ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Quasar Absorption Spectra: Probes of the Baryonic Gas at High Redshift

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Valentina D'Odorico

SUPERVISORS

Prof. Luigi Danese

Prof. Stefano Cristiani

October 1999

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

TRIESTE
Via Beirut 2-4

TRIESTE

SISSA ISAS
SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Quasar Absorption Spectra: Probes of the Baryonic Gas at High Redshift

Thesis submitted for the degree of
"Doctor Philosophiæ"

CANDIDATE

Valentina D'Odorico

SUPERVISORS

Prof. Luigi Danese
Prof. Stefano Cristiani

October 1999

Contents

List of Figures

Acknowledgments

Firts of all my thank goes to Stefano Cristiani, he always supported me with trust and friendship, and I owe mainly to him my formation as a scientist. Then I am grateful to the staff of the astrophysics sector in SISSA, and in particular to Gigi Danese.

I wish to thank Andrea Ferrara, Adriano Fontana, Emanuele Giallongo, Paolo Molaro, Sandra Savaglio and Giovanni Vladilo: I enjoyed their friendship and I was scientifically enriched by their knowledge. The three months I spent this year as a visitor at the Institut d'Astrophysics de Paris were very fruitful for me, especially thanks to Patrick Petitjean, I learned a lot from the discussions we had.

I cannot forget the people of the SISSA secretariat, Alex, Andrea, Claudia, Cristina, Giuliana and Riccardo, and the system managers, Davide, Fabio, Luisa, and Marina for their efficiency and availability.

Alice, Antonella, Benedetta, Bruce, Catia, Cecilia, Chiara, Christian, Cristiano, Daniele, Enrica, Emanuela, Ezio, Fabio, Fabio, Francesca, Gaspare, Gianluca, Giovanni, Giulietta, Giuliano, Ilaria, Joachim, Laura, Lorenzo, Manuela, Marco, MariaStella, Marta, Massimo, Matteo, Matteo, Patrizia, Pierluigi, Rodica, Ujin, Vanessa, and those that I forgot, I got something from all of you, I hope that the links of friendship will be stronger than time and distance.

Anna, Barbara, Cecilia, Diego, Elena, Elisabetta, Ettore, Fabrizio, Giovanni, Goran, Laura, Marcella, Mathias, Montse, Patrizia, Sebastien, Shashi, Simone, Stefano, Steve, Thomas were my companions during three intense, wonderful months in Paris.

I will always be indebted to Paola and Stefano, you fed my spirit and my stomach, you were always there when I needed you, a friendship like yours is rare and precious and I hope it will last forever!

A last thought goes to my parents, in different ways but both with love, they were always at my side.

Chapter 1

Introduction

1.1 Structure of the Thesis

The thesis is planned in the following way. This chapter introduces the main features of the spectrum of Active Galactic Nuclei (AGN) and, in particular, of quasars. A schematic description of the unified model for AGN and quasars is also given. In Chapter 2, I focus the attention on the narrow absorption lines in QSO spectra. I present the main issues related to the observation and the interpretation of the lines. Then, I report how absorption systems are in general classified and what are the potentialities of such a study. Chapter 3 is dedicated to a review of the state of the art of the observational properties of absorption systems. We contributed to this topic adding to the statistics the results from the analysis of the high redshift quasar PKS 2126-158 ($z_{\rm em} = 3.27$) high resolution spectrum: they are described in § 3.7.

The main interest of my PhD work has been devoted to the study of the spatial distribution and of the dimension of QSO absorbers. We based our analysis on two large samples of data: (i) a collection of high resolution Lyα forests for the study of clustering, (ii) the whole QSO pair spectra sample at intermediate to high resolution available in the literature, for the study of the transverse dimension and geometry. We stated once and for all the presence of clustering of Lyα absorbers along a single line of sight; and we confirmed the large size of absorbers with a more reliable estimating algorithm. These results, together with a review of pre-

vious and following studies on the same topic, are displayed in Chapters 4 and 5.

In Chapter 6, I sketch the outline of the theoretical models proposed in the course of the years for the Lyα clouds and, briefly, for metal absorbers. In the second part of the chapter, I dedicate more space to a review of the results obtained by the numerical and semi-analytical simulations.

Finally, I sum up our results and especially, I trace the lines of future works in Chapter 7.

Most of the review sections are based on the following sources: Weedman 1986; Frank, King & Raine 1992; Emerson 1996; Petitjean 1998; Rauch 1998b; Weinberg et al. 1998; Kembhavi & Narlikar 1999.

All through the thesis, if not otherwise specified, I will use the following values for the cosmological parameters, $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$, $q_0 = 0.5, \Lambda = 0, \Omega_0 = 1$.

1.2 The Discovery of Quasars

The discovery of quasi-stellar objects (QSOs or quasars) in 1963 made a new science out of astronomy. There were two factors in this invigorating revolution. One was the conceptual shock of learning that some very important sources of energy exist in the universe that are not related to the nuclear fusion processes in stars. The other was the fact that the discovery was made with radio astronomy. For the theorist, there was suddenly an open season for wide ranging and creative speculations on cosmological processes, energy generation, and radiation physics. For the technologist, there was proof that the coordinate use of different observational spectral windows could reveal extraordinary and totally unanticipated things.

1.3 The Spectrum of Quasars

1.3.1 Continuum

Quasars are unique among objects in the universe in the observable span of their continuous spectra. In some cases, the same quasar can be seen at wavelengths from at least $\sim 10^8$ Hz in the radio region to $\sim 10^{27}$ Hz, which corresponds

to extremely high energy γ-ray photons. This vast range spans many different physical processes and the emitting region can vary in size from the $\sim$ Mpc scale of extended radio emission to the scale of $\sim 10^2$ light seconds on which high energy radiation is emitted.

The continuum spectrum of quasars is often, over limited frequency ranges, represented by simple power law forms. The luminosity as a function of frequency $L(\nu)$ is then given by

$$L(\nu) = A\,\nu^{-\alpha}, \tag{1.1}$$

where A and α are constant.

Data presently available show that there are basically two kinds of continuum shape for quasars and AGN.

1. Blazars, which include quasars and BL Lacs with bright radio cores and high degrees of polarization and variability, indicating bulk relativistic motion in these cores, have non-thermal continuum radiation from the radio to the ultraviolet band, possibly produced by synchrotron emission. The non-thermal emission continues in the high energy region, where it can be affected by the effects of inverse Compton scattering and pair production.
2. In radio-quiet quasars, Seyfert galaxies and even broad line radio galaxies the dominant emission is in the $\sim 0.01 - 1$ μm region. It is of thermal origin and is produced in an accretion disc or as free-free optically thin emission. Emission in the $\sim 1 - 100$ μm region, which is the second most dominant component, is produced as thermal emission from dust with a wide range of temperatures. Emission in the radio region, when detectable, has a steep spectrum and is of non-thermal origin. X-ray emission is often detectable in these objects and could have a strong non-thermal component.

In this thesis, I will concentrate on the spectrum of thermally dominated objects (see Fig. 1.1). Indeed, a fraction of thermally dominated quasars ($\sim 10 - 20$ %) are actually radio-loud at the level of a few mJy, while others remain radio-quiet even when probed at the μJy level. Objects detected in radio surveys usually have steep radio spectra indicative of optically thin synchrotron emission. The extrapolation of the radio spectrum to infrared wavelengths lies far below the actual

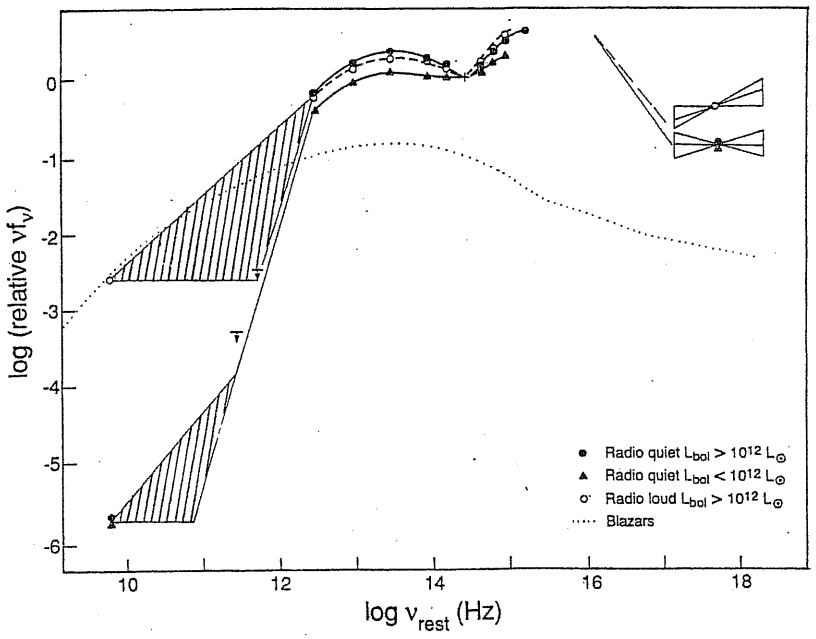

Figure 1.1: Average continuum energy distributions for radio-quiet and radio-loud quasars from the sample of Sanders et al. (1989) which is a subset of the PBQ survey. Each data point represents the mean over the corresponding sample. The X-ray slopes are adopted from Wilkes and Elvis (1987). The hatched region at radio-submillimeter wavelength represents the range of spectral indices observed. The infrared bump and the big blue bump are clearly seen. Shown for comparison (*dotted line*) is the mean continuum energy distribution for blazars (Impey & Neugebauer 1988), normalized at 6 cm. Reproduced from Sanders et al. (1989)

observed flux, indicating that there are other processes that become dominant in that region.

