. J. Boyle, R. F. Carswell, P. C. Hewett, and R. E. Williams. Echelle spectrograph observations of the Hubble Deep Field South QSO J2233-606. **MNRAS**, 305:685–692, May 1999.

J. A. Peacock. *Cosmological Physics.* ©Cambridge University Press, 1999.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. **ApJ**, 142:419–421, July 1965.

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. **ApJ**, 517:565–586, June 1999.

C. Péroux, M. Dessauges-Zavadsky, S. D'Odorico, T.-S. Kim, and R. G. McMahon. A homogeneous sample of sub-damped Lyman systems - IV. Global metallicity evolution. **MNRAS**, 382:177–193, Nov. 2007.

M. Pettini, S. L. Ellison, C. C. Steidel, and D. V. Bowen. Metal Abundances at $z<1.5$: Fresh Clues to the Chemical Enrichment History of Damped Lyα Systems. **ApJ**, 510:576–589, Jan. 1999.

M. Pettini, A. E. Shapley, C. C. Steidel, J.-G. Cuby, M. Dickinson, A. F. M. Moorwood, K. L. Adelberger, and M. Giavalisco. The Rest-Frame Optical Spectra of Lyman Break Galaxies: Star Formation, Extinction, Abundances, and Kinematics. **ApJ**, 554:981–1000, June 2001.

M. Pierleoni, E. Branchini, and M. Viel. The relation between Lyman α absorbers and gas-rich galaxies in the local Universe. **MNRAS**, 388:282–292, July 2008.

A. Pollo, B. Meneux, L. Guzzo, O. Le Fèvre, J. Blaizot, A. Cappi, A. Iovino, C. Marinoni, H. J. McCracken, D. Bottini, B. Garilli, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, L. Tresse, G. Vettolani, A. Zanichelli, C. Adami, M. Arnaboldi, S. Arnouts, S. Bardelli, M. Bolzonella, S. Charlot, P. Ciliegi, T. Contini, S. Foucaud, P. Franzetti, I. Gavignaud, O. Ilbert, B. Marano, G. Mathez, A. Mazure, R. Merighi, S. Paltani, R. Pellò, L. Pozzetti, M. Radovich, G. Zamorani, E. Zucca, M. Bondi, A. Bongiorno, G. Busarello, L. Gregorini, F. Lamareille, Y. Mellier, P. Merluzzi, V. Ripepi, and D. Rizzo. The VIMOS VLT deep survey. Computing the two point correlation statistics and associated uncertainties. **A&A**, 439:887–900, Sept. 2005.

C. Porciani and P. Madau. The Origin of Intergalactic Metals around Lyman Break Galaxies. **ApJL**, 625:L43–L46, May 2005.

N. Prause, D. Reimers, C. Fechner, and E. Janknecht. The baryon density at z = 0.9-1.9. Tracing the warm-hot intergalactic medium with broad Lyman α absorption. **A&A**, 470:67–72, July 2007.

S. M. Rao, D. B. Nestor, D. A. Turnshek, W. M. Lane, E. M. Monier, and J. Bergeron. Low-Redshift Damped Lyα Galaxies toward the Quasars B2 0827+243, PKS 0952+179, PKS 1127-145, and PKS 1629+120. **ApJ**, 595:94–108, Sept. 2003a.

S. M. Rao, D. B. Nestor, D. A. Turnshek, W. M. Lane, E. M. Monier, and J. Bergeron. Low-Redshift Damped Lyα Galaxies toward the Quasars B2 0827+243, PKS 0952+179, PKS 1127-145, and PKS 1629+120. **ApJ**, 595:94–108, Sept. 2003b.

S. M. Rao, D. A. Turnshek, and F. H. Briggs. The Incidence of Damped Lyα Systems in the Redshift Interval $0 < Z < 4$. **ApJ**, 449:488, Aug. 1995.

M. Rauch. The Lyman Alpha Forest in the Spectra of QSOs. **Annu. Rev. Astron. Astrophys**, 36:267–316, 1998.

M. J. Rees. Lyman absorption lines in quasar spectra - Evidence for gravitationally-confined gas in dark minihaloes. **MNRAS**, 218:25P–30P, Jan. 1986.

M. J. Rees and J. P. Ostriker. Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters. **MNRAS**, 179: 541–559, June 1977.

P. Richter, B. D. Savage, T. M. Tripp, and K. R. Sembach. FUSE and STIS Observations of the Warm-hot Intergalactic Medium toward PG 1259+593. **ApJS**, 153:165–204, July 2004.

J. L. Rosenberg. The HI properties and environment of Lyman-α absorbers. In P. Williams, C.-G. Shu, and B. Menard, editors, *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, pages 98–103, Mar. 2005.

A. I. Ryabinkov, A. D. Kaminker, and D. A. Varshalovich. A catalogue of absorption-line systems in QSO spectra. **A&A**, 412:707–709, Dec. 2003.

E. V. Ryan-Weber. Cross-correlation of Lyman α absorbers with gas-rich galaxies. **MNRAS**, 367:1251–1260, Apr. 2006.

G. Rybicki and A. Lightman. *Radiative Processes in Astrophysics.* ©John Wiley & Sons, Inc, 1979.

W. L. W. Sargent, P. J. Young, A. Boksenberg, and D. Tytler. The distribution of Lyman-alpha absorption lines in the spectra of six QSOs - Evidence for an intergalactic origin. **ApJS**, 42:41–81, Jan. 1980.

S. Savaglio, H. C. Ferguson, T. M. Brown, B. R. Espey, K. C. Sahu, S. A. Baum, C. M. Carollo, M. E. Kaiser, M. Stiavelli, R. E. Williams, and J. Wilson. The Lyα Forest of the Quasar in the Hubble Deep Field South. **ApJL**, 515:L5–L8, Apr. 1999.

E. Scannapieco, C. Pichon, B. Aracil, P. Petitjean, R. J. Thacker, D. Pogosyan, J. Bergeron, and H. M. P. Couchman. The sources of intergalactic metals. **MNRAS**, 365:615–637, Jan. 2006.

J. Schaye. Model-independent Insights into the Nature of the Lyα Forest and the Distribution of Matter in the Universe. **ApJ**, 559:507–515, Oct. 2001.

J. Schaye, A. Aguirre, T.-S. Kim, T. Theuns, M. Rauch, and W. L. W. Sargent. Metallicity of the Intergalactic Medium Using Pixel Statistics. II. The Distribution of Metals as Traced by C IV. **ApJ**, 596:768–796, Oct. 2003.

J. Schaye, R. F. Carswell, and T.-S. Kim. A large population of metal-rich, compact, intergalactic CIV absorbers - evidence for poor small-scale metal mixing. **MNRAS**, 379:1169–1194, Aug. 2007.

J. Schaye and C. Dalla Vecchia. On the relation between the Schmidt and Kennicutt-Schmidt star formation laws and its implications for numerical simulations. **MNRAS**, 383:1210–1222, Jan. 2008.

D. P. Schneider, G. F. Hartig, B. T. Jannuzi, S. Kirhakos, D. H. Saxe, R. J. Weymann, J. N. Bahcall, J. Bergeron, A. Boksenberg, W. L. W. Sargent, B. D. Savage, D. A. Turnshek, and A. M. Wolfe. The Hubble Space Telescope quasar absorption line key project. II - Data calibration and absorption-line selection. **ApJS**, 87: 45–62, July 1993.

J. Scott, J. Bechtold, and A. Dobrzycki. A Uniform Analysis of the Lyα Forest at Z=0-5. I. The Sample and Distribution of Clouds at Z¿1.7. **ApJS**, 130:37–66, Sept. 2000.

K. R. Sembach, T. M. Tripp, B. D. Savage, and P. Richter. Physical Properties and Baryonic Content of Low-Redshift Intergalactic Lyα and O VI Absorption Line Systems: The PG 1116+215 Sight Line. **ApJS**, 155:351–393, Dec. 2004.

K. R. Sembach, B. P. Wakker, B. D. Savage, P. Richter, M. Meade, J. M. Shull, E. B. Jenkins, G. Sonneborn, and H. W. Moos. Highly Ionized High-Velocity Gas in the Vicinity of the Galaxy. **ApJS**, 146:165–208, May 2003.

A. Smette, J. G. Robertson, P. A. Shaver, D. Reimers, L. Wisotzki, and T. Koehler. The gravitational lens candidate HE 1104-1805 and the size of absorption systems. **A&AS**, 113:199, Oct. 1995.

G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. **ApJL**, 396:L1–L5, Sept. 1992.

Y. Sofue and V. Rubin. Rotation Curves of Spiral Galaxies. **Annu. Rev. Astron. Astrophys**, 39:137–174, 2001.

V. Springel. The cosmological simulation code GADGET-2. **MNRAS**, 364:1105–1134, Dec. 2005a.

V. Springel. The cosmological simulation code GADGET-2. **MNRAS**, 364:1105–1134, Dec. 2005b.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. **Nature**, 435:629–636, June 2005.

J. T. Stocke, C. W. Danforth, J. M. Shull, S. V. Penton, and M. L. Giroux. The Metallicity of Intergalactic Gas in Cosmic Voids. **ApJ**, 671:146–152, Dec. 2007.

J. T. Stocke, S. V. Penton, C. W. Danforth, J. M. Shull, J. Tumlinson, and K. M. McLin. The Galaxy Environment of O VI Absorption Systems. **ApJ**, 641:217–228, Apr. 2006.

T. Theuns, A. Leonard, and G. Efstathiou. The low-redshift evolution of the Lyman-alpha forest. **MNRAS**, 297:L49–L52, June 1998.

J. Tonry and M. Davis. A survey of galaxy redshifts. I - Data reduction techniques. **AJ**, 84:1511–1525, Oct. 1979.

T. M. Tripp, B. Aracil, D. V. Bowen, and E. B. Jenkins. The O VI Absorbers toward PG 0953+415: High-Metallicity, Cosmic-Web Gas Far from Luminous Galaxies. **ApJL**, 643:L77–L82, June 2006.

D. Tytler, J. M. O'Meara, N. Suzuki, D. Kirkman, D. Lubin, and A. Orin. The Kast Ground-based Ultraviolet Spectral Survey of 79 Quasi-stellar Objects at Redshift 2 for Lyα Forest and Metal Absorption. **AJ**, 128:1058–1077, Sept. 2004.

J. H. van Gorkom, C. L. Carilli, J. T. Stocke, E. S. Perlman, and J. M. Shull. The HI Environment of Nearby Lyman-Alpha Absorbers. **AJ**, 112:1397, Oct. 1996.

S. Veilleux, G. Cecil, and J. Bland-Hawthorn. Galactic Winds. **Annu. Rev. Astron. Astrophys**, 43:769–826, Sept. 2005.

D. A. Verner, P. D. Barthel, and D. Tytler. Atomic data for absorption lines from the ground level at wavelengths greater than 228A. **A&AS**, 108:287–340, Dec. 1994.

R. J. Weymann, B. T. Jannuzi, L. Lu, J. N. Bahcall, J. Bergeron, A. Boksenberg, G. F. Hartig, S. Kirhakos, W. L. W. Sargent, B. D. Savage, D. P. Schneider, D. A. Turnshek, and A. M. Wolfe. The Hubble Space Telescope Quasar Absorption Line Key Project. XIV. The Evolution of Lyalpha Absorption Lines in the Redshift Interval Z = 0-1.5. **ApJ**, 506:1–18, Oct. 1998.