In the infrared region, the emitted power per unit logarithmic interval $\nu L(\nu)$ rises sharply between 10 and 100 μm and peaks somewhere in this range. The continuum in the 2-300 μm region is known as the *infrared bump* and typically contains a third of the total detected (i.e., bolometric) luminosity of the source, though this fraction can be as high as $\sim$ 90 %. This continuum portion in radio-quiet quasars is dominated by thermal emission from dust grains. Radio-loud quasars can have substantial contribution from thermal as well as non-thermal processes. There is a local minimum in $\nu L(\nu)$ at $\sim$ 1 μm that can be explained

by the opposite falls off of the emission from the accretion disc and from the dust, which sublimates at temperatures above $\sim 1500-2000$ K. Shortwards of 1 μm the continuum rises into the optical and ultraviolet, with a prominent peak in $\nu L(\nu)$, called the *big blue bump* (BBB) or the *UV excess*, occurring between ~ 0.3 μm and ~ 10 nm ($\sim 10^{15} - 3 \times 10^{16}$ Hz). In some objects the peak power occurs in the $10.0 - 122.0$ nm region, while in others $\nu L(\nu)$ continues to rise into the extreme ultraviolet. The magnitude of the energy in the bump is an indication that it has primary origin, in the sense that it is not high energy radiation that has been reprocessed to lower energy through thermalization. If the IR power law, with index ~ 1, extends to higher energies and is subtracted from the UV flux, then the residual has the form $F(\nu) \sim \nu^{1/3}$ which is remarkably similar to the shape of the continuum expected from a geometrically thin but optically thick accretion disc (Shields 1978). This has led to detailed models for the emission of the BBB from an accretion disc. An alternative class of models interprets the BBB as produced by optically thin thermal bremsstrahlung (free-free).

Riding on the big blue bump is a subsidiary "300 nm bump", which extends from 180 to ~ 400 nm. This is made up of contributions from (1) the Balmer continuum, (2) a pseudocontinuum formed from higher order Balmer lines that blend into each other and contribute at longer wavelengths than the Balmer continuum limit of 364.6 nm and (3) a pseudocontinuum due to a forest of iron lines, mainly Fe II, with the contribution occurring in the $\sim 180 - 350$ nm region.

Sanders et al. (1989; see also Elvis et al. 1994) have discussed continuum observations, in the $\sim$ 0.3 nm - 6 cm (5 GHZ - 10^{18} Hz) range of 105 quasars from the Palomar Bright Quasar (PBQ) survey. They find that the strenghts of the infrared and big blue bump do not show any correlation with the bolometric luminosity. The ratio of the bump luminosities $L_{\rm IR}/L_{\rm UV}$ has a spread of a factor ~ 10 and $< L_{\rm IR}/L_{\rm UV} >= 0.4$ with a spread of 0.15. In contrast, the ratio of the radio to the 1 μm luminosity has a spread of a factor $\sim 10^5$. This has clearly a bimodal distribution. Sanders et al. define the boundary between radio-loud and radio-quiet quasars at

$$\nu L(\nu)(5\ \mathrm{GHz})/\nu L(\nu)(1\ \mu\mathrm{m}) = 10^{-4}. \qquad (1.2)$$

Observations in the extreme ultraviolet range (EUV) between the Lyman limit

of 91.2 nm and $\sim$ 10 nm are difficult due to the absorption of the Galactic interstellar medium and can only be made through specific low opacity windows. The continuum in this region is responsible for ionisation of the broad line and narrow line emitting region, and therefore some idea of its shape can be obtained from the study of emission lines.

The spectra in the X-ray region, up to a few keV, are of power law form, with radio-loud quasars having flatter spectra than the radio-quiet ones. Beyond this range and up to a few tens of keV the spectra have an approximate power law shape with a canonical spectral index of $\sim$ 0.7. Recent observations have shown that Seyferts and radio-quiet quasars are not strong emitters above some hundreds of keV, while the radio-loud blazars often show high energy γ-ray emission.

1.3.2 Emission Lines

At a variance with most galaxies, the visible spectra of quasars are characterized by strong broad emission lines.

Emission lines can provide diagnostics of velocities, temperatures and densities unavailable from any other technique. The lines which can be seen represent a wide range of ionisation, so line fluxes also provide indirect measurements of unobserved portions of the continuum (e.g., the already mentioned far ultraviolet portion). Not least is the fact that emission lines are spectroscopically conspicuous, calling attention to locations where unusual events are occurring.

The emission lines represent re-radiation by individual ions of energy derived from a more primary continuum. We deal with two fundamental categories of emission lines: those produced by decays occurring during ionic recombination following ionisation, and those produced by decays following collisional excitations by other particles in the gas. Recombinations proceed at a low rate because of the low density, thus they are detectable only for the most abundant elements, hydrogen, helium and, sometimes, carbon. Collisionally excited lines from the first excited levels to the ground states are called *resonance lines*. Examples include the C IV doublet at 154.8 nm and 155.0 nm, the Mg II doublet at 279.8 nm and 280.2 nm, and the H I Lyman α, which is both collisionally excited and a recombination line. Other collisionally excited lines are *forbidden lines*, like the

doublet [O II] 372.6 nm and 372.9 nm.

A further classification of the emission lines observed in QSO spectra is between the *broad* lines with full width at half maximum (FWHM) of the order of 1000 to 5000 km s^{-1} and the *narrow* lines with FWHM of the order of 500 km s^{-1}. Forbidden transitions are only found as narrow lines whereas permitted lines may have both broad and narrow components. Lines that are collisionally excited and then decay radiatively have intensities that are proportional to the square of the density as do recombination lines. However, if a collisionally excited line normally decays by collisional de-excitation and only occasionally by radiative de-excitation (so the level population tends towards its Boltzmann equilibrium value), the intensity of the line will only be proportional to the density. Hence at high densities, forbidden lines are weak compared with both permitted collisionally excited lines and recombination lines. It must be concluded that the narrow lines in AGN come from a lower velocity and density region than the broad lines, respectively the *narrow line region* (NLR) and the *broad line region* (BLR).

In the following, a brief description of the characteristics of these two regions is given, as can be gained from the analysis of the corresponding emission lines.

For the NLR, the line ratios can be used to determine the kinetic temperatures, by exploiting the fact that the relative collisional excitation rates of levels with considerably different excitation potentials are strongly dependent on temperatures. The density can be obtained from lines corresponding to closely spaced upper levels. These are clearly insensitive to temperature but depend on the extent of collisional de-excitation. Typically, temperatures in the range 10^4 to 2×10^4 K, and densities in the range 10^3 to 1.6×10^4 cm^{-3}, are found. Considerations on the observed ions and estimated temperatures lead to the conclusion that the NLR is photoionised.

The Balmer line decrement (i.e., the ratio $I(\mathrm{H}\alpha) : I(\mathrm{H}\beta) : I(\mathrm{H}\gamma)$) in the NLR is steeper than predicted by recombination theory, implying reddening by dust. Analysis of line strengths using detailed ionisation models suggest abundances similar or probably larger than solar.

The volume of ionised gas can be determined as follows (Emerson 1996). The total number of recombinations equals the total number of ionising photons. We know that a fraction f_β of all Balmer recombination photons are Hβ photons (from

the Balmer decrement). Hence if the volume of the gas is V:

$$\alpha_{\rm rec}\, N_{\rm e}\, N({\rm H}^+)\, V = (\text{number of H}\beta \text{ photons})/f_\beta. \tag{1.3}$$

The flux $F(\mathrm{H}\beta)$ equals the Hβ luminosity, $L(\mathrm{H}\beta)$, divided by $4\pi d^2$, where d is the distance of the cloud. $L(\mathrm{H}\beta)$ is the number of Hβ photons produced in the whole cloud per second multiplied by the energy of an Hβ photon, $h\nu_\beta$. Hence

$$V = 1.2\,\frac{F(\mathrm{H}\beta)4\pi d^2}{h\,\nu_\beta\,\alpha_{\rm rec}\,f_\beta\,N_{\rm e}^2}, \tag{1.4}$$

where, since the gas is fully ionised, $N_{\rm e} = N(\mathrm{H}) + 2N(\mathrm{He})$ with $N(\mathrm{H}) = N(\mathrm{H}^+)$ and for normal abundances $N(\mathrm{He}) = 0.1N(\mathrm{H})$. The mass is given by

$$\begin{aligned} M &= [N(\mathrm{H}) + 4N(\mathrm{He})]m_{\rm H}V = 1.4N(\mathrm{H})m_{\rm H}V \\ &= \frac{1.4}{1.2}\,N_{\rm e}m_{\rm H}V. \end{aligned} \tag{1.5}$$

Masses of the order 10^5 $\mathrm{M}_\odot$ and volumes of 3×10^4 pc^3 are typically obtained.

Direct observations of some nearby Seyfert 2 galaxies have given radii of the overall NLR of several hundred parsecs, and there is some evidence that the radii of NLR may be considerably larger for QSOs. It follows that only a small fraction of the NLR, less than 1 %, actually contains gas of any significant density. An arrangement of the gas in a number of individually optically thick clouds in a largely empty $\sim$ 100 parsec radius sphere seems the more likely scenario, i.e., the filling factor of the NLR is $\sim 10^{-3}$ or less.

The BLR spectrum is dominated, in the visible, by the Balmer lines of H I and the lines of He II, and sometimes the lines of Fe II. In the UV the lines of Mg II at 279.8 nm are strong and C III] at 190.9 nm is often present, while at shorter wavelengths we find the Lyman lines and a variety of collisionally excited lines, those of C IV at 154.9 nm being particularly conspicuous, with other strong lines including Si IV at 140.0 nm, N V at 124.0 nm, and O VI at 103.5 nm.

From the absence of forbidden lines and the presence of some semi-forbidden ones, a representative density of about 10^9 cm^{-3} is found for the BLR.

The fact that iron is present in singly ionised form means that the temperature of the BLR is less than 3.5×10^4 K. The absence of any C III λ97.7 nm emission and the presence of strong C III] λ190.9 nm is the most compelling evidence that the broad emission lines arise from photoionised gas. Infact, collisional excitation would populate significantly the upper level of C III λ97.7 nm but would need a temperature $T \gtrsim 3 \times 10^5$ K. Photoionisation is the only way in which high ionisation states can be produced at low temperatures. The ratio of the C III line strengths gives an upper limit of 2.5×10^4 K for the temperature of the BLR. Eventually, a temperature of 2×10^4 K would seem reasonable for the BLR of most AGN.

Further evidence for photoionisation comes from a general overall correlation between the strenghts of the broad lines and in particular of the hydrogen broad lines with the overall level of the continuum flux (extrapolated to the ultraviolet). In individual sources, time variations of the continuum do not always correlate well with line variability, even allowing for light travel time delays. This may be taken to indicate complex dependence of the line-emitting gas on the continuum power.

Direct observations of NGC 4151 shows that the BLR lies within 0.025 pc of the central source. Overall, the radial distances scale with luminosity, such that $r \approx 0.1(L/10^{46}\ \mathrm{ergs}^{-1})^{1/2}$ (Peterson 1993).

The observation of the strong broad lines of Fe II and Mg II, whose ionisation potentials are similar to that of H I, and the absence of a Lyman discontinuity in the spectra of high redshift AGN can be reconciled if the BLR clouds are optically thick and therefore ionisation bounded, but only block a small fraction of the continuum source, that is the covering factor of the BLR is low. If the Balmer lines are mainly produced by recombination, broad component Balmer line luminosities together with electron density give the total volume and mass of the broad line gas by the same method that was employed in the case of the NLR. Typical values are 3×10^6 pc^3 and 50 solar masses for Seyfert 1 galaxies and rather higher values for QSOs. As a consequence, the filling factor is small, perhaps around 0.01.

1.4 The Unified Model for AGN

The observed luminosities and spectral features in quasars and, in general, in AGN supply fundamental evidences for the structure that is generating them. The single most significant observational datum is the extraordinary similarity of AGN spectra. This imply that they have fundamentally similar central engines, but ones whose power can vary by a factor of 10^7 without changing the kind of physical processes taking place.

The distinction of all quasars is also their ability to generate extraordinary luminosities within small volumes. The most shocking discovery came with the realization that these volumes were far too small to contain the stars needed to account for energy generation by nuclear processes.

In a pioneering paper appeared in *Nature* in early 1963, Hoyle and Fowler proposed that the energy for such sources was of gravitational origin, being derived from the collapse of very massive objects under their own strong gravitational fields. This early idea found a modified expression in the black hole accretion disc paradigm, a few years later (see Rees 1984 for an excellent account on this topic).

The most popular version of the unified model for AGN identifies four main zones. These are (Fig. 1.2):

1. The supermassive black hole at the very heart of the quasar surrounded by the accretion disc. Hurled away from the plane of the disc are jets of matter, moving at velocities close to the speed of light. This region is typically a few light-days across.
2. The broad line region at about 100 light days from the central source. I already described some of its properties in the previous section. There BLR is formed of two parts. The inner zone has a relatively high density, it is probably spherical and highly ionised. The outer zone is less dense, the clouds move more slowly, it is flattened and the gas is in a state of low-ionisation.
3. The molecular torus at about 100 light years across. This "doughnut" is made up of many clouds of dusty gas. A similar torus has been imaged directly by the Hubble Space Telescope in the Seyfert galaxy NGC 4261. The torus is optically thick, so if the torus is edge-on, the central regions (the accretion disc, the region

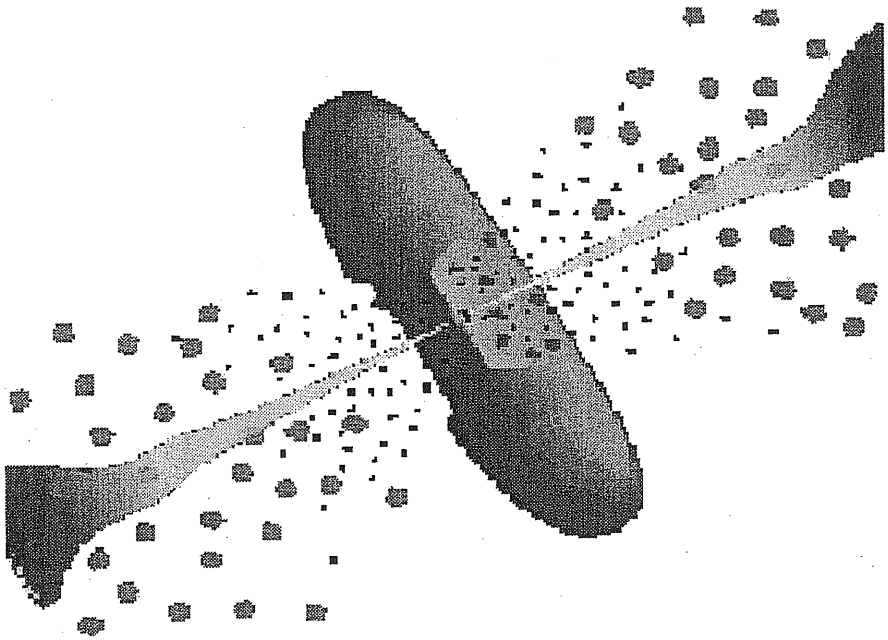

Figure 1.2: Cartoon of the unified model for AGN. Reproduced from Urry & Padovani (1995)

around the black hole and the broad line region) will be blocked out. This means that AGN will look different according to the angle at which it lies - this is the principle behind the Unified Model of AGN.

4. The narrow line region. These are the very outer regions of an active nucleus, before the surrounding galaxy is reached. As already said, the NLR is similar to the BLR, but it has a lower density and the clouds move with a lower velocity. Because of its low density, this gas can produce strong forbidden lines. Clouds that are in the cone of light from the inner nucleus are more ionised than those that lie in the shadow of the torus. These ionisation cones are seen in many nearby galaxies.

1.5 Absorption Lines

Once it is established that quasars are at cosmological distances, they can be used as useful probes of the intervening universe - like background searchlights - even

if the mechanism giving rise to their extreme luminosity is not totally understood.

The absorption line systems observed in QSO spectra can be classified in two main categories: the broad absorption line systems (BAL) and the narrow absorption line systems (NAL).

Broad absorption lines are observed in about 12 % of optically-selected quasars (Weymann et al. 1991). They are characterized by distinctive absorption troughs arising from various ions of differing excitation. There is general consensus that they are intrinsic to the quasar and caused by gas within the quasar itself with outflow velocity up to 60,000 km s^{-1}, relative to the emission redshift. They are often described as P Cygni profiles, by analogy to similar features seen in stars that are ejecting material toward the observer. The lines C IV, Al III, Al II, Si II, Si III, Fe II and Fe III show structures in the broad troughs.

Analysis of BAL systems have suggested that there is a link between the BAL gas and the broad emission line region, and detailed studies of BAL objects could provide diagnostic tools for the investigation of the nuclear region of the QSO (Wampler, Chugai, Petitjean 1995; Turnshek et al. 1996).

Another interesting fact is that BAL systems are only found in the radio-quiet population of quasars, whereas the narrow associated systems (with $z_{\rm abs} \approx z_{\rm em}$) of high metal abundance and high degree of ionisation are seen in both radio-quiet and radio-loud quasars.

More controversial are the discrete, narrow absorption line features (NAL). They can form in a variety of locations, ranging from very near the QSOs, as in ejecta like the BALs, to unrelated gas or galaxies at cosmological distances (Weymann et al. 1979). The so called "associated" or $z_{\rm abs} \approx z_{\rm em}$ absorbers (with redshift close to the emission redshift) have strengths and frequency of occurrence that appear to correlate with the QSO luminosities or radio properties, suggesting some physical relationship (Weymann et al. 1981; Foltz et al. 1988). These correlations may extend to NALs at blueshifts of 30,000 km s^{-1} or more (Richards et al. 1999).

On the other hand, the question of whether the majority of the NALs (with $z_{\rm abs} \ll z_{\rm em}$) are truly intervening at cosmological distances from the quasar, or ejected by it, which had received considerable interest in the earlier days, is now

settled in favor of the intervening hypothesis. The huge momentum requirements for ejection (Goldreich & Sargent 1976), the outcome of the Bahcall & Peebles test (1969; Young et al. 1982a) for a random redshift distribution of absorbers to different QSOs, the discovery of galaxies at the same redshifts as metal absorption systems (Bergeron 1986), and the detection of high metallicity gas in associated systems and low metallicities more than $30,000$ km s^{-1} away from the QSO (Petitjean, Rauch & Carswell 1994) leave no doubt that most of the systems are not physically related to the QSO against which they are observed.

The properties of these lines and the nature of the objects that give rise to them will be the topic of the following chapters.

Chapter 2

The Absorption Spectrum of Quasars

In this chapter, I summarize the fundamentals of line modelling and fitting. Then, I introduce the standard classification for narrow absorption lines. Eventually, I focus on the favourable distinctive features of QSO absorption spectra and on their relevance in the solution of specific cosmological questions.

2.1 Observational Approaches

The progress in the study of QSO absorption spectra has always been tightly linked to the evolution of telescope and detector technology.

The first quantitative results have been obtained with the University College London Image Photon Counting System (Boksenberg 1972) mounted on the RGO spectrograph at the 3.9m Anglo-Australian telescope, which allowed to obtain spectra with resolution $\lesssim 0.1$ nm.

Instrumental facilities whose introduction has marked significative steps in the overall knowledge of QSO absorption systems have been:

• The spectrographs (FOS, GHRS, STIS) mounted on the Hubble Space Telescope. The QSO Absorption Line Survey, a Key Project during the first four years of observations with the telescope, has resulted in the acquisition of ultraviolet spectra of a sample of 92 bright QSOs. Around 15 papers have been published

describing the properties of QSO absorbers at low and intermediate redshifts, and invetigating the relationship between absorbers and galaxies.

• The Echelle Multi Mode Instrument (EMMI, D'Odorico 1990) at the New Technology Telescope of ESO. The ESO key programme devoted to the study of QSO absorption systems at high redshift, produced a database of high resolution spectra which allowed to derive determinant results on the clustering properties of Lyα lines and on the proximity effect.