S. D. M. White and M. J. Rees. Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. **MNRAS**, 183:341–358, May 1978.

R. P. C. Wiersma, J. Schaye, and B. D. Smith. The effect of photoionization on the cooling rates of enriched, astrophysical plasmas. **MNRAS**, 393:99–107, Feb. 2009.

G. M. Williger, S. R. Heap, R. J. Weymann, R. Davé, E. Ellingson, R. F. Carswell, T. M. Tripp, and E. B. Jenkins. The Low-Redshift Lyα Forest toward PKS 0405-123. **ApJ**, 636:631–653, Jan. 2006.

R. J. Wilman, S. L. Morris, B. T. Jannuzi, R. Davé, and A. M. Shone. The association between gas and galaxies - II. The two-point correlation function. **MNRAS**, 375:735–744, Feb. 2007.

D. S. Womble, W. L. W. Sargent, and R. S. Lyons. Heavy Elements in the Lyman - α Forest: Abundances and Clustering at z = 3. In M. N. Bremer and N. Malcolm, editors, *Cold Gas at High Redshift*, volume 206 of *Astrophysics and Space Science Library*, page 249, 1996.

N. Yoshida, K. Omukai, L. Hernquist, and T. Abel. Formation of Primordial Stars in a ΛCDM Universe. **ApJ**, 652:6–25, Nov. 2006.

P. Young, W. L. W. Sargent, and A. Boksenberg. A high-resolution study of the absorption spectra of three QSOs - Evidence for cosmological evolution in the Lyman-alpha lines. **ApJ**, 252:10–31, Jan. 1982.

P. A. Young, C. D. Impey, and C. B. Foltz. Observations of Lyα Absorption in a Triple Quasar System. **ApJ**, 549:76–89, Mar. 2001.

M. Zwaan, F. Walter, E. Ryan-Weber, E. Brinks, W. J. G. de Blok, and R. C. Kennicutt. Are the Kinematics of DLAs in Agreement with Their Arising in the Gas Disks of Galaxies? **AJ**, 136:2886–2896, Dec. 2008.

M. A. Zwaan, L. Staveley-Smith, B. S. Koribalski, P. A. Henning, V. A. Kilborn, S. D. Ryder, D. G. Barnes, R. Bhathal, P. J. Boyce, W. J. G. de Blok, M. J. Disney, M. J. Drinkwater, R. D. Ekers, K. C. Freeman, B. K. Gibson, A. J. Green, R. F. Haynes, H. Jerjen, S. Juraszek, M. J. Kesteven, P. M. Knezek, R. C. Kraan-Korteweg, S. Mader, M. Marquarding, M. Meyer, R. F. Minchin, J. R. Mould, J. O'Brien, T. Oosterloo, R. M. Price, M. E. Putman, E. Ryan-Weber, E. M. Sadler, A. Schröder, I. M. Stewart, F. Stootman, B. Warren, M. Waugh, R. L. Webster, and A. E. Wright. The 1000 Brightest HIPASS Galaxies: The H I Mass Function and $\Omega_{\rm HI}$. **AJ**, 125:2842–2858, June 2003.

F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

B. J. Zych, M. T. Murphy, M. Pettini, P. C. Hewett, E. V. Ryan-Weber, and S. L. Ellison. The host galaxies of strong CaII quasar absorption systems at z < 0.5. **MNRAS**, 379:1409–1422, Aug. 2007.

Appendix A

The Galaxy Catalog

A.1 The HE 1122-1648 Galaxies

Table A.1: The FORS2 galaxy sample in the field-of-view of HE 1122-1648

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
0.18756	0.00041	11 24 37.06	-17 06 10.40	22.12	-16.40	-15.01	Emission	2
0.25417	0.00041	11 24 36.54	-17 09 17.82	22.03	-17.09	-15.65	Emission	2
0.29554	0.00013	11 24 36.01	-17 08 01.07	21.97	-17.45	-15.98	Emission	1
0.30467	0.00056	11 24 44.60	-17 09 22.76	22.35	-17.12	-15.66	Emission	1
0.34186	0.00007	11 24 45.54	-17 03 18.88	22.00	-17.72	-16.21	Emission	3
0.34649	0.00008	11 24 35.99	-17 05 45.78	22.95	-16.80	-15.28	Emission	1
0.37733	0.00017	11 24 44.35	-17 04 26.94	22.03	-17.90	-16.34	Emission	1
0.37914	0.00020	11 24 40.32	-17 07 37.52	22.50	-17.45	-15.89	Emission	1
0.38256	0.00011	11 24 35.37	-17 06 54.39	22.82	-17.14	-15.57	Emission	1
0.40595	0.00004	11 24 39.63	-17 02 39.56	22.69	-17.41	-15.80	Emission	3
0.40719	0.00057	11 24 34.80	-17 08 32.28	23.22	-16.88	-15.27	Emission	3
0.41215	0.00014	11 24 50.59	-17 03 11.03	21.65	-18.48	-16.86	Emission	1
0.44624	0.00014	11 24 47.28	-17 03 02.58	23.35	-16.95	-15.28	Emission	1
0.46276	0.00025	11 24 36.98	-17 04 55.59	21.67	-18.73	-17.02	Emission	1
0.46545	0.00002	11 24 49.06	-17 02 28.28	23.01	-17.40	-15.69	Emission	2
0.50538	0.00013	11 24 37.69	-17 02 51.20	22.43	-18.20	-16.42	Emission	1
0.52913	0.00198	11 24 39.57	-17 03 43.02	22.18	-18.55	-16.76	Absorption	4
0.54120	0.00049	11 24 47.82	-17 07 21.95	21.76	-19.03	-17.22	Absorption	4
0.54166	0.00029	11 24 38.59	-17 05 01.91	22.73	-18.06	-16.26	Emission	1
0.54204	0.00045	11 24 39.63	-17 06 44.00	21.96	-18.83	-17.02	Absorption	2
0.54523	0.00018	11 24 47.09	-17 07 30.53	22.48	-18.33	-16.52	Emission	1
0.54721	0.00018	11 24 35.58	-17 05 05.74	21.99	-18.83	-17.02	Emission	1
0.55121	0.00008	11 24 36.54	-17 07 04.62	23.29	-17.54	-15.73	Emission	1
0.57845	0.01839	11 24 46.99	-17 04 47.93	23.20	-17.77	-15.93	Emission	1
0.59406	0.00073	11 24 47.46	-17 03 22.26	21.56	-19.47	-17.62	Emission	2
0.59512	0.00022	11 24 50.08	-17 02 48.21	22.39	-18.66	-16.80	Emission	1

Continued on next page

---

[a] The RMS error determined by the redshift of multiple lines $\sigma_z = \sqrt{\frac{\sum_{i=1}^{n}(z_{feature}-\bar{z})^2}{n-1}}$. Thus single line error is not applicable.

[b] MAG_AUTO that was computed from SExtractor on the pre-imaging.

[c] Maximum and minimum M_B based on synthetic colours, see Section 3.6.

[d] The confidence in the redshift based on criteria described in Table 3.6.

Table A.1 – continued from previous page

z	RMS Error^a	RA (J2000)	DEC (J2000)	R mag^b	Max^c M_B	Min M_B	Emission/ Absorption	Conf.^{d}
0.59619	0.00019	11 24 40.58	-17 07 45.85	23.48	-17.57	-15.72	Emission	2
0.59670	···	11 24 41.46	-17 06 42.99	21.84	-19.20	-17.35	Emission	5
0.59824	0.00036	11 24 40.85	-17 05 18.37	22.24	-18.82	-16.96	Emission	1
0.59951	0.00013	11 24 35.28	-17 07 24.90	23.24	-17.82	-15.97	Emission	2
0.61141	0.00059	11 24 44.93	-17 03 56.22	22.39	-18.72	-16.86	Emission	2
0.61853	0.00031	11 24 44.53	-17 03 50.77	22.99	-18.16	-16.29	Emission	2
0.62907	···	11 24 35.81	-17 06 27.72	22.16	-19.04	-17.16	Emission	5
0.63112	···	11 24 47.90	-17 05 54.07	21.70	-19.51	-17.63	Emission	5
0.64270	0.00006	11 24 43.84	-17 05 28.79	23.40	-17.86	-15.97	Emission	1
0.64456	···	11 24 46.26	-17 05 12.57	22.33	-18.94	-17.05	Emission	5
0.65144	0.00030	11 24 39.03	-17 07 28.44	22.55	-18.75	-16.86	Emission	2
0.66628	···	11 24 38.27	-17 02 35.63	22.48	-18.89	-16.99	Emission	5
0.68212	0.00052	11 24 45.07	-17 05 17.18	23.07	-18.36	-16.45	Emission	1
0.68279	0.00055	11 24 43.00	-17 06 00.10	22.19	-19.25	-17.34	Emission	2
0.73646	0.00164	11 24 49.04	-17 08 38.33	22.64	-19.02	-17.06	Absorption	3
0.74561	0.00084	11 24 49.32	-17 08 47.58	21.70	-20.00	-18.04	Absorption	2
0.76737	···	11 24 43.57	-17 06 44.55	23.19	-18.60	-16.62	Emission	5
0.76840	0.00060	11 24 49.25	-17 08 51.46	22.61	-19.19	-17.20	Absorption	3
0.76844	0.00045	11 24 35.22	-17 09 06.25	22.16	-19.64	-17.65	Emission	2
0.77271	0.00010	11 24 42.43	-17 03 04.85	23.14	-18.68	-16.69	Emission	3
0.77349	···	11 24 35.79	-17 08 17.70	21.62	-20.20	-18.21	Emission	5
0.77411	···	11 24 36.16	-17 08 19.60	22.78	-19.04	-17.05	Emission	5
0.77530	···	11 24 35.38	-17 04 58.30	22.27	-19.56	-17.56	Emission	5
0.80445	0.00024	11 24 47.21	-17 07 56.82	22.75	-19.20	-17.19	Absorption	4
0.80686	···	11 24 40.79	-17 05 23.76	21.84	-20.12	-18.11	Emission	5
0.80777	···	11 24 35.57	-17 06 09.24	22.83	-19.14	-17.12	Emission	5
0.81635	···	11 24 35.25	-17 08 09.60	21.91	-20.09	-18.07	Emission	5
0.82440	0.00054	11 24 50.74	-17 05 21.41	21.98	-20.05	-18.03	Emission	1
0.82565	0.00027	11 24 39.57	-17 07 28.17	23.07	-18.96	-16.94	Emission	2
0.82581	0.00020	11 24 39.52	-17 07 45.49	23.16	-18.88	-16.85	Emission	1
0.82603	0.00007	11 24 37.89	-17 07 12.04	22.05	-19.99	-17.97	Emission	1
0.82779	0.00003	11 24 35.32	-17 09 16.91	22.66	-19.39	-17.36	Emission	2
0.83136	0.00068	11 24 36.03	-17 06 56.21	22.13	-19.93	-17.90	Emission	2
0.83173	···	11 24 48.97	-17 07 01.64	22.23	-19.83	-17.80	Emission	5
0.83201	···	11 24 33.89	-17 08 03.06	23.27	-18.79	-16.76	Emission	5
0.83687	0.00036	11 24 36.81	-17 05 46.48	23.30	-18.78	-16.75	Emission	2
0.83910	0.00062	11 24 35.55	-17 06 35.87	22.50	-19.59	-17.56	Emission	2
0.84008	···	11 24 38.12	-17 03 25.41	22.66	-19.44	-17.40	Emission	5
0.86042	0.00034	11 24 35.15	-17 02 25.13	22.71	-19.46	-17.41	Emission	1
0.87629	0.00036	11 24 41.43	-17 08 51.78	23.32	-18.92	-16.84	Emission	2
0.89710	···	11 24 50.75	-17 06 11.29	22.33	-19.99	-17.88	Emission	5
0.90230	0.02365	11 24 37.55	-17 06 47.61	23.40	-18.94	-16.81	Emission	1
0.90747	···	11 24 34.64	-17 05 35.56	23.22	-19.14	-17.01	Emission	5
0.90795	0.00045	11 24 48.68	-17 03 49.34	22.90	-19.47	-17.33	Emission	2
0.90888	0.00043	11 24 48.08	-17 03 34.04	21.93	-20.43	-18.30	Emission	2
0.93033	···	11 24 39.53	-17 05 42.74	22.62	-19.83	-17.66	Emission	5
0.94291	···	11 24 38.49	-17 02 25.51	22.93	-19.56	-17.37	Emission	5
0.97659	···	11 24 35.62	-17 04 35.95	22.23	-20.39	-18.15	Emission	5