• The High Resolution Spectrograph (HIRES, Vogt et al. 1994) at the Keck telescope. In addition to interesting studies on the distribution of the number of lines with redshift and on the distribution of line parameters, striking results have been obtained on Deuterium abundances and on the metallicity of Lyα forest lines.

Great expectations are now placed on the Ultraviolet and Visual Echelle Spectrograph (UVES, D'Odorico 1997), the very high resolution spectrograph that in Fall 1999 will be mounted on one of the ESO Very Large Telescopes. UVES will reach a higher sensitivity in the UV and blue band and thus cover the gap left by HIRES in the Lyα forest at $z = 2 \div 3$. There will also be the possibility to use UVES in multiobject-mode and thus obtain simultaneously the spectroscopy of group of quasars covering regions of sky of one square degree.

Besides the evolution of technologies, the quality of final spectra depends strongly on the applied reduction algorithms. Eventually, the results on the properties of the absorbers are influenced by the way absorption lines are defined and fitted.

I will not dwell upon spectra reduction processes because I believe this topic is too technical for this context, although I want to underline the extreme influence it has on the resulting spectrum.

2.2 Absorption Line Definition

To gain relevant information from the lines of a spectrum, we have to model them with some analytical profile whose parameters are related to the physical proper-

ties of the gas from which lines originated.

The energy intensity of radiation, I_ν, will vary when passing through a cloud of gas due to absorption and emission, this process is described by the *equation of radiative transfer*,

$$dI_\nu = -\kappa_\nu \rho I_\nu dx + j_\nu \rho \, dx, \tag{2.1}$$

where κ_ν is the absorption coefficient per unit mass, and the emission coefficient j_ν is defined as the emission per unit mass, per unit frequency interval, into a unit solid angle. Equation (2.1) becomes,

$$\begin{aligned} \frac{dI_\nu}{\kappa_\nu \rho \, dx} &= -I_\nu + \frac{j_\nu \rho}{\kappa_\nu \rho} \\ \frac{dI_\nu}{d\tau_\nu} &= -I_\nu + S_\nu(\tau), \end{aligned} \tag{2.2}$$

where $d\tau_\nu = \kappa_\nu \rho \, dx$ defines the *optical depth* through the cloud and $S_\nu(\tau)$ is the source function.

Since we are concerned only with absorption lines, we can put $S_\nu(\tau) = 0$ and thus easily solve Eq. (2.2) as $I_\nu = I_0 \exp(-\tau_\nu)$.

The absorption coefficient, κ_ν, is related to the absorption cross-section per atom, σ_ν, by $\kappa_\nu \rho = \sigma_\nu N_{\text{abs}}$, where N_{abs} is the number of atoms per unit volume capable of absorbing radiation at the appropriate fequency. The cross-section in turn can be expressed as the product of the line profile and a constant for a particular line:

$$\sigma_\nu = \sigma_{\nu 0} \phi(\nu) \tag{2.3}$$

where the line profile is normalized so $\int \phi(\nu) d\nu = 1$ and hence $\int \sigma_\nu d\nu = \sigma_{\nu 0}$. The oscillator strength f is defined by

$$\sigma_{\nu 0} = \frac{\pi e^2}{4\pi \epsilon_0 mc} f \tag{2.4}$$

where m and e are the mass and charge of the electron.

The line profile is determined by the combined effects of different broadening.

• The Doppler broadening due to the motion of individual absorbing atoms. The frequency ν of a line absorbed by an atom moving with velocity v_r in the line of sight is $\nu = \nu_0(1 - v_r/c)$, where ν_0 is the rest frequency of the line. The sign of $\Delta\nu = \nu - \nu_0$ is negative (i.e., a redshift) if v_r is positive, i.e. away from the observer. The normalized profile of the line $\phi(\nu)$ will be given by the distribution of line of sight velocities v_r since the radiation that we observe is normally the result of the radiation of many atoms. If we assume that the absorbing atoms have a Gaussian velocity dispersion (a Maxwell-Boltzmann thermal distribution), we obtain for the thermal Doppler broadening,

$$\phi(\nu) = \frac{1}{\sqrt{\pi}\Delta\nu_D} e^{-(\Delta\nu/\Delta\nu_D)^2} \tag{2.5}$$

with the *Doppler width*

$$\Delta\nu_D = \frac{\nu_0}{c}\sqrt{\frac{2kT}{m}} = \frac{\nu_0}{c} b_{th} \tag{2.6}$$

where ν_0 is the frequency of the line centre, and $b_{th} = \sqrt{2kT/m}$ is the thermal Doppler parameter.

There can be also larger scale collective motions of group of atoms, called turbulence, that contribute to the Doppler line broadening. In case the turbulence can be modelled as a Gaussian velocity distribution with quadratic mean velocity $\sqrt{\frac{3}{2}}V$ and mean velocity $2V/\sqrt{\pi}$, then the resulting profile is still given by Eq. (2.5), but with Doppler parameter $b_{tot}^2 = b_{th}^2 + V^2$.

•The natural and pressure line broadening. Every spectral line has an intrinsic width due to the fact that excited states have in general very short lifetimes and by Heisenberg's uncertainty principle, $\Delta E \Delta t \sim h/2\pi$, so there will be a spread in energy level. The pressure or collisional broadening is instead the effect of de-excitation of the upper level caused by interactions with other particles. The two effects give a similar line shape, the so called Lorentzian profile

$$\phi(\nu) = \frac{\Gamma}{4\pi^2} \frac{1}{(\Delta\nu)^2 + \left(\frac{\Gamma}{4\pi}\right)^2} \tag{2.7}$$

where the Lorentzian width Γ is the sum of the collisional and natural line widths and microscopically is the de-excitation rate of the upper level, or total damping constant.

The overall line profile always contains both components. The line centre (core) is dominated by the Doppler profile, while the collisional/natural component is dominating the wings of the line. The convoluted profile is given by

$$\phi(\nu) = \frac{a}{\pi^{3/2}\Delta\nu_{\rm D}} \int_{-\infty}^{\infty} \frac{{\rm e}^{-x^2}}{(u-x)^2 + a^2} dx \tag{2.8}$$

where $u = \Delta\nu/\Delta\nu_{\rm D}$ is the frequency distance from the line centre in Doppler widths and $a = \Gamma/(4\pi\Delta\nu_{\rm D})$ is the ratio of the Lorentz width to the Doppler width. The convolved profile is called the *Voigt* profile. The integral has to be evaluated numerically, and a/π times the integral in (2.8) is called the Voigt function

$$\mathcal{H}(a,u) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{{\rm e}^{-x^2}}{(u-x)^2 + a^2} dx. \tag{2.9}$$

Now, going back to the optical depth definition and switching from frequency to wavelength, we obtain the expression

$$\begin{aligned} \tau(\lambda) &= \frac{\pi e^2}{4\pi\epsilon_0 mc} f N \frac{\lambda_0}{\sqrt{\pi} b} \mathcal{H}(a,u) \\ &= 1.498 \times 10^{-2} \frac{fN\lambda_0}{b} \mathcal{H}(a,u) \end{aligned} \tag{2.10}$$

where, λ_0 is the wavelength of the centre of the line and N is the column density, i.e., the number of atoms along the line of sight in a cilindric volume of unitary base. The optical depth at the centre of the line is:

$$\tau_0 = 1.498 \times 10^{-15} \frac{fN\,({\rm cm}^{-2})\lambda_0\,({\rm \AA})}{b\,({\rm km\ s}^{-1})}. \tag{2.11}$$

Absorption lines are observed against a continuous background. For each point in the profile, observations give the depth of the line,

$$r = (I_{\rm c} - I/I_{\rm c}) \tag{2.12}$$

where I is the observed spectral intensity and I_c the interpolation of the absorption-free continuum over the absorption feature. The overall line strength is given by the *equivalent width*, w_{obs}, defined as:

$$w_{obs} = \int \frac{I_c - I}{I_c} d\lambda = \int (1 - e^{-\tau(\lambda)}) d\lambda \tag{2.13}$$

If r_0 is the normalized depth at the line centre and $\Delta\lambda_{1/2}$ is the full width at half minimum (FWHM), for a rectangular profile $w_{obs} = r_0 \Delta\lambda_{1/2}$. Hence the equivalent width can also be defined as the width in wavelength of a rectangular-profiled line 100 % deep which has the same area in a flux-wavelength plot as the actual line.

For redshifted absorption lines, $w_{obs} = w_r \times (1 + z_{abs})$.

2.2.1 The Curve of Growth

It is apparent from Eq. (2.13) that the observed equivalent width does not depend on the spectral resolution. Thus, this is the parameter that can be measured for lines in spectra of low to intermediate resolution, when the instrumental width of lines is larger than the typical Doppler parameter, b, of QSO absorption lines.