Continued on next page

Table A.1 – continued from previous page

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
0.98497	···	11 24 37.00	-17 02 28.88	23.18	-19.48	-17.22	Emission	5
1.01020	···	11 24 41.27	-17 06 12.46	23.08	-19.66	-17.39	Emission	5
1.11580	···	11 24 41.08	-17 03 08.91	22.92	-20.12	-17.83	Emission	5
1.16870	···	11 24 43.36	-17 04 34.10	22.04	-21.14	-18.87	Emission	5
1.21190	···	11 24 43.06	-17 08 32.74	23.41	-19.86	-17.63	Emission	5
1.30490	···	11 24 45.04	-17 09 06.21	22.69	-20.76	-18.65	Emission	5
1.33700	···	11 24 41.97	-17 02 50.21	22.87	-20.63	-18.57	Emission	5
1.46160	···	11 24 35.80	-17 02 43.87	22.75	-20.92	-19.09	Emission	5
1.46860	···	11 24 35.76	-17 07 55.75	23.39	-20.29	-18.47	Emission	5

A.2 The PKS 1127-145 Galaxies

Table A.2: The FORS2 galaxy sample in the field-of-view of PKS 1127-145

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
0.05279	0.00006	11 30 12.22	-14 51 41.29	22.49	-14.31	-14.31	Emission	1
0.21085	0.00078	11 30 01.30	-14 48 44.44	22.67	-16.10	-14.67	Emission	3
0.25315	0.00039	11 29 59.89	-14 51 27.11	22.75	-16.35	-14.92	Emission	1
0.25778	0.00054	11 30 07.83	-14 47 22.79	23.23	-15.92	-14.47	Emission	2
0.27298	0.00034	11 30 11.90	-14 52 17.39	21.80	-17.44	-16.01	Emission	1
0.27313	0.00046	11 30 00.27	-14 50 28.47	23.34	-15.90	-14.47	Emission	3
0.27408	0.00022	11 29 59.43	-14 50 36.79	23.13	-16.12	-14.69	Emission	1
0.27918	0.00018	11 30 01.71	-14 52 48.01	21.72	-17.58	-16.13	Emission	3
0.29138	0.00051	11 30 02.72	-14 52 38.92	22.56	-16.84	-15.36	Emission	3
0.30509	0.00003	11 30 09.69	-14 51 10.54	22.02	-17.45	-15.99	Emission	1
0.33313	0.00022	11 29 58.15	-14 52 46.94	21.61	-18.05	-16.55	Emission	1
0.35425	0.00018	11 30 03.36	-14 46 58.73	21.72	-18.07	-16.55	Emission	1
0.35847	0.00027	11 30 08.18	-14 46 27.59	22.89	-16.93	-15.40	Emission	1
0.36420	0.00018	11 30 09.27	-14 47 05.81	21.74	-18.11	-16.58	Emission	2
0.40274	···	11 30 05.55	-14 52 36.38	23.43	-16.65	-15.05	Emission	5
0.40921	0.00030	11 30 10.19	-14 50 47.29	23.19	-16.92	-15.31	Emission	1
0.42210	0.00014	11 30 03.47	-14 52 31.02	23.40	-16.78	-15.15	Emission	1

Continued on next page

---

[a] The RMS error determined by the redshift of multiple lines $\sigma_z = \sqrt{\frac{\sum_{i=1}^{n}(z_{\mathrm{feature}}-\bar{z})^2}{n-1}}$. Thus single line error is not applicable.

[b] MAG_BEST that was computed from SExtractor on the pre-imaging.

[c] Maximum and minimum M_B based on synthetic colours, see Section 3.6.

[d] The confidence in the redshift based on criteria described in Table 3.6.

Table A.2 – continued from previous page

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
0.42315	0.00012	11 30 09.31	-14 48 28.08	22.02	-18.17	-16.53	Emission	1
0.42992	0.00022	11 30 02.40	-14 47 28.76	21.55	-18.67	-17.02	Emission	1
0.42995	···	11 30 13.65	-14 51 29.68	23.21	-17.01	-15.36	Emission	5
0.43037	0.00027	11 30 02.31	-14 48 58.76	22.43	-17.79	-16.14	Emission	1
0.43166	0.00020	11 30 09.57	-14 45 59.49	23.36	-16.87	-15.22	Emission	1
0.43168	0.00015	11 30 01.44	-14 49 16.56	22.24	-17.99	-16.34	Emission	1
0.43240	0.00023	11 30 00.75	-14 46 26.53	22.42	-17.82	-16.16	Emission	3
0.43259	0.00014	11 30 09.44	-14 46 01.03	22.60	-17.64	-15.98	Emission	1
0.44461	0.00025	11 29 58.27	-14 48 58.64	21.56	-18.74	-17.07	Emission	1
0.50216	0.00027	11 29 59.47	-14 50 39.12	23.26	-17.34	-15.57	Emission	2
0.51862	0.00021	11 30 14.23	-14 50 13.95	21.99	-18.70	-16.91	Emission	1
0.53750	···	11 30 02.71	-14 51 02.44	22.33	-18.44	-16.64	Emission	5
0.55646	0.00031	11 30 05.74	-14 49 30.70	22.52	-18.34	-16.52	Emission	1
0.58031	0.00020	11 30 04.98	-14 46 55.30	23.30	-17.67	-15.83	Emission	3
0.61623	···	11 30 03.58	-14 51 30.77	21.87	-19.27	-17.41	Emission	5
0.61707	0.00029	11 30 00.03	-14 52 19.62	22.52	-18.62	-16.75	Emission	2
0.62562	0.00028	11 30 10.34	-14 47 51.97	21.95	-19.23	-17.35	Emission	1
0.64545	0.00019	11 30 13.55	-14 47 44.69	21.72	-19.55	-17.67	Emission	1
0.74235	0.00029	11 30 00.15	-14 47 53.38	22.79	-18.90	-16.94	Emission	1
0.74851	0.00013	11 30 10.77	-14 47 23.52	22.09	-19.63	-17.66	Emission	2
0.74993	0.00028	11 30 09.30	-14 49 32.96	22.99	-18.74	-16.76	Absorption	4
0.76420	···	11 30 00.43	-14 52 48.74	21.74	-20.04	-18.06	Emission	5
0.76736	···	11 30 12.60	-14 51 26.24	22.15	-19.65	-17.66	Emission	5
0.78175	0.00010	11 29 58.87	-14 50 24.10	22.20	-19.66	-17.66	Emission	1
0.79654	···	11 30 10.70	-14 46 02.74	23.89	-18.02	-16.02	Emission	5
0.79689	0.00039	11 30 14.89	-14 51 46.90	22.20	-19.72	-17.71	Emission	1
0.79815	0.00031	11 30 04.51	-14 46 14.24	22.76	-19.16	-17.15	Emission	1
0.79819	0.00031	11 30 02.64	-14 50 08.86	22.70	-19.22	-17.22	Emission	1
0.79875	0.00041	11 30 02.41	-14 50 32.72	22.63	-19.29	-17.28	Emission	1
0.79909	···	11 30 13.31	-14 46 15.27	22.43	-19.49	-17.48	Emission	5
0.79959	0.00029	11 29 58.96	-14 47 33.29	22.08	-19.85	-17.84	Emission	1
0.80084	0.00010	11 30 06.83	-14 49 19.81	23.10	-18.83	-16.82	Emission	2
0.80155	0.00022	11 30 12.92	-14 48 25.42	21.70	-20.24	-18.23	Emission	2
0.80184	···	11 30 14.75	-14 52 47.21	24.17	-17.77	-15.76	Emission	5
0.80187	···	11 30 14.16	-14 48 04.05	23.45	-18.49	-16.48	Emission	5
0.81298	···	11 30 10.32	-14 47 01.34	23.76	-18.23	-16.21	Emission	5
0.81745	0.00075	11 30 07.42	-14 50 47.54	22.52	-19.49	-17.47	Absorption	4
0.82084	0.00046	11 30 02.44	-14 51 59.22	22.82	-19.19	-17.17	Emission	2
0.82855	···	11 30 08.62	-14 52 55.48	22.40	-19.65	-17.62	Emission	5
0.83903	0.00012	11 30 06.42	-14 50 52.95	23.39	-18.70	-16.67	Absorption	3
0.84849	0.00017	11 30 01.00	-14 47 08.29	22.29	-19.84	-17.79	Absorption	4
0.84904	···	11 30 14.73	-14 50 23.14	23.63	-18.50	-16.46	Emission	5
0.84909	0.00021	11 30 12.36	-14 51 34.63	23.03	-19.10	-17.06	Emission	1
0.85377	···	11 29 59.52	-14 48 28.77	22.51	-19.64	-17.60	Emission	5
0.85469	0.00026	11 30 00.75	-14 49 34.03	23.21	-18.94	-16.89	Emission	2
0.85634	···	11 30 03.51	-14 52 56.52	22.95	-19.21	-17.16	Emission	5
0.86960	···	11 30 06.65	-14 52 08.45	23.25	-18.96	-16.89	Emission	5
0.89411	···	11 30 14.50	-14 47 31.92	22.45	-19.85	-17.75	Emission	5

Continued on next page

Table A.2 – continued from previous page

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
0.90275	0.00022	11 30 10.33	-14 50 23.31	22.94	-19.40	-17.28	Emission	1
0.93048	···	11 30 01.94	-14 48 51.56	24.14	-18.31	-16.14	Emission	5
0.94074	0.00099	11 30 05.58	-14 50 23.82	22.50	-19.98	-17.80	Emission	2
0.94896	0.00023	11 30 11.79	-14 51 20.40	22.50	-20.02	-17.82	Emission	2
0.95031	0.00040	11 30 06.20	-14 48 34.11	23.27	-19.25	-17.05	Emission	3
0.95149	0.00027	11 30 10.37	-14 48 02.44	23.44	-19.09	-16.88	Emission	2
0.95566	···	11 29 58.74	-14 49 41.66	23.03	-19.51	-17.30	Emission	5
1.00280	0.00035	11 30 10.46	-14 52 12.57	22.03	-20.69	-18.42	Absorption	4
1.01870	···	11 30 07.35	-14 48 43.31	22.18	-20.60	-18.32	Emission	5
1.04670	···	11 30 04.25	-14 51 16.91	22.98	-19.87	-17.58	Emission	5
1.04970	···	11 30 08.88	-14 52 11.37	23.33	-19.53	-17.24	Emission	5
1.05070	0.00096	11 30 08.66	-14 52 22.33	22.04	-20.83	-18.54	Absorption	4
1.08980	···	11 30 09.54	-14 46 07.58	22.53	-20.45	-18.15	Emission	5
1.09290	···	11 30 11.99	-14 52 22.73	22.64	-20.34	-18.05	Emission	5
1.09480	···	11 30 04.77	-14 46 21.64	21.84	-21.15	-18.85	Emission	5
1.09600	···	11 30 09.99	-14 48 36.58	23.02	-19.97	-17.67	Emission	5
1.09620	···	11 30 08.42	-14 50 16.11	22.68	-20.31	-18.01	Emission	5
1.09750	···	11 30 09.41	-14 48 30.99	22.13	-20.86	-18.56	Emission	5
1.09820	···	11 30 04.79	-14 48 14.17	22.97	-20.02	-17.73	Emission	5
1.11890	···	11 30 13.90	-14 51 04.74	23.09	-19.96	-17.67	Emission	5
1.12660	···	11 30 09.06	-14 49 11.27	23.38	-19.69	-17.41	Emission	5
1.14260	···	11 30 13.35	-14 52 20.95	23.71	-19.40	-17.12	Emission	5
1.14280	···	11 30 07.99	-14 51 18.94	22.67	-20.44	-18.16	Emission	5
1.16040	···	11 29 58.37	-14 48 15.90	22.64	-20.51	-18.24	Emission	5
1.18930	···	11 30 14.40	-14 48 00.48	22.32	-20.90	-18.65	Emission	5
1.18980	···	11 30 06.81	-14 49 12.36	23.36	-19.87	-17.62	Emission	5
1.19030	···	11 30 01.25	-14 46 46.05	22.65	-20.57	-18.33	Emission	5
1.21120	···	11 29 59.64	-14 50 35.52	22.16	-21.11	-18.88	Emission	5
1.21350	···	11 30 14.72	-14 50 07.63	23.02	-20.26	-18.03	Emission	5
1.21400	···	11 30 06.45	-14 51 10.42	23.80	-19.47	-17.24	Emission	5
1.21560	···	11 30 04.37	-14 52 34.80	22.73	-20.55	-18.32	Emission	5
1.21570	···	11 30 11.76	-14 52 10.17	22.83	-20.45	-18.22	Emission	5
1.21670	···	11 30 07.59	-14 52 27.91	22.57	-20.71	-18.48	Emission	5
1.24780	···	11 30 15.00	-14 52 56.38	23.24	-20.10	-17.91	Emission	5
1.26970	···	11 30 06.89	-14 52 30.08	23.30	-20.09	-17.93	Emission	5
1.27310	···	11 30 05.69	-14 47 32.72	23.23	-20.16	-18.01	Emission	5
1.27330	···	11 30 06.00	-14 47 18.47	22.31	-21.08	-18.93	Emission	5
1.27840	···	11 30 11.07	-14 51 17.63	23.72	-19.68	-17.53	Emission	5
1.27850	···	11 30 07.28	-14 47 42.64	23.34	-20.06	-17.92	Emission	5
1.27990	···	11 30 14.50	-14 46 34.66	23.36	-20.05	-17.90	Emission	5
1.28840	···	11 30 10.71	-14 51 12.26	22.81	-20.61	-18.48	Emission	5
1.33780	···	11 29 59.10	-14 48 47.57	22.37	-21.14	-19.08	Emission	5
1.35650	···	11 30 03.91	-14 50 09.45	23.47	-20.07	-18.04	Emission	5
1.36140	···	11 29 58.69	-14 49 47.01	22.88	-20.66	-18.64	Emission	5
1.36260	···	11 30 13.38	-14 50 08.31	22.58	-20.97	-18.95	Emission	5
1.36270	···	11 29 59.24	-14 47 07.42	22.65	-20.89	-18.87	Emission	5
1.38460	···	11 30 12.68	-14 51 12.57	22.88	-20.70	-18.72	Emission	5
1.41350	···	11 30 14.78	-14 51 35.37	22.93	-20.69	-18.76	Emission	5

Continued on next page

Table A.2 – continued from previous page

z	RMS Error[a]	RA (J2000)	DEC (J2000)	R mag[b]	Max[c] M_B	Min M_B	Emission/ Absorption	Conf.[d]
1.44390	···	11 30 10.34	-14 50 05.23	22.64	-21.02	-19.15	Emission	5
1.44900	···	11 30 04.64	-14 50 40.73	22.06	-21.60	-19.74	Emission	5
1.46880	···	11 29 59.58	-14 52 10.40	24.03	-19.65	-17.83	Emission	5
1.47320	···	11 30 09.61	-14 51 51.09	23.00	-20.69	-18.87	Emission	5

Appendix B

The QSO Echelle Absorption Spectra

B.1 STIS E230M Spectra

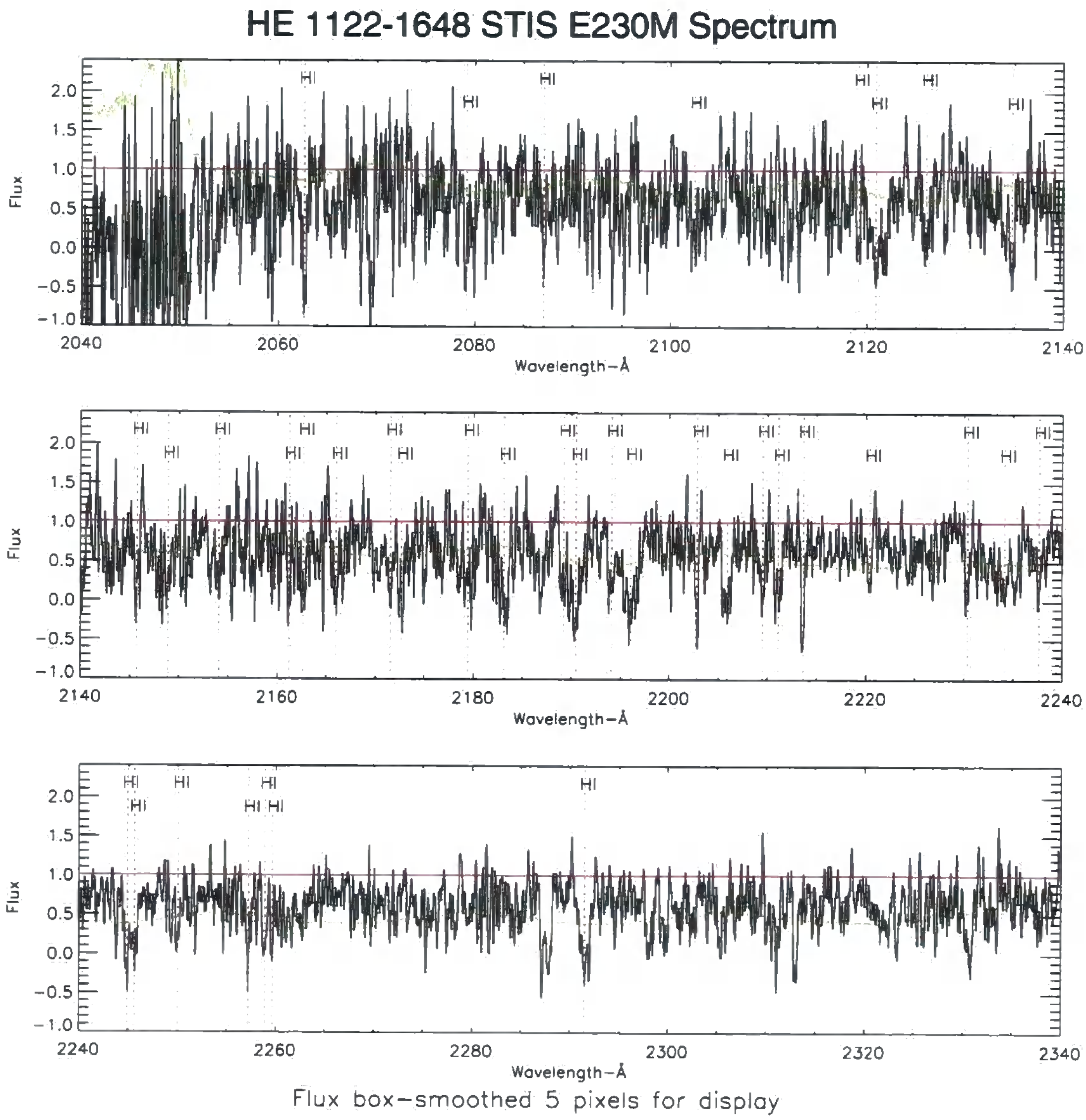


Figure B.1 The Blue HE 1122-1648 E230M STIS Spectrum between 2050-2280 Å. This was the spectrum where the Lyα absorbers were identified by the method in Section 4.4.3.

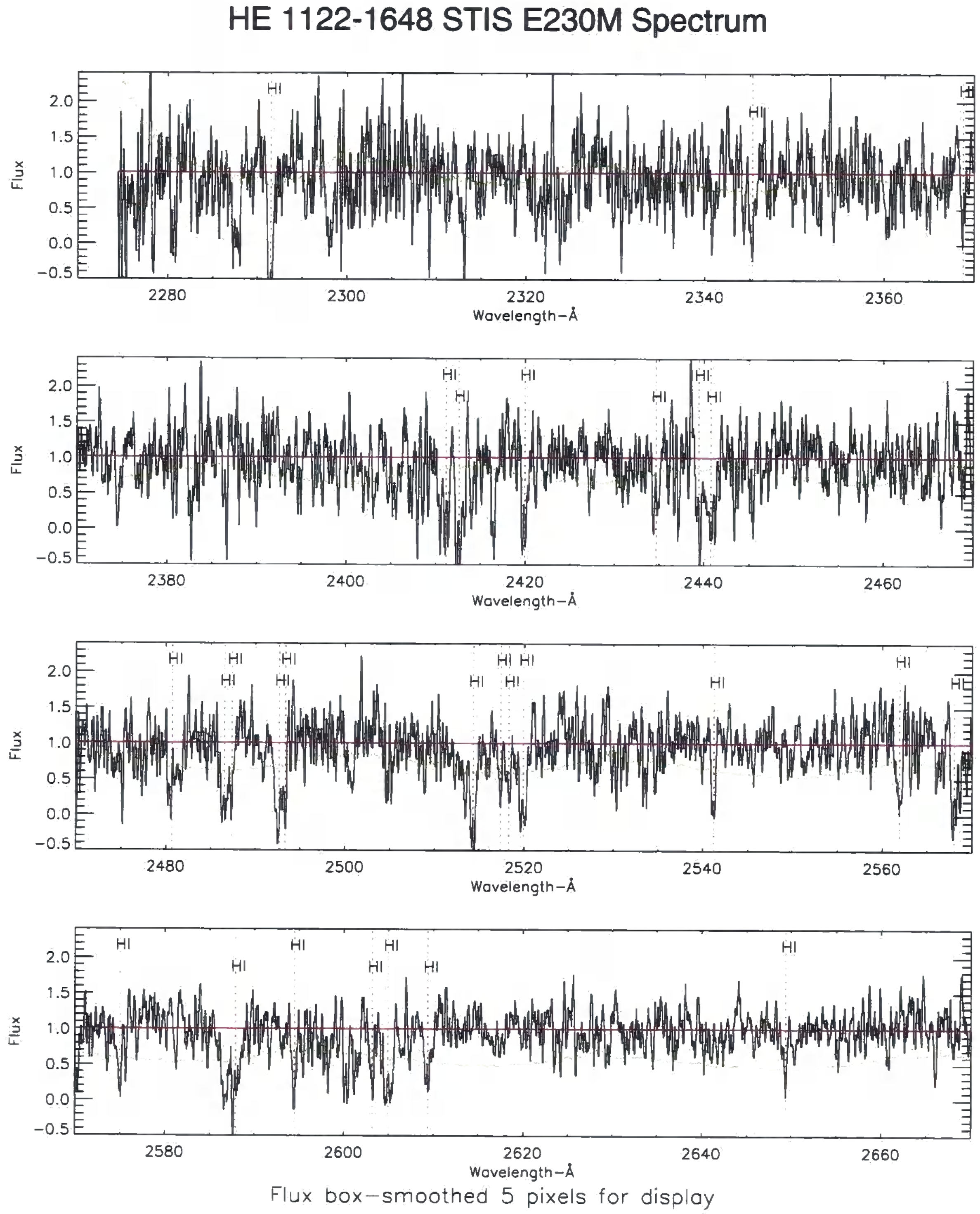


Figure B.2 The Red HE 1122-1648 E230M STIS Spectrum between 2280-2670 Å.