From Eqs. (2.13), (2.10), (2.9) it is possible to relate the equivalent width to the column density for different values of the Doppler parameter. The function that gives this relation is called the *curve of growth* and an example corresponding to the H I Lyα transition is plotted on Fig. 2.1 for $b = 5, 10, 20$ and 30 km s^{-1} (Petitjean 1998). Three distinct regimes can be identified:

• When the column density is small ($\tau_0 < 0.1$), the absorption line is optically thin, the Voigt function reduces to a Gaussian profile and the equivalent width does not depend on b. This is the linear part of the curve of growth, where the determination of N from w is easy and reliable. For any transition,

$$N(\mathrm{cm}^{-2}) = 1.13 \times 10^{20} \frac{w_r(\text{Å})}{\lambda_0^2(\text{Å}) f} \tag{2.14}$$

•The logarithmic or flat-part of the curve of growth is characterized by the large dependence of N on b at a given w. In this regime, the determination of b and N

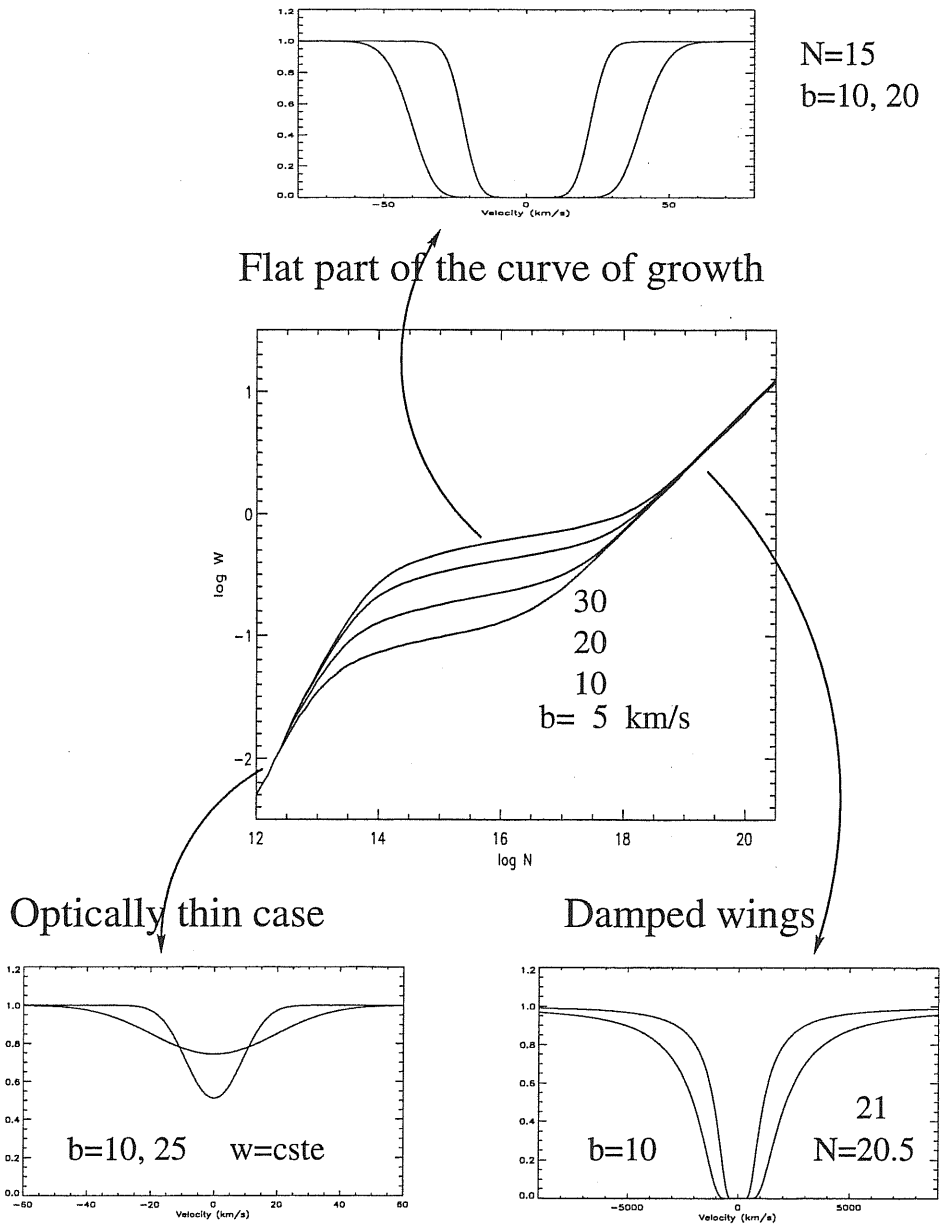

Figure 2.1: Curve of growth: logarithm of the equivalent width (w in Å) versus logarithm of the column density (N in cm^{-2}) for different values of the Doppler parameter (b in km s^{-1}). The curves are calculated for H I Lyα. The three characteristic regimes are illustrated (see the text). Reproduced from Petitjean 1998

are very uncertain except when several lines of the same ion are used. Equivalent width and optical depth at the centre of the line (see Eq. (2.11)) are related by:

$$\frac{w}{\lambda_0} = 2\frac{b}{c}\sqrt{\ln(\tau_0)} \tag{2.15}$$

• We have so far taken the line as having a Gaussian profile, since tha ratio $\Gamma/(4\pi\Delta\nu_D)$ is usually much less than one and the Gaussian element dominates except in the wings. However, as the line becomes stronger, the core will saturate while the wings can still grow, eventually they will supply the major part of the equivalent width. In this case the equivalent width can be found approximately by assuming that the line has a pure Lorentzian profile. This gives rise to the damping portion of the curve of growth in which the equivalent width no longer depends on b and the column density determination is again very accurate. In that case (see Eqs. (2.9) and (2.11)),

$$\frac{w}{\lambda_0} = 2.64\frac{b\sqrt{a}}{c}\sqrt{\tau_0} \tag{2.16}$$

For H I 121.6 nm, $\gamma = 6.265 \times 10^8$ s^{-1} (Morton & Smith 1973) and:

$$N(\mathrm{cm}^{-2}) = 1.88 \times 10^{18} w_r^2(\text{Å}) \tag{2.17}$$

2.2.2 High Resolution Spectroscopy: Voigt Profile Decomposition

The introduction of echelle spectrographs on 4m telescopes has allowed the observation of quasars at high resolution ($R \equiv \lambda/\Delta\lambda > 20\,000$). With this resolution narrow absorption lines can be observed with an instrumental FWHM lower than their intrinsic width. This makes it possible to independently derive the column density N and the Doppler parameter b from line profile fitting, rather than from the curve of growth analysis. The statistics of QSO absorption lines from high resolution studies largely have been cast in terms of the distribution functions of these quantities, together with the central redshift of the line, and their correlations. The main advantage of the high resolution approach is the opportunity of determining the shape of these distribution functions without parametric prejudices, by directly counting lines with parameters in a certain range.

The fit of heavy element systems may require the simultaneous fit of different lines spread all over the spectrum. For these lines, it is often necessary to test different configurations (i.e., number of components and constraints on them) before a satisfying fit is achieved. In the analysis of the Lyα forest, on the contrary, hundreds of independent lines are fitted. Even if, for strong Lyα lines the higher order Lyman lines can provide additional constraints when fitted simultaneously. The standard approach to this problem (Webb 1987; Carswell et al. 1991; Giallongo et al. 1993; Fontana & Ballester 1995) relies on χ^2 minimization to achieve a complete decomposition of the spectrum into as many independent Voigt profile components as necessary to make the χ^2 probability consistent with random fluctuations. The absorption lines are measured against a QSO continuum estimated locally from polynomial fits to spectral regions deemed free of absorption. A local high order continuum fit (as compared to a global extrapolation with a physical model for the QSO continuum) is necessary because the spectra are patched together from many individual echelle orders with strong variations in sensitivity. These variations do not divide out completely when dividing by the flux of a standard star because the light going through a slit narrow enough to ensure slit-width limited resolution varies with the seeing conditions and with the position of the object on the slit. When applying a local fit to the continuum the zeroth order contribution tends to be underestimated, i.e., the continuum is drawn too low, which is the main drawback of this method.

Given sufficient spectral resolution, and assuming that absorbers are discrete entities (which is debatable in the case of the Lyα forest, as will be discussed in section § 4.5) the profile fitting approach is the most physically meaningful way of extracting information from QSO absorption lines.

The Voigt profile model for spectral lines is based on the assumption that absorbing atoms have a Gaussian velocity dispersion, unfortunately, in most more realistic models of the absorbing gas finite velocity and density gradients invalidate this assumption, and the line parameters may have less immediate physical meaning. Departures of the absorption line shape from a Voigt profile may contain valuable information about the underlying nature of the absorption systems, and different scenarios may have quite different observational signatures. Rota-

tional motion (Weisheit 1978; Prochaska & Wolfe 1997), gravitational collapse (McGill 1990; Meiksin 1994; Rauch 1996) and galactic outflows (Fransson & Epstein 1982; Wang 1995) have been discussed in terms of the likely absorption line shapes they produce. As yet, the quantitative application of these results has proven difficult, because of the lack of realistic prototypical models for the actual line formation, the rather subtle departures from Voigt profiles expected, and the wide variety of profiles actually encountered.

Non-Voigt profiles can still be fitted as blends of several Voigt profiles, but the information about the non-thermal motion is encoded in the spatial correlations between the individual profiles (Rauch 1996). Also, there is no guarantee that the number of components necessary for a good fit converges with increasing signal-to-noise ratio.

2.3 The Classification of QSO Absorption Lines

In an earlier review, Weymann et al. (1981) introduced a classification of absorption systems that is still useful, although some of the distinctions introduced have been blurred by the most recent research (Tytler et al. 1995; Cowie et al. 1995). In particular, Weymann and collaborators distinguished two classes of absorption systems, physically separated from the QSO environment (i.e., with $z_{\rm abs} \ll z_{\rm em}$), according to whether they do, or do not, show metal absorption lines associated to the ubiquituous Lyα. For most of the Lyα clouds detectable with current technology ($N_{\rm HI} \gtrsim 10^{12}$ cm^{-2}) metal lines with metallicities common at high redshifts ($Z \lesssim 10^{-2}\ Z\odot$) are simply below the detection threshold. Therefore this classification is an observational one and is subject to variation linked to technology improvements.

A more refined classification that embeds the previous one, is based on the H I column density of Lyα lines. It goes as follow. At column densities $N_{\rm HI} \sim 10^{12}$ cm^{-2} we start to detect absorption lines; at $N_{\rm HI} \gtrsim 3 \times 10^{14}$ cm^{-2}, C IV doublets are found associated to roughly half of the Lyα forest lines (Cowie et al. 1995). At column densities $N_{\rm HI}$ exceeding 10^{17} cm^{-2} the gas becomes optically thick to the ionising radiation and a discontinuity at the Lyman limit (91.2 nm) is detectable. These systems are called in fact Lyman limit systems and they coincide with the

systems detected by the Mg II 279.8, 280.2 nm doublet. In systems with $N_{\rm HI}$ larger than $\sim 10^{19}$ cm^{-2}, selfshielding renders the gas predominantly neutral. The damping wings of the Lorentzian component of the absorption profile begin to be detected, and the absorption features are called damped Lyα systems. The definition of damped systems as absorbers with $N_{\rm HI} \gtrsim 2 \times 10^{20}$ cm^{-2} is artificial and was introduced to search for damped candidates in low spectral resolution data (see Wolfe et al. 1986). Indeed, for such column densities, the absorption line has $w_{\rm obs} > 1.5$ nm at $z > 2$.