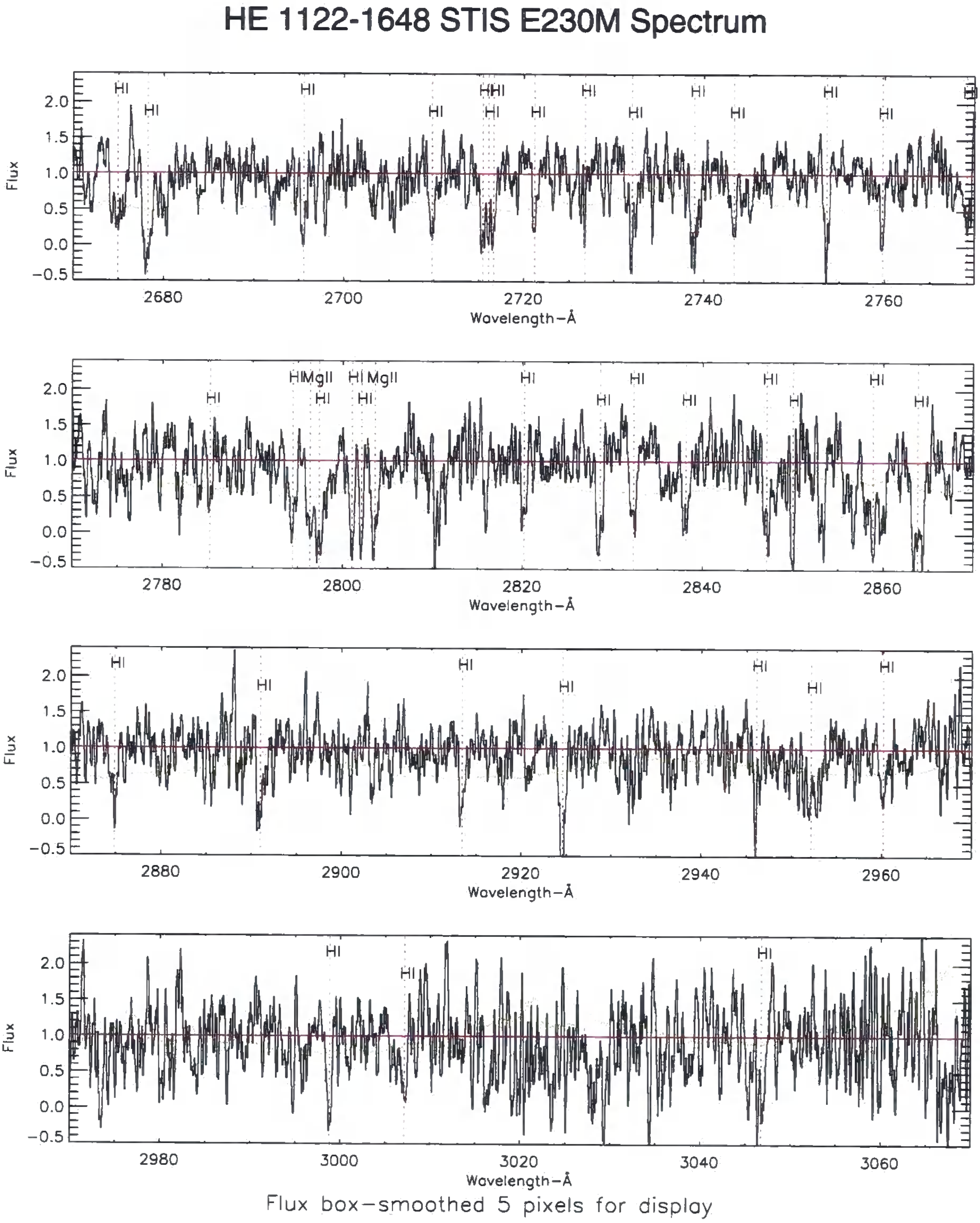


Figure B.3 The Red HE 1122-1648 E230M STIS Spectrum between 2670-3050 Å.

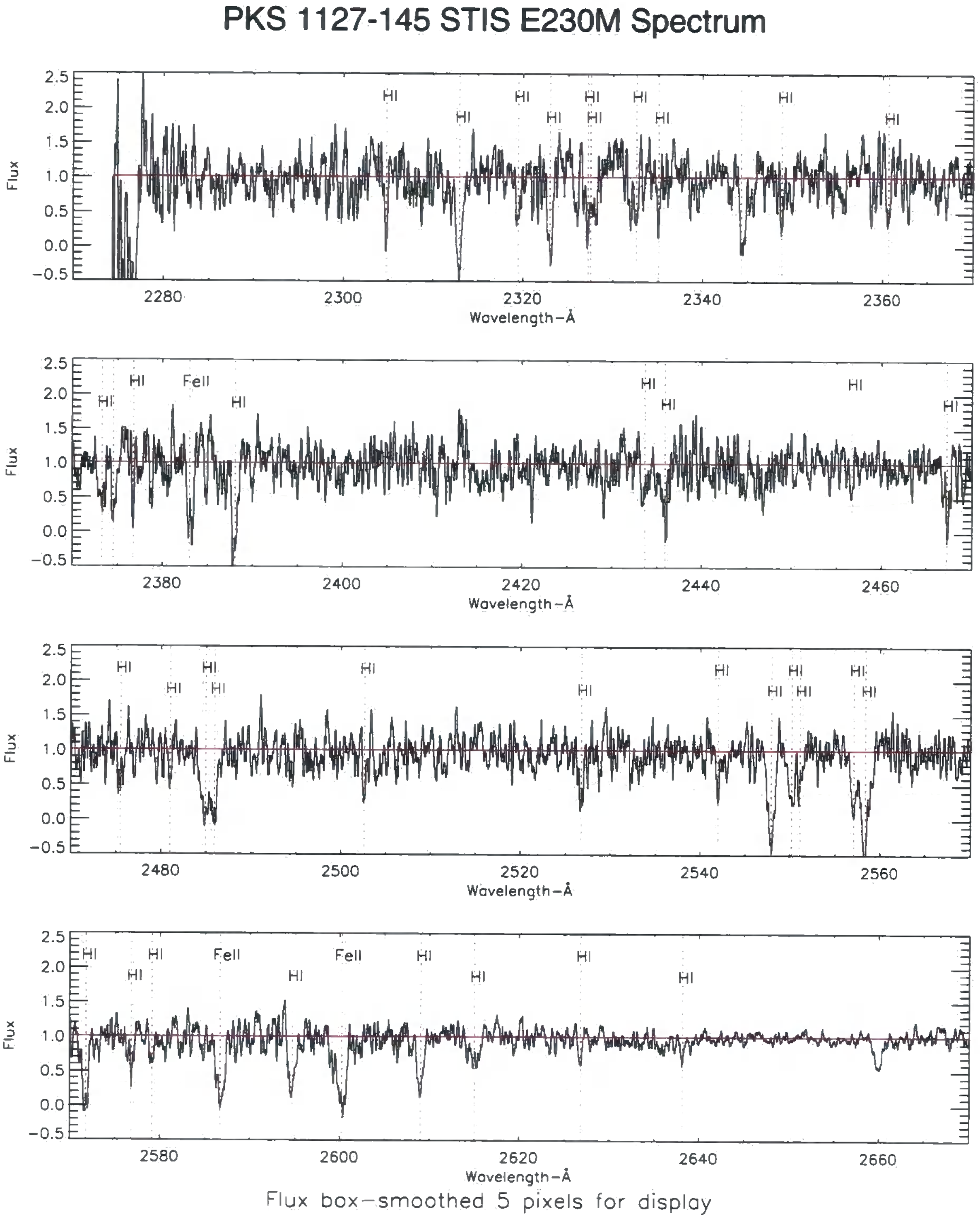


Figure B.4 The PKS 1127-145 E230M STIS Spectrum between 2280-2670 Å.

Appendix C

The QSO STIS Spectra Absorption Line-lists

C.1 Line-list for HE 1122-1648

All lines used in the cross-correlation had a variance weighted EW significance ≥ 3.

Table C.1: Redshift ordered HE 1122-1648 line-list for the absorptions detected in STIS E230M spectra