The Lyman limit systems and the damped Lyα systems are two extremely interesting classes of absorbers. They are tightly linked to galaxies, in particular, various samples of low redshift ($z_{\rm abs} \sim 1$) Mg II galaxies have been directly observed and studied (Bergeron & Boissé 1991; Steidel, Dickinson & Persson 1994; Churchill, Steidel & Vogt 1996; Guillemin & Bergeron 1997).

Imaging searches for the galaxies responsible of producing damped Lyα systems have revealed a highly diverse population of absorbers, low surface brightness galaxies, dwarfs, as well as spirals (Steidel et al. 1994; Le Brun et al. 1997; Lanzetta et al. 1997; Rao & Turnshek 1998). The damped systems are valuable probes of the high redshift universe since they are the reservoir of neutral hydrogen. They have been studied extensively to determine the consumption of interstellar gas by stars (Lanzetta, Wolfe & Turnshek 1995; Storrie-Lombardi, McMahon & Irwin 1996), to measure abundances of metals and dust (Fall& Pei 1993; Lu et al. 1996; Pettini et al. 1997; Boissé et al. 1998; Vladilo 1998; Pettini et al. 1999; Prochaska & Wolfe 1999), and to explore the kinematics of forming galaxies (Prochaska & Wolfe 1998; Haehnelt, Steinmetz, Rauch 1998).

Unfortunately, I have not had the opportunity to work on this subject up to now, for this reason I will not dwell further upon it in the following chapters and I refer the interested reader to the previously listed references.

2.4 The Cosmological Relevance of QSO Absorption Spectra

In the previous chapter, I have given first an overview on the spectrum as a whole of quasars and, more in general, of AGN to get acquainted with these powerful cosmic lighthouses. Then, I have focussed on the absorption features that populate the spectra of quasars and briefly described the theoretical tools which allow us to relate the observed modulation of the QSO transmitted flux to the physical properties of the absorbing material. At this point, it should start to become selfevident why it is more than twenty years that QSO absorption spectra cover a primary role in extragalactic astronomy and are the target of at least one Key Project of any ground-based or spatial new telescope.

Actually, the position of QSO absorption spectra has changed in the last years. The detection of the population of Lyman Break Galaxies (LBG) at $z \sim 3$ (Steidel et al. 1996a, 1996b) and at $z \sim 4$ (Steidel et al. 1999) has allowed to study accurately and directly luminosity functions, color distributions, star-formation rate and clustering properties (the latter only for the sample at $z \sim 3$) of star-forming galaxies at high redshifts. Thus, absorption systems in QSO spectra are no longer the only tool we have to investigate the distribution of matter at early epochs. Indeed, they are still a unique source of information, infact for their intrinsic nature they probe regions of the universe rich in gas and thus complementary to the LBG star-forming environment.

In the following, I want to make clear what are the favourable characteristic offered by absorption features as probes of the physical state of the universe during the early stages of cosmic evolution.

- One of the most relevant features is the extremely high sensitivity to the presence of neutral hydrogen along the line of sight due to the large cross section of the Lyα resonance transition. The present instrumental capabilities allow to detect lines with $N_{\rm HI} \gtrsim 10^{12}$ cm^{-2}. To make a comparison, the neutral hydrogen column density necessary to observe the 21 cm radiation is about 8 orders of magnitude larger.
- The range of sampled parameters is wide: H I column density can vary between 10^{12} cm^{-2} and 10^{22} cm^{-2}, metal abundances have been found from above solar,

in systems associated to the QSO, to 10^{-3} solar, temperatures from 10^2 to 10^6 K. The redshift interval over which absorption systems have been observed goes from $z \sim 0$, thanks to HST observations, to the highest redshift QSO at $z \sim 5$. In Big Bang cosmologies, these redshifts correspond to times when the Universe itself was just ~ 1 Gyr old.

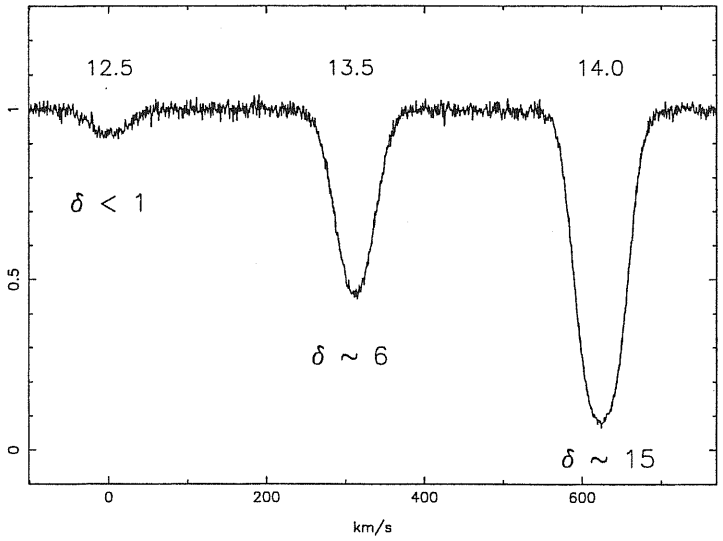

Figure 2.2: The observable range of the linear part of the curve of growth. The logarithmic column densities and corresponding typical baryonic overdensities δ are shown for three artificial $z = 3$ H I Lyα lines. Reproduced from Rauch 1998a.

• The lines are associated with typical structures as opposed to the "rare" structures (such as galaxies and clusters) that arise from the statistical tail of density perturbations. In particular, the linear density regime of gravitational collapse near turnaround for galactic scales (several hundred kpc at $z \sim 3$) coincides roughly with the linear part of the curve of growth for the Lyα absorption line (see Fig. 2.2, Rauch 1998a). This property has been exploited in the inversion methods described in section § 4.5.

• The observational selection biases can be understood and quantified. Absorption lines are due to gas (a) large in extent, (b) of low ionisation to place the main lines at long wavelengths, (c) with high density and low radiation field to favour low ionisation, (d) in front of lensed QSOs and (e) dust extinction free.

• They are observed along random directions in the sky and may not be biased with respect to the local structure distribution.

• They potentially comprise an enormous database, with approximately 10^3 discovered quasars each generating tens to hundreds of lines, presenting a distinct statistical advantage compared with other astronomical observations.

Now, the first topic addressed with QSO absorption spectra is the investigation of the physical properties of the population of objects from which absorption lines themselves get originated. The nature of the absorbers, besides giving us an insight over the fauna populating the high redshift universe, allows us to effectively compare observations with model predictions.

Furthermore, the following open issues in cosmology can be tackled.

• The cosmological baryon density. From measures of Deuterium abundances and from the Lyα forest.

• The distribution of gas in the intergalactic medium and in galactic and/or protogalactic halos, hot and cold.

• The formation of structure. From the kinetic properties of absorptions measured in line profiles.

• The formation of heavy elements and the early enrichment of the intergalactic medium.

• The UV radiation field. Origin, spectrum and intensity from the proximity effect and the abundance ratios of elements like C IV and Si IV.

• The temperature of the cosmic microwave background. From the presence and strength of the excitation produced by the CMB in certain atomic transitions (e.g., carbon fine-structure).

• The opacity of the Universe. From the Gunn-Peterson trough and the number density of lines due to H I and He II.

Chapter 3

Observational Results

This chapter try to picture the historical evolution and the state of the art of the study of absorption systems in QSO spectra. Given the vastness of the subject, I chose to focus mainly on the aspects I am more interested in and/or I directly faced in the course of my research work.

I mainly describe the results of the statistical analysis of the observational parameters of the lines distinguishing between metal systems and Lyα forest lines. At the end of this review I inserted the study of the high resolution spectrum of the QSO PKS 2126–158, based on the paper by D'Odorico et al. (1998a).

The spatial distribution of absorbers and their size properties, that constitute my main original contribution to the subject are described in Chapters 4 and 5.

3.1 An Historical Overview

The large cross-section for the Lyα UV resonance line in neutral hydrogen (121.6 nm) suggested that the gas homogeneously distributed in an expanding universe would produce an absorption trough in the spectrum of a QSO due to the continuously redshifted Lyα line (independent predictions by Gunn & Peterson 1965; Scheuer 1965; Shklovski 1965). Gunn and Peterson found such a spectral region of reduced flux blueward of the QSO's Lyα emission line, and used this measurement to put upper limits on the amount of intergalactic neutral hydrogen ($\Omega_{\rm IGM} \lesssim 0.01$ in the hypothesis that quasars are the major contributors to the

ionising UV flux, Giallongo et al. 1994).

Bahcall & Salpeter (1965) stated that a population of discrete absorption lines should also be observed from a more clumpy gas distribution, specifically from intervening groups of galaxies. Discrete lines were detected shortly thereafter (Lynds & Stockton 1966; Burbidge et al. 1966; Stockton & Lynds 1966; Kinman 1966). The simultaneous detection of higher order lines of the Lyman series (e.g. Baldwin et al. 1974) had confirmed the suggestion (Lynds 1971) that most of the absorption is indeed from H I Lyα. At higher spectral resolution, the Lyα forest can be resolved into hundreds (in $z > 2$ QSO spectra) of distinct absorption lines, the appearance of which gave rise to the label *Lyα forest* (Weymann et al. 1981).

The study of metal line systems have proceeded more straightforwardly. As soon as spectra could be obtained with a FWHM < 0.25 nm, the common doublets of C IV and Mg II could be resolved and thus easily and reliably detected. Their association with galaxies has never been questioned, even if the kinetic structure of their absorption profiles has not yet found a satisfactory explanation (Churchill, Steidel & Vogt 1996; Haehnelt et al. 1996; Rauch et al. 1996; see also the references in Chapter 4). Unfortunately, heavy-element redshifts are not as common as Lyα, thus their statistical properties are still poorly known. Furthermore, dedicated campaign of observations have not been carried out recently, and the largest homogeneous sample of high redshift C IV still remains the compilation of 55 QSO spectra by Sargent, Boksenberg and Steidel (1988).

The basic observational properties of the QSO absorption systems were established in the late 1970s and early 1980s when the combination of 4m telescopes (e.g., the AAT, KPNO, MMT, Palomar) and sensitive photon counting electronic detectors (e.g. the University College London's IPCS) first permitted quantitative spectroscopy on high redshift QSOs to be performed. Making use of the new technology, the work by Sargent et al. (1980) set the stage for what for many years has been the standard picture: Lyα absorbers were found to be consistent with a new class of astronomical objects, intergalactic gas clouds, because of their large rate of incidence ($d\mathcal{N}/dz$), their pristine composition and their weak clustering, distinct from metal absorption systems which were related to galaxies.

3.2 The Redshift Evolution of the Line Population

In an individual QSO line of sight, observations of the Lyα forest or of the region between the Lyα and the CIV emissions at high redshift ($z \sim 3$) can extend over a redshift range Δz greater than unity. Then we are sampling a significant fraction of a Hubble time, and it is natural to look for changes in the rate of incidence of absorption lines with redshift.

An analytic expression (Wagoner 1967; Bahcall & Peebles 1969) can be given for the number of absorption systems per unit redshift, $d\mathcal{N}/dz$, in terms of the comoving number density $n_0(z)$ of absorbers, the geometric absorption cross section $\sigma(z)$, and the Hubble constant H_0:

$$\frac{d\mathcal{N}}{dz} = \frac{c\, n_0(z)\sigma(z)}{H_0} \frac{1+z}{(1+q_0\, z)^{1/2}}. \tag{3.1}$$

For absorbers with no intrinsic evolution,

$$\frac{d\mathcal{N}}{dz} \propto \begin{cases} 1+z, & q_0 = 0 \\ (1+z)^{1/2}, & q_0 = \frac{1}{2} \end{cases}. \tag{3.2}$$

Peterson (1978) first found evidence of an increase in the number of Lyα clouds with redshift beyond what was expected for a population of objects with invariant cross-section and uniform comoving number density. Sargent et al. (1980) drew attention to the need of defining a consistent sample by setting a common lower limit to the intrinsic equivalent width $w_{\rm r}$ of the Lyα lines, and they found lack of evolution for their sample. The presence of evolution was at first subject to some debate (see the summary by Murdoch et al. 1986), but it is now clear that the Lyα forest as a whole evolves quite strongly with z. The observationally determined evolution in the number of absorbers above a certain column density threshold is usually expressed in the form

$$\frac{d\mathcal{N}}{dz} = \left(\frac{d\mathcal{N}}{dz}\right)_0 (1+z)^{\gamma} \tag{3.3}$$

(Sargent et al. 1980; Young et al. 1982b), where the exponent includes the cosmological dependence. Much observational effort has been devoted to studying the redshift number density evolution, but unfortunately the resulting conclusions

are far from robust. This is because Lyα cloud column densities N are distributed according to a power law $N^{-\beta}$ with index $\beta \sim 1.5$ (see below) so the majority of lines in any column density limited sample are always close to the threshold, and small variations in imposing the threshold can cause large changes in the estimated numbers of lines, and in γ. Moreover, line blending, and its dependence on spectral resolution, data quality, and redshift can make a large difference in the normalization $(d\mathcal{N}/dz)_0$, with individual studies differing by a factor of two or more.

At lower resolution (FWHM $\sim 50 - 100$ km s^{-1}) large samples of lines have been used to investigate this topic. The γ values tend to lie between the low value $\gamma = 1.89 \pm 0.28$ obtained by Bechtold (1994), (for $w_r > 0.032$ nm), and the high $\gamma = 2.75 \pm 0.29$ (for $w_r > 0.036$ nm) from the study by Lu, Wolfe and Turnshek (1991). High resolution spectra, usually confined to $z > 2$, give in general high values: $\gamma = 2.65 \pm 0.21$ ($\log N_{\rm HI} > 13.3$; $1.7 < z < 4.1$; Giallongo et al. 1996); 2.9 ± 0.3 ($2 < z < 4.5$; Cooke et al. 1997); $\gamma = 2.78 \pm 0.71$ ($\log N_{\rm HI} > 13.77$; $2 < z < 3.5$; Kim et al. 1997). It was also shown that in this same redshift range, weak lines ($\log N_{\rm HI} < 14$) evolve more slowly than strong ones, with $\gamma = 1.29 \pm 0.45$ consistent with a non-evolving model (Kim et al. 1997). At $z > 4$ the evolution appears to be accelerating, with γ increasing from a value just below 3 to 5.5 (Williger et al. 1994).

The Lyα forest observed at low redshift ($z < 1.5$) by HST (Bahcall et al. 1993, 1996; Impey et al. 1996; Jannuzi et al. 1998) shows more absorption systems than expected from a naive extrapolation of any of the high z power law exponents. The last results of the HST Key Project on this topic (Weymann et al. 1998), confirm the low $\gamma = 0.16 \pm 0.16$ valid from $z = 0$ to $z \sim 1.5$ which gives a $d\mathcal{N}/dz$ consistent with weak negative evolution. Also in this redshift regime an increase in γ with increasing column density threshold has been found. Single power law fits attempting to explain both high and low z absorbers fail to account for the rapid upturn around $z \sim 1.5 - 1.7$ (e.g., Impey et al. 1996).

It has to be noted that the comparison of high and low redshift data presents a caveat due to the different techniques used to measure column densities. At high redshifts, high resolution observations allow line fitting of Voigt profiles, and those give more robust results than the curve-of-growth technique used to estimate

H I from equivalent width measurements of low resolution data at low redshift.

Data at higher resolution from STIS and UCLES observations in the redshift range $1.5 < z < 2$ have been obtained for the QSO J2233-606 in the Hubble Deep Field South, just falling in the gap between the two redshift regimes described before. Figure 3.1, reproduced from Savaglio et al. (1999), shows the number density evolution for lines with $\log N_{\rm HI} > 14$. The Lyα forest of J2233-606 shows on average a higher number of strong lines than expected from studies at lower and higher redshifts. A larger sample of lines is needed to confirm if this line density is typical at that redshift or if it is peculiar to this line of sight.

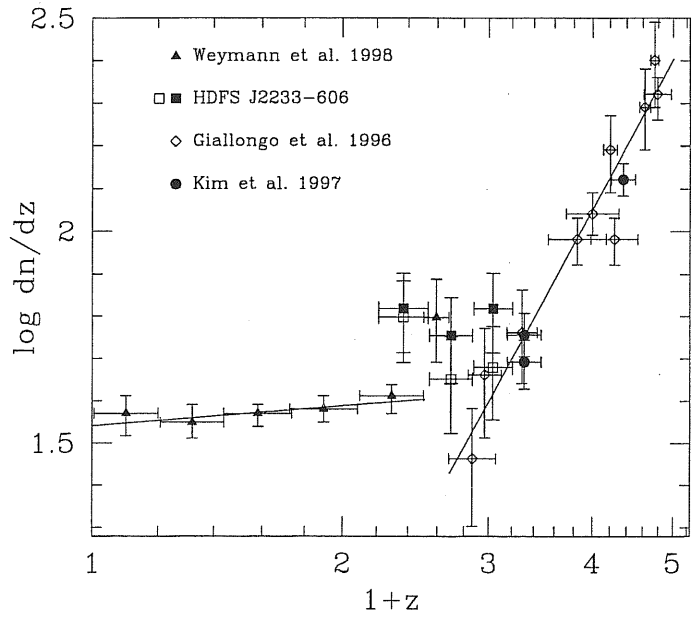

Figure 3.1: Number density evolution of the Lyα clouds with $\log N_{\rm HI} \geq 14.0$ from $z = 0$ to $z = 4$. Filled symbols are for samples that include metal systems, open symbols do not. In the sample of Kim et al. (1997), only lines with $\log N_{\rm HI} \leq 16.0$ are included. The low redshift line is the fit with $\gamma = 0.16$ obtained by a sample of Lyα clouds with equivalent width $w_{\rm r} \geq 0.024$ nm (Weymann et al. 1998) which corresponds to $\log N_{\rm HI} \geq 14$ for $b \sim 26$ km s^{-1}. The high redshift line is the fit with $\gamma = 3.62$. Reproduced from Savaglio et al. (1999)

A possible explanation of the observed redshift evolution of Lyα lines has

been proposed by Riediger, Petitjean and Mücket (1998) who carried out N-body simulations down to redshift $z = 0.0$. They distinguish two population of absorbers: P_u in which material has not been shocked and it is found primarily in underdense regions and in the more outlying regions of condensations, and P_s, the shocked population which is found in the vicinity of regions undergoing condensations, in big halos and elongated filamentary structures. The simulations show that population P_s evolves slowly with redshift and, for $\log N_{HI} > 14$, dominates the number density of lines at $z < 3$; the unshocked population for the same column densities, disappears rapidly at low redshifts. The total number of lines with $\log N_{HI} \gtrsim 12$ remains approximately constant from $z \sim 4$ to $z = 1$ and decrease slowly at lower redshift.

Finally, an interesting similarity is revealed between the volume density distributions of Lyman limit systems, damped systems and Lyα systems with $\log N_{HI} \gtrsim 14$ as derived on the basis of their estimated sizes and geometries. Also their cosmological evolutions appear consistent in the overall redshift range (Giallongo et al. 1996). This represents a further evidence of a physical association between the strong Lyα absorptions and the halos of protogalactic systems.