Species	z	σ Z	b (km s^{-1})	σ b[a] (km s^{-1})	log(N_{HI}) (cm^{-2})	σ log(N_{HI}) (cm^{-2})	EW (Å)	σ EW (Å)	Sig
Mg II	-0.000001	0.000007	27.4	1.7	14.23	0.34	5.4E-01	8.3E-02	6.49
H I	0.696630	0.000029	26.5	0.0	14.94	0.31	3.6E-01	1.2E-01	3.09
H I	0.710265	0.000027	26.5	0.0	14.48	0.23	3.5E-01	1.0E-01	3.47
H I	0.716755	0.000032	26.5	0.0	13.96	0.16	3.4E-01	1.0E-01	3.24
H I	0.729448	0.000063	26.5	0.0	13.83	0.23	3.1E-01	8.9E-02	3.50
H I	0.743138	0.000034	26.5	0.0	13.79	0.14	3.6E-01	1.1E-01	3.40
H I	0.744602	0.000054	26.5	0.0	14.62	0.17	4.2E-01	1.0E-01	4.19
H I	0.748819	0.000021	26.5	0.0	14.41	0.17	3.4E-01	8.8E-02	3.91
H I	0.756011	0.000028	26.5	0.0	14.56	0.25	4.0E-01	1.2E-01	3.47
H I	0.765010	0.000024	26.5	0.0	13.90	0.11	3.3E-01	8.6E-02	3.86
H I	0.767598	0.000024	26.5	0.0	13.94	0.12	3.6E-01	8.6E-02	4.17
H I	0.771882	0.000061	26.5	0.0	13.72	0.15	3.2E-01	1.0E-01	3.11
H I	0.777769	0.000055	26.5	0.0	13.78	0.22	3.4E-01	1.1E-01	3.14
H I	0.778936	0.000026	26.5	0.0	14.06	0.10	3.4E-01	1.1E-01	3.07
H I	0.781680	0.000035	26.5	0.0	13.98	0.19	3.7E-01	9.0E-02	4.15
H I	0.786259	0.000016	26.5	0.0	13.89	0.06	3.8E-01	7.9E-02	4.81
H I	0.787201	0.000025	26.5	0.0	14.61	0.22	3.5E-01	7.8E-02	4.52
H I	0.792752	0.000069	26.5	0.0	14.02	0.07	3.1E-01	9.6E-02	3.26
H I	0.795778	0.000031	26.5	0.0	16.60	0.42	3.9E-01	1.0E-01	3.87
H I	0.800821	0.000036	26.5	0.0	13.99	0.16	3.3E-01	7.8E-02	4.28
H I	0.801825	0.000051	26.5	0.0	16.31	0.48	4.4E-01	7.4E-02	5.92
H I	0.804798	0.000022	26.5	0.0	13.92	0.06	3.4E-01	7.0E-02	4.86
H I	0.806409	0.000109	26.5	0.0	15.11	0.43	4.5E-01	6.9E-02	6.52
H I	0.811996	0.000033	26.5	0.0	14.00	0.17	3.8E-01	8.6E-02	4.46
H I	0.814514	0.000028	26.5	0.0	16.05	0.42	4.2E-01	8.8E-02	4.79
H I	0.817497	0.000032	26.5	0.0	13.96	0.16	3.8E-01	7.8E-02	4.85
H I	0.818814	0.000029	26.5	0.0	14.38	0.35	3.6E-01	7.8E-02	4.65
H I	0.820897	0.000023	26.5	0.0	14.24	0.16	4.3E-01	6.2E-02	6.93
H I	0.826473	0.000025	26.5	0.0	13.68	0.15	3.5E-01	6.3E-02	5.49
H I	0.834646	0.000019	26.5	0.0	13.93	0.18	3.7E-01	7.6E-02	4.82
H I	0.837759	0.000035	26.5	0.0	14.08	0.08	3.9E-01	6.9E-02	5.69
H I	0.840629	0.000020	26.5	0.0	13.77	0.08	3.8E-01	6.0E-02	6.33
H I	0.846568	0.000038	26.5	0.0	14.31	0.26	4.3E-01	5.6E-02	7.82
H I	0.847168	0.000043	26.5	0.0	14.30	0.26	4.1E-01	5.6E-02	7.31
H I	0.850822	0.000024	26.5	0.0	13.66	0.08	3.5E-01	6.7E-02	5.21

Continued on next page

---

[a] A zero error in either or both the redshift and b-parameter means that the value was fixed in VPFIT in order to constrain the error in the column density.

Table C.1 – continued from previous page

Species	z	σ z	b (km s^{-1})	σ b^{a} (km s^{-1})	log(N_{HI}) (cm^{-2})	σ log(N_{HI}) (cm^{-2})	EW (Å)	σ EW (Å)	Sig
H I	0.856721	0.000026	26.5	0.0	14.06	0.21	3.7E-01	7.2E-02	5.16
H I	0.858092	0.000016	26.5	0.0	14.10	0.27	4.1E-01	6.8E-02	6.07
H I	0.858710	0.000016	26.5	0.0	13.80	0.30	3.6E-01	6.7E-02	5.30
H I	0.884913	0.000036	26.8	17.8	15.16	0.23	5.0E-01	1.1E-01	4.40
H I	0.929168	0.000061	39.7	13.2	14.11	0.22	5.4E-01	1.2E-01	4.38
H I	0.948528	0.000070	27.2	15.0	13.69	0.24	3.2E-01	1.0E-01	3.04
H I	0.983369	0.005187	32.0	10.4	14.33	0.45	4.3E-01	1.3E-01	3.21
H I	0.984562	0.000324	26.5	0.0	14.96	0.34	4.0E-01	1.2E-01	3.28
H I	0.990582	0.000113	31.9	13.0	14.15	0.36	5.0E-01	1.2E-01	4.20
H I	1.002700	0.000132	47.6	34.1	13.96	0.19	5.3E-01	1.5E-01	3.61
H I	1.006660	0.000036	24.6	13.0	14.66	0.20	3.8E-01	1.2E-01	3.20
H I	1.007723	0.000040	66.4	8.9	14.36	0.08	9.2E-01	2.0E-01	4.74
H I	1.040552	0.000216	52.5	24.7	13.99	0.21	6.2E-01	1.4E-01	4.53
H I	1.045378	0.000031	45.3	7.2	14.32	0.11	6.7E-01	1.1E-01	5.86
H I	1.046038	0.000026	15.9	7.0	13.87	0.27	2.2E-01	7.0E-02	3.07
H I	1.050396	0.000025	43.4	6.9	14.63	0.20	7.6E-01	1.3E-01	5.98
H I	1.050988	0.000018	10.5	2.5	13.94	0.19	2.0E-01	6.3E-02	3.12
H I	1.068243	0.000034	26.5	0.0	14.76	0.28	4.9E-01	8.7E-02	5.63
H I	1.070746	0.000043	16.9	9.3	13.71	0.30	2.3E-01	6.6E-02	3.58
H I	1.071475	0.000066	31.8	14.1	13.79	0.20	3.7E-01	9.6E-02	3.88
H I	1.072811	0.000033	33.0	11.8	14.73	0.35	5.9E-01	9.6E-02	6.13
H I	1.090345	0.000121	20.3	11.9	14.04	0.27	3.1E-01	8.0E-02	3.91
H I	1.107416	0.000067	28.8	15.1	14.07	0.37	4.3E-01	1.3E-01	3.40
H I	1.112397	0.000018	18.4	10.3	15.55	0.35	3.1E-01	8.6E-02	3.60
H I	1.118118	0.000051	25.5	10.0	13.78	0.20	3.8E-01	8.4E-02	4.56
H I	1.128718	0.000075	59.4	14.9	14.35	0.14	8.5E-01	1.3E-01	6.56
H I	1.134174	0.000028	24.6	5.3	13.98	0.17	3.6E-01	1.1E-01	3.17
H I	1.141312	0.000047	21.0	10.0	13.89	0.33	3.0E-01	8.2E-02	3.70
H I	1.142771	0.000061	51.8	13.6	14.46	0.28	8.1E-01	1.2E-01	6.57
H I	1.146429	0.000067	45.3	12.3	13.94	0.13	5.4E-01	1.0E-01	5.20
H I	1.179346	0.000259	17.4	7.7	13.72	0.27	2.6E-01	5.6E-02	4.70
H I	1.200343	0.000054	69.5	9.8	14.05	0.06	8.4E-01	1.4E-01	6.17
H I	1.203080	0.000033	39.7	9.5	14.65	0.39	8.0E-01	9.2E-02	8.78
H I	1.217305	0.000102	29.3	11.4	14.04	0.26	4.8E-01	1.1E-01	4.51
H I	1.229029	0.000060	32.5	10.9	13.93	0.19	4.6E-01	1.0E-01	4.54
H I	1.233690	0.000013	20.5	5.3	14.37	0.37	3.9E-01	7.4E-02	5.32
H I	1.234207	0.000029	10.4	2.9	14.49	0.25	2.1E-01	4.9E-02	4.30
H I	1.234662	0.000032	22.6	3.5	13.96	0.09	3.6E-01	7.1E-02	5.13
H I	1.238392	0.000036	17.1	8.6	13.93	0.43	2.9E-01	6.3E-02	4.51
H I	1.242975	0.000014	10.5	10.1	13.51	0.26	2.1E-01	5.9E-02	3.50
H I	1.247300	0.000062	25.1	11.7	13.91	0.30	4.3E-01	1.1E-01	3.99
H I	1.252999	0.000057	28.2	2.9	14.76	0.24	5.1E-01	1.2E-01	4.41
H I	1.256624	0.000066	30.1	11.8	13.94	0.24	4.7E-01	1.1E-01	4.17
H I	1.265073	0.000032	25.3	8.7	14.30	0.46	4.5E-01	8.3E-02	5.43
H I	1.270146	0.000031	26.5	0.0	13.99	0.13	4.7E-01	8.5E-02	5.56
H I	1.277879	0.000107	43.8	18.5	13.83	0.19	5.0E-01	1.5E-01	3.38
H I	1.291091	0.000078	24.9	14.1	13.53	0.22	2.9E-01	9.2E-02	3.12
H I	1.298652	0.000056	38.3	10.5	14.08	0.17	6.5E-01	1.0E-01	6.52
H I	1.301080	0.000127	21.0	13.8	14.52	0.11	4.3E-01	7.6E-02	5.71
H I	1.304075	0.000023	26.5	0.0	14.21	0.12	5.1E-01	8.8E-02	5.88
H I	1.304932	0.000026	26.5	0.0	14.01	0.10	4.7E-01	9.1E-02	5.13
H I	1.319819	0.000061	25.3	12.2	14.04	0.40	4.0E-01	1.2E-01	3.37
H I	1.326740	0.000109	21.0	14.5	15.41	0.33	4.4E-01	8.2E-02	5.37
H I	1.329801	0.000030	31.7	5.2	13.99	0.11	5.0E-01	1.0E-01	4.86
H I	1.334597	0.000221	34.0	13.9	14.17	0.31	5.8E-01	1.1E-01	5.47
H I	1.341977	0.000018	26.5	0.0	15.82	0.26	4.7E-01	1.2E-01	3.88
H I	1.344336	0.000083	19.1	8.4	14.06	0.12	4.2E-01	1.2E-01	3.49
H I	1.351678	0.000038	32.1	9.1	14.27	0.29	6.6E-01	1.3E-01	5.11
H I	1.355746	0.000046	53.9	15.1	14.80	0.47	1.1E+00	1.6E-01	7.19
H I	1.364782	0.000070	37.9	11.7	13.93	0.16	5.7E-01	1.1E-01	5.08
H I	1.378067	0.000037	33.4	7.4	14.19	0.20	5.6E-01	1.7E-01	3.25
H I	1.396513	0.000036	22.6	8.5	14.16	0.42	4.2E-01	8.4E-02	5.02

Continued on next page

Table C.1 – continued from previous page

Species	z	σ z	b (km s^{-1})	σ b[a] (km s^{-1})	log(N_{HI}) (cm^{-2})	σ log(N_{HI}) (cm^{-2})	EW (Å)	σ EW (Å)	Sig
H I	1.405692	0.000052	24.6	15.9	15.22	0.25	5.6E-01	1.1E-01	5.11
H I	1.423414	0.000017	31.5	3.0	13.86	0.05	4.3E-01	1.4E-01	3.08
H I	1.428392	0.000146	27.9	9.9	13.76	0.19	4.0E-01	1.2E-01	3.45
H I	1.434985	0.000078	37.5	12.6	13.81	0.15	5.2E-01	1.4E-01	3.82
H I	1.466779	0.000048	21.9	5.3	14.47	0.20	4.9E-01	1.2E-01	4.08
H I	1.473605	0.000078	27.0	15.8	14.00	0.33	5.0E-01	1.5E-01	3.34
H I	1.506200	0.000031	39.6	5.5	14.28	0.12	6.4E-01	2.1E-01	3.04

C.2 Line-list for PKS 1127-145

Table C.2: Redshift ordered PKS 1127-145 line-list for the absorptions detected in STIS E230M spectra

Species	z	σ z	b (km s^{-1})	σ b[a] (km s^{-1})	log(N_{HI}) (cm^{-2})	σ log(N_{HI}) (cm^{-2})	EW (Å)	σ EW (Å)	Sig
Fe II	0.000079	0.000006	23.9	3.5	15.59	0.50	3.6E-01	8.9E-02	4.08
Mg II	0.000113	0.000014	49.0	5.6	13.88	0.06	9.4E-01	8.3E-02	11.29
H I	0.895887	0.000036	15.4	25.3	13.85	0.43	2.5E-01	7.2E-02	3.41
H I	0.902568	0.000032	39.0	9.8	14.50	0.35	6.7E-01	8.3E-02	8.07
H I	0.907956	0.000045	30.2	9.8	13.65	0.13	3.0E-01	7.3E-02	4.10
H I	0.910844	0.000031	24.7	14.8	14.78	0.30	4.2E-01	7.3E-02	5.80
H I	0.914341	0.000039	30.5	9.9	13.85	0.10	3.5E-01	9.3E-02	3.82
H I	0.914574	0.000097	77.5	20.2	14.02	0.11	6.8E-01	1.5E-01	4.67
H I	0.918717	0.000050	26.2	10.7	13.67	0.18	2.9E-01	8.5E-02	3.37
H I	0.920784	0.000034	26.5	0.0	13.60	0.10	2.5E-01	7.9E-02	3.19
H I	0.932106	0.000044	14.0	7.7	13.75	0.37	2.0E-01	5.0E-02	3.99
H I	0.941859	0.000046	24.1	9.8	13.63	0.17	2.6E-01	7.6E-02	3.47
H I	0.952229	0.000065	49.9	13.1	13.81	0.11	4.8E-01	9.9E-02	4.86
H I	0.955103	0.000068	26.5	0.0	13.33	0.16	2.6E-01	7.8E-02	3.28
H I	0.964353	0.000043	26.5	0.0	14.95	0.33	5.0E-01	8.7E-02	5.77
H I	1.001837	0.000075	45.4	15.1	13.65	0.13	3.9E-01	9.2E-02	4.17
H I	1.003717	0.000058	50.4	11.2	13.99	0.11	5.6E-01	1.1E-01	5.15
H I	1.020791	0.000042	20.5	9.2	13.35	0.15	1.7E-01	5.2E-02	3.30
H I	1.029523	0.000062	45.5	11.8	13.92	0.13	5.2E-01	1.1E-01	4.66
H I	1.036288	0.000089	39.7	12.7	13.55	0.32	3.4E-01	9.4E-02	3.60
H I	1.040834	0.000037	16.8	8.4	13.37	0.17	1.6E-01	5.1E-02	3.08
H I	1.044049	0.000042	57.8	8.0	14.34	0.07	9.5E-01	8.4E-02	11.32
H I	1.044873	0.000029	26.5	0.0	14.56	0.22	4.6E-01	5.7E-02	8.11
H I	1.058587	0.000033	14.8	7.7	13.73	0.33	2.0E-01	6.4E-02	3.21
H I	1.078454	0.000043	33.7	8.4	13.85	0.12	4.1E-01	8.1E-02	5.07
H I	1.090986	0.000026	13.9	6.0	13.59	0.21	1.9E-01	4.6E-02	4.18
H I	1.095878	0.000024	26.5	0.0	15.83	0.27	5.2E-01	5.4E-02	9.58
H I	1.097770	0.000039	40.1	7.8	13.93	0.08	5.3E-01	6.8E-02	7.88
H I	1.098522	0.000049	31.6	10.2	13.58	0.11	2.9E-01	6.2E-02	4.75
H I	1.103384	0.000039	35.3	7.6	14.03	0.12	5.5E-01	6.9E-02	8.01
H I	1.104456	0.000041	57.4	10.8	14.57	0.20	1.1E+00	9.1E-02	11.75
H I	1.115500	0.000022	11.6	1.8	17.42	0.20	2.1E-01	3.6E-02	5.97
H I	1.119656	0.000052	40.4	10.1	13.54	0.10	3.3E-01	5.9E-02	5.56
H I	1.121533	0.000045	27.3	9.0	13.33	0.12	2.0E-01	4.7E-02	4.16

Continued on next page

---

[a] A zero error in either or both the redshift and b-parameter means that the value was fixed in VPFIT in order to constrain the error in the column density.

Table C.2 – continued from previous page

Species	z	σ z	b (km s^{-1})	σ b[a] (km s^{-1})	log(N_{HI}) (cm^{-2})	σ log(N_{HI}) (cm^{-2})	EW (Å)	σ EW (Å)	Sig
H I	1.134350	0.000043	41.6	7.9	13.98	0.10	5.8E-01	8.4E-02	6.86
H I	1.146119	0.000026	39.4	4.9	13.95	0.06	5.5E-01	5.1E-02	10.92
H I	1.151076	0.000070	79.7	13.0	13.68	0.06	4.6E-01	6.2E-02	7.53
H I	1.160740	0.000035	20.8	7.2	13.26	0.11	1.6E-01	3.8E-02	4.14
H I	1.170093	0.000027	20.9	5.6	13.18	0.09	1.4E-01	2.5E-02	5.49
H I	1.188028	0.000026	59.8	4.8	13.69	0.03	4.6E-01	3.0E-02	15.16

Appendix D

The Metal and Associated HI lines

D.1 Metal Line-list for Carbon IV

All lines used in the metallicity calculation had a variance weighted EW significance ≥ 3.

Table D.1: CIV and paired HI lines from the UVES and STIS Spectra

Species	z	log(N_{HI}) (cm^{-2})	EW (Å)	Field	Metallicity $Z_\odot$
C IV	1.007919 ± 0.000009	14.16 ± 0.31	0.12 ± 0.01	HE 1122-1648	$-0.39^{+0.47}_{-0.24}$
H I	1.007723 ± 0.000040	14.36 ± 0.08	0.92 ± 0.20		
C IV	1.045339 ± 0.000004	13.17 ± 0.03	0.09 ± 0.01	HE 1122-1648	$-1.23^{+0.10}_{-0.13}$
H I	1.045378 ± 0.000031	14.32 ± 0.11	0.67 ± 0.11		
C IV	1.047084 ± 0.000006	12.39 ± 0.08	0.02 ± 0.00	HE 1122-1648	$-1.69^{+0.22}_{-0.39}$
H I	1.046038 ± 0.000026	13.87 ± 0.27	0.22 ± 0.07		
C IV	1.067739 ± 0.000001	13.55 ± 0.01	0.14 ± 0.00	HE 1122-1648	$-1.33^{+0.21}_{-0.40}$
H I	1.068243 ± 0.000034	14.76 ± 0.28	0.49 ± 0.09		
C IV	1.068022 ± 0.000005	13.60 ± 0.07	0.17 ± 0.00	HE 1122-1648	$-1.28^{+0.22}_{-0.40}$
H I	1.068243 ± 0.000034	14.76 ± 0.28	0.49 ± 0.09		
C IV	1.068145 ± 0.000008	13.83 ± 0.04	0.30 ± 0.00	HE 1122-1648	$-1.26^{+0.21}_{-0.40}$
H I	1.068243 ± 0.000034	14.76 ± 0.28	0.49 ± 0.09		
C IV	1.068692 ± 0.000003	13.10 ± 0.03	0.10 ± 0.00	HE 1122-1648	$-1.82^{+0.21}_{-0.40}$
H I	1.068243 ± 0.000034	14.76 ± 0.28	0.49 ± 0.09		
C IV	1.233794 ± 0.000021	12.57 ± 0.33	0.08 ± 0.00	HE 1122-1648	$-1.97^{+0.55}_{-0.63}$
H I	1.233690 ± 0.000013	14.37 ± 0.37	0.39 ± 0.07		
C IV	1.233847 ± 0.000000	13.61 ± 0.03	0.17 ± 0.00	HE 1122-1648	$-0.86^{+0.25}_{-0.59}$
H I	1.233690 ± 0.000013	14.37 ± 0.37	0.39 ± 0.07		
C IV	1.234071 ± 0.000043	13.85 ± 0.39	0.35 ± 0.00	HE 1122-1648	$-0.68^{+0.65}_{-0.42}$
H I	1.234207 ± 0.000029	14.49 ± 0.25	0.21 ± 0.05		
C IV	1.352006 ± 0.000028	13.23 ± 0.11	0.20 ± 0.00	HE 1122-1648	$-0.37^{+0.25}_{-0.42}$
H I	1.351678 ± 0.000038	14.27 ± 0.29	0.66 ± 0.13		
C IV	1.352190 ± 0.000013	13.48 ± 0.09	0.26 ± 0.00	HE 1122-1648	$-0.68^{+0.23}_{-0.42}$
H I	1.351678 ± 0.000038	14.27 ± 0.29	0.66 ± 0.13		
C IV	1.352209 ± 0.000005	13.01 ± 0.41	0.15 ± 0.00	HE 1122-1648	$-1.36^{+0.71}_{-0.49}$
H I	1.351678 ± 0.000038	14.27 ± 0.29	0.66 ± 0.13		
C IV	1.352521 ± 0.000005	12.96 ± 0.02	0.09 ± 0.00	HE 1122-1648	$-1.15^{+0.21}_{-0.41}$
H I	1.351678 ± 0.000038	14.27 ± 0.29	0.66 ± 0.13		

Continued on next page

Table D.1 – continued from previous page

Species	z	log(N_{HI}) (cm^{-2})	EW (Å)	Field	Metallicity $Z_{\odot}$
C IV	1.364664 ± 0.000003	13.06 ± 0.01	0.09 ± 0.00	HE 1122-1648	$-0.42^{+0.13}_{-0.19}$
H I	1.364782 ± 0.000070	13.93 ± 0.16	0.57 ± 0.11		
C IV	1.364914 ± 0.000006	12.38 ± 0.05	0.02 ± 0.00	HE 1122-1648	$-1.43^{+0.14}_{-0.20}$
H I	1.364782 ± 0.000070	13.93 ± 0.16	0.57 ± 0.11		
C IV	1.379066 ± 0.000002	12.76 ± 0.05	0.05 ± 0.00	HE 1122-1648	$-1.43^{+0.17}_{-0.26}$
H I	1.378067 ± 0.000037	14.19 ± 0.20	0.56 ± 0.17		
C IV	1.641839 ± 0.000100	12.97 ± 0.34	0.10 ± 0.00	HE 1122-1648	$-1.11^{+0.52}_{-0.24}$
H I	1.641854 ± 0.000004	14.95 ± 0.06	0.79 ± 0.01		
C IV	1.641842 ± 0.000003	12.64 ± 0.15	0.03 ± 0.00	HE 1122-1648	$-2.46^{+0.19}_{-0.15}$
H I	1.641854 ± 0.000004	14.95 ± 0.06	0.79 ± 0.01		
C IV	1.642030 ± 0.000006	13.16 ± 0.14	0.09 ± 0.00	HE 1122-1648	$-1.95^{+0.17}_{-0.14}$
H I	1.641854 ± 0.000004	14.95 ± 0.06	0.79 ± 0.01		
C IV	1.724331 ± 0.000004	12.46 ± 0.03	0.04 ± 0.00	HE 1122-1648	$-2.45^{+0.05}_{-0.05}$
H I	1.724331 ± 0.000000	14.90 ± 0.04	0.83 ± 0.00		
C IV	1.724449 ± 0.000007	12.01 ± 0.07	0.02 ± 0.00	HE 1122-1648	$-3.02^{+0.08}_{-0.08}$
H I	1.724331 ± 0.000000	14.90 ± 0.04	0.83 ± 0.00		
C IV	2.007109 ± 0.000000	13.10 ± 0.01	0.15 ± 0.00	HE 1122-1648	$-2.56^{+0.09}_{-0.11}$
H I	2.007109 ± 0.000004	15.44 ± 0.10	0.89 ± 0.00		
C IV	2.007205 ± 0.000000	12.31 ± 0.07	0.08 ± 0.00	HE 1122-1648	$-2.94^{+0.24}_{-0.48}$
H I	2.007205 ± 0.000000	15.10 ± 0.32	0.71 ± 0.00		
C IV	2.030096 ± 0.000000	12.27 ± 0.04	0.02 ± 0.00	HE 1122-1648	$0.64^{+0.05}_{-0.05}$
H I	2.029300 ± 0.000021	12.97 ± 0.02	0.17 ± 0.00		
C IV	2.033133 ± 0.000000	12.05 ± 0.00	0.02 ± 0.00	HE 1122-1648	$-2.13^{+0.01}_{-0.01}$
H I	2.033133 ± 0.000001	14.68 ± 0.01	1.01 ± 0.00		
C IV	2.064414 ± 0.000044	12.56 ± 0.08	0.08 ± 0.00	HE 1122-1648	$-0.13^{+0.17}_{-0.24}$
H I	2.064414 ± 0.000000	14.24 ± 0.19	1.16 ± 0.00		
C IV	2.064787 ± 0.000016	12.52 ± 0.07	0.06 ± 0.00	HE 1122-1648	$-0.14^{+0.30}_{-0.84}$
H I	2.064787 ± 0.000000	14.05 ± 0.47	1.86 ± 0.00		
C IV	2.065682 ± 0.000000	12.20 ± 0.04	0.02 ± 0.00	HE 1122-1648	$0.53^{+0.21}_{-0.39}$
H I	2.065682 ± 0.000014	13.34 ± 0.28	0.84 ± 0.00		
C IV	2.080162 ± 0.000000	12.10 ± 0.06	0.02 ± 0.00	HE 1122-1648	$-1.88^{+0.10}_{-0.10}$
H I	2.080162 ± 0.000000	14.70 ± 0.08	0.71 ± 0.00		
C IV	2.080527 ± 0.000000	11.86 ± 0.09	0.02 ± 0.00	HE 1122-1648	$-1.61^{+0.17}_{-0.20}$
H I	2.080527 ± 0.000000	14.24 ± 0.15	0.54 ± 0.00		
C IV	2.101764 ± 0.000000	12.69 ± 0.02	0.07 ± 0.00	HE 1122-1648	$-1.01^{+0.14}_{-0.21}$
H I	2.101764 ± 0.000006	14.88 ± 0.17	1.01 ± 0.00		
C IV	2.206459 ± 0.000000	11.96 ± 0.10	0.02 ± 0.00	HE 1122-1648	$-3.41^{+0.12}_{-0.09}$
H I	2.206459 ± 0.000000	15.13 ± 0.01	1.00 ± 0.00		
C IV	2.215271 ± 0.000000	11.87 ± 0.06	0.01 ± 0.00	HE 1122-1648	$-2.92^{+0.07}_{-0.07}$
H I	2.215271 ± 0.000000	14.71 ± 0.04	0.84 ± 0.00		
C IV	2.215494 ± 0.000000	12.11 ± 0.06	0.03 ± 0.00	HE 1122-1648	$-0.72^{+0.15}_{-0.20}$
H I	2.215494 ± 0.000000	14.34 ± 0.16	1.19 ± 0.00		

Continued on next page

Table D.1 – continued from previous page

Species	z	log(N_{HI}) (cm^{-2})	EW (Å)	Field	Metallicity $Z_\odot$
C IV	2.339260 ± 0.000000	11.93 ± 0.04	0.01 ± 0.00	HE 1122-1648	$0.61^{+0.08}_{-0.09}$
H I	2.337401 ± 0.000029	12.87 ± 0.07	0.01 ± 0.00		
C IV	2.352170 ± 0.000000	12.38 ± 0.02	0.03 ± 0.00	HE 1122-1648	$-1.87^{+0.04}_{-0.05}$
H I	2.352170 ± 0.000005	14.41 ± 0.04	0.71 ± 0.00		
C IV	2.352464 ± 0.000007	12.15 ± 0.06	0.02 ± 0.00	HE 1122-1648	$-2.85^{+0.20}_{-0.34}$
H I	2.352695 ± 0.000000	14.87 ± 0.25	1.59 ± 0.00		
C IV	2.352762 ± 0.000005	12.80 ± 0.02	0.09 ± 0.00	HE 1122-1648	$-1.86^{+0.19}_{-0.34}$
H I	2.352695 ± 0.000000	14.87 ± 0.25	1.59 ± 0.00		
C IV	2.353383 ± 0.000016	12.67 ± 0.03	0.08 ± 0.00	HE 1122-1648	$1.00^{+0.08}_{-0.10}$
H I	2.353946 ± 0.000072	13.74 ± 0.09	0.91 ± 0.00		
C IV	2.354078 ± 0.000003	12.23 ± 0.02	0.02 ± 0.00	HE 1122-1648	$-0.66^{+0.10}_{-0.13}$
H I	2.354095 ± 0.000007	13.44 ± 0.11	0.44 ± 0.00		
C IV	1.057877 ± 0.000039	12.31 ± 0.19	0.03 ± 0.01	PKS 1127-145	$-1.08^{+0.33}_{-0.51}$
H I	1.058587 ± 0.000033	13.73 ± 0.33	0.20 ± 0.06		
C IV	1.095874 ± 0.000101	13.00 ± 0.50	0.04 ± 0.01	PKS 1127-145	$-2.52^{+0.96}_{-0.47}$
H I	1.095878 ± 0.000024	15.83 ± 0.27	0.52 ± 0.05		
C IV	1.098493 ± 0.000008	13.21 ± 0.06	0.05 ± 0.00	PKS 1127-145	$-0.36^{+0.12}_{-0.14}$
H I	1.098522 ± 0.000049	13.58 ± 0.11	0.29 ± 0.06		

D.2 Metal Line-list for Oxygen VI

All lines used in the metallicity calculation had a variance weighted EW significance ≥ 3.

Table D.2: OVI and paired HI lines from the UVES and STIS Spectra

Species	z	log(N_{HI}) (cm^{-2})	EW (Å)	Field	Metallicity $Z_\odot$
O VI	2.007109 ± 0.000000	13.60 ± 0.00	0.13 ± 0.01	HE 1122-1648	$-2.93^{+0.09}_{-0.11}$
H I	2.007109 ± 0.000004	15.44 ± 0.10	0.89 ± 0.00		
O VI	2.007205 ± 0.000005	13.42 ± 0.08	0.07 ± 0.01	HE 1122-1648	$-2.18^{+0.24}_{-0.48}$
H I	2.007205 ± 0.000000	15.10 ± 0.32	0.71 ± 0.00		
O VI	2.030096 ± 0.000000	13.40 ± 0.07	0.14 ± 0.01	HE 1122-1648	$-0.38^{+0.08}_{-0.07}$
H I	2.029300 ± 0.000021	12.97 ± 0.02	0.17 ± 0.00		
O VI	2.033133 ± 0.000000	13.68 ± 0.03	0.15 ± 0.01	HE 1122-1648	$-2.38^{+0.03}_{-0.03}$
H I	2.033133 ± 0.000001	14.68 ± 0.01	1.01 ± 0.00		
O VI	2.064414 ± 0.000000	13.60 ± 0.00	0.10 ± 0.01	HE 1122-1648	$-0.84^{+0.15}_{-0.23}$
H I	2.064414 ± 0.000000	14.24 ± 0.19	1.16 ± 0.00		
O VI	2.064787 ± 0.000000	13.50 ± 0.00	0.09 ± 0.01	HE 1122-1648	$-1.64^{+0.29}_{-0.83}$

Continued on next page

Table D.2 – continued from previous page

Species	z	$\log(N_{HI})$ (cm^{-2})	EW (Å)	Field	Metallicity $Z_\odot$
H I	2.064787 ± 0.000000	14.05 ± 0.47	1.86 ± 0.00		
O VI	2.065682 ± 0.000000	13.03 ± 0.10	0.10 ± 0.01	HE 1122-1648	$-1.00^{+0.24}_{-0.40}$
H I	2.065682 ± 0.000014	13.34 ± 0.28	0.84 ± 0.00		
O VI	2.080162 ± 0.000018	13.50 ± 0.05	0.14 ± 0.01	HE 1122-1648	$-2.68^{+0.09}_{-0.10}$
H I	2.080162 ± 0.000000	14.70 ± 0.08	0.71 ± 0.00		
O VI	2.080527 ± 0.000021	13.35 ± 0.07	0.11 ± 0.01	HE 1122-1648	$-2.17^{+0.15}_{-0.20}$
H I	2.080527 ± 0.000000	14.24 ± 0.15	0.54 ± 0.00		
O VI	2.101764 ± 0.000000	14.13 ± 0.02	0.41 ± 0.01	HE 1122-1648	$-1.69^{+0.14}_{-0.21}$
H I	2.101764 ± 0.000006	14.88 ± 0.17	1.01 ± 0.00		
O VI	2.206459 ± 0.000005	13.12 ± 0.10	0.08 ± 0.00	HE 1122-1648	$-2.57^{+0.12}_{-0.09}$
H I	2.206459 ± 0.000000	15.13 ± 0.01	1.00 ± 0.00		
O VI	2.215271 ± 0.000005	13.08 ± 0.08	0.07 ± 0.00	HE 1122-1648	$-2.49^{+0.10}_{-0.09}$
H I	2.215271 ± 0.000000	14.71 ± 0.04	0.84 ± 0.00		
O VI	2.215494 ± 0.000031	13.35 ± 0.16	0.19 ± 0.01	HE 1122-1648	$-1.21^{+0.24}_{-0.23}$
H I	2.215494 ± 0.000000	14.34 ± 0.16	1.19 ± 0.00		
O VI	2.339260 ± 0.000007	12.99 ± 0.05	0.04 ± 0.00	HE 1122-1648	$-0.52^{+0.08}_{-0.09}$
H I	2.337401 ± 0.000029	12.87 ± 0.07	0.01 ± 0.00		
O VI	2.352170 ± 0.000000	12.77 ± 0.13	0.23 ± 0.00	HE 1122-1648	$-2.74^{+0.15}_{-0.12}$
H I	2.352170 ± 0.000005	14.41 ± 0.04	0.71 ± 0.00		
O VI	2.352695 ± 0.000015	13.97 ± 0.01	0.59 ± 0.00	HE 1122-1648	$-1.32^{+0.19}_{-0.34}$
H I	2.352695 ± 0.000000	14.87 ± 0.25	1.59 ± 0.00		
O VI	2.369205 ± 0.000005	13.02 ± 0.06	0.10 ± 0.00	HE 1122-1648	$-2.43^{+0.07}_{-0.06}$
H I	2.369161 ± 0.000009	15.43 ± 0.03	1.06 ± 0.00		
O VI	2.388362 ± 0.000000	13.02 ± 0.14	0.18 ± 0.00	HE 1122-1648	$-0.30^{+0.31}_{-0.73}$
H I	2.388362 ± 0.000085	13.57 ± 0.42	1.12 ± 0.00		