The evolution in redshift of the number density of C IV lines has been investigated by Bergeron and Boissé (1984) on the basis of the data by Young, Sargent and Boksenberg (1982a), with a small extension. The unbiased sample consisted of only 38 absorption systems along 39 lines of sight over the redshift range $1.2 \leq z \leq 2.2$. Bergeron and Boissé found that the best fit to the power law (3.3) has $\gamma = 1.8 \pm 1.5$; the result is therefore consistent with no-evolution for $q_0 = 0$, and shows very marginal evidence for evolution with $q_0 = 1/2$. Sargent, Boksenberg and Steidel (1988) verified with a K-S test that data are consistent with the hypothesis that the absorption redshifts are produced by cosmologically distributed intervening objects, i.e. their distribution is described by Eq. (3.1). They obtained for a sample of 86 CIV absorptions in the redshift interval $1.2 < z < 3.7$ a maximum likelihood estimation of $\gamma = -1.2 \pm 0.7$ ($w_r > 0.015$ nm). Including the data from Young, Sargent and Boksenberg (1982a), and Foltz et al. (1986), they find $\gamma = -1.56 \pm 0.79$ ($w_r > 0.03$ nm). At lower redshifts, $0.15 < z < 2.2$, the number density of Mg II systems increases with a slope $\gamma = 0.78 \pm 0.42$ for

$w_r > 0.03$ nm (Steidel & Sargent 1992). The level of evolution for both C IV and Mg II systems is strongly dependent on the w_r threshold, the stronger systems evolving more rapidly. By combining the behaviour of Mg II systems at low redshifts and of CIV systems at high redshifts can be inferred that the density of heavy element redshifts first rises up to $z \sim 1$, then flattens off, and finally declines beyond $z \sim 2.5$. This effect if interpreted as being due to the combined effects of cosmological expansion and of a gradual formation of the heavy elements over cosmic time (Sargent, Steidel & Boksenberg 1988, Bergeron 1995) it is in conflict with the results obtained both from Damped Lyα systems (see e.g., Pettini et al. 1999) and galaxies (see e.g., Steidel et al. 1999) which show a constant metallicity up to redshift $z \sim 4$.

3.3 The Proximity Effect and the Measure of the Ionising UV Background

The UV radiation from QSOs has been considered as the most natural origin for the ionisation of the intergalactic gas (Arons & McCray 1969; Rees & Setti 1970).

Carswell et al. (1982) studying the redshift distribution of absorption lines, realized that the general increase with redshift was accompanied by a simultaneous decrease of $d\mathcal{N}/dz$ in each individual QSO spectrum when approaching the QSO's emission redshift. Murdoch et al. (1986) calling it the inverse effect, confirmed this result and proposed it as a possible explanation of the inconsistencies in the different determinations of γ. They also considered as a possible mechanism the fact that the Lyman continuum flux from each QSO may substantially enhance the degree of ionisation of the hydrogen clouds nearest in redshift, leading to correspondingly lower column densities of neutral hydrogen. Tytler (1987b) referring to the effect as the "anomaly", questioned the specific assumption that the reduced absorption is caused by the QSO's radiation field. Hunstead et al. (1988), and Webb and Larsen (1988) defended the earlier interpretation, and in particular the claim that the effect is local to the neighbourhood of the QSO.

If the inverse effect is indeed caused by enhanced ionisation, measuring the intensity of the ionising UV background becomes possible from observations of

the density of lines, $d\mathcal{N}/dz$, as a function of the distance from the QSO. Given that the clouds are highly ionised by the general background of UV photons, the neutral hydrogen density in a cloud is inversely proportional to the ionising flux impingent on it. In particular, for a given cloud near a bright quasar,

$$N_{\rm HI} = N_\infty (1+\omega)^{-1}, \tag{3.4}$$

where N_∞ is what the column density would be at an infinite distance from the quasar, and

$$\omega = \frac{F_\nu^Q}{4\pi J_\nu}. \tag{3.5}$$

In this equation, $J_\nu(z)$ is the Lyman limit background radiation intensity (erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}) at the cloud redshift z, and $F_\nu^Q = L_\nu^Q/4\pi r_L^2$, is the local Lyman limit flux density (erg cm^{-2} s^{-1} Hz^{-1}) due to the QSO. In the hypothesis of a power law form for the column density distribution (see below), it is possible to relate the decrease in column density to the corresponding decrease in $d\mathcal{N}/dz$ for lines above a given detection threshold. In this way Carswell et al. (1987) performed the first crude measurement of the UV background radiation field, obtaining $J_{-21} = 3$, where $J_\nu = J_{-21} \times 10^{-21}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}. Bajtlik, Duncan and Ostriker (1988) assembled a line list from 19 QSO spectra and presented evidence that the inverse effect was correlated with the intrinsic luminosity of the QSOs. They called this the "proximity effect" and developed a quantitative physical model. Their measurement procedure (adopted by most later studies) consists of fitting the number density of lines per *coevolving redshift interval*, $dX = (1+z)^\gamma dz$

$$\frac{d\mathcal{N}}{dX} = \left(\frac{d\mathcal{N}}{dz}\right)_0 (1+\omega)^{1-\beta}. \tag{3.6}$$

The quantity β is the exponent of the power law distribution of column densities. They obtained $J_{-21} = 1^{+3.2}_{-0.7}$. Lu et al. (1991) with a larger sample get identical results. Bechtold (1994) in her even larger dataset finds $J_{-21} = 3$.

The largest compilations of high resolution data was assembled by Giallongo et al. (1996). They applied a maximum likelyhood analysis to estimate simultaneously the best-fit parameters of the Lyα statistic and of the UV background.

Considering Eqs. (3.3), (3.4) and a double power law distribution for column densities, they wrote,

$$\frac{\partial^2 \mathcal{N}}{\partial z\, \partial N_{\rm HI}} = \mathcal{A}_0 (1+z)^{\gamma} \quad (1+\omega)^{1-\beta_f} \times \begin{cases} N_{\rm HI}^{-\beta_f} & N_{\rm HI} < N_{\rm break} \\ N_{\rm HI}^{-\beta_s}\, N_{\rm break}^{\beta_s-\beta_f} & N_{\rm HI} \geq N_{\rm break} \end{cases} \quad (3.7)$$

where $N_{\rm break}(z) = N_{\infty,b}/(1+\omega(z))$ is the observed break in the column density distribution (see below), which is shifted to lower and lower $N_{\rm HI}$ as the QSO emission redshift is approached. After correcting for the blanketing of weak lines, they find a value, $J_{-21} = 0.5 \pm 0.1$, a factor of 6 lower than previous estimates, marginally consistent with $J_{-21}=1^{+0.5}_{-0.3}$ obtained from another high resolution sample (Cooke et al. 1997).

None of these studies has found evidence for a significant change with redshift (for $1.6 < z < 4.1$). However, Williger et al. (1994) and Lu et al. (1996) both found lower values ($J_{-21} = 0.2$) just above $z \sim 4$, in contrast to Savaglio et al.'s (1997) value ($J_{-21} = 0.7$) for the same redshift which is consistent with no evolution. In any case, when going to lower redshifts there appears to be a drastic decline in intensity below $z \sim 1.6$. At $z \sim 0.5$ Kulkarni and Fall (1993) obtained $J_{-21} = 6 \times 10^{-3}$ from HST FOS data (Bahcall et al. 1993).

Assuming that the proximity effect really allows to measure the background ionising intensity, we still know this quantity only to within an order of magnitude, and the uncertainty may even be larger than that. This is because, in addition to the errors from line counting discussed in the previous section, there are other systematic uncertainties sometimes difficult to quantify (some of them have been reviewed by Bechtold 1994, 1995).

A major but measurable effect arises from the systematic difference between the QSO emission redshift measured by high ionisation and low ionisation lines. The blueshift of the peak of the Lyα, C IV λ154.9 nm and C III] λ190.9 nm with respect to the Balmer lines and Mg II λ279.8 nm may typically amount to $\sim$ 1500 km s^{-1}. Downward corrections for J_ν correspond to a factor of 2–3 in this case (Bechtold 1994; Espey 1993). These corrections have been taken into account in the work by Williger et al. (1994), Lu et al. (1996), Cooke et al. (1997) and

Savaglio et al. (1997) and they would reduce for example the Bechtold value by a factor of three.

The effect of the UV radiation field of a foreground QSO on the absorption spectrum of a background QSO provides an important way of testing the "proximity effect" interpretation of the inverse effect, no significant results have been obtained up to now.

The proximity effect measurements beg the question whether there are enough QSOs to produce the ionising background seen, or whether an additional population of sources is needed. The values derived for the Lyman limit UV background flux are not far from that predicted for the quasar population at the same redshift $J_{-21} \sim 0.1 - 0.2$ (Haardt & Madau 1996), although there is room for a contribution by other kind of ionising sources like primeval galaxies.

Disentangling between these two possibilities requires the knowledge of the shape of the UV background at the He II edge at 4 Rydberg (54.4 eV or 22.8 nm). Infact, absorption by Lyα clouds produces a break in the ionising background at the He II edge whose amplitude and redshift variation dipends on the source of ionisation. It is larger and its amplitude increases with redshift if galaxies are the dominant contributors to the emissivity; in contrast, if quasars dominate, the break is smaller and approximately constant in redshift (Miralda-Escudé & Ostriker 1990). The possible diagnostics are either the direct measure of the quasar flux at 4 Ryd, e.g. by measuring the He II Gunn-Peterson absorption in the few cases where the QSO spectrum can be observed in this region (Jakobsen et al. 1994; Tytler et al. 1995; Davidsen, Kriss & Zheng 1996; Hogan, Anderson & Rugers 1997; Reimers et al. 1997; Zheng, Davidsen & Kriss 1998) or in an indirect way, the measure of the relative abundances of ions like C IV and Si IV whose ionisation potentials are near the He II edge.

Songaila and Cowie (1996) and Songaila (1998) considered metal lines associated to the low column density $N_{\rm HI} > 5 \times 10^{14}$ cm^{-2} Lyα lines. The Si/C ratios in the forest clouds are similar to those in low metallicity stars in our Galaxy, i.e., about a factor of 3 higher than solar. They found an abrupt transition for the value of SiIV/CIV at a redshift slightly lower than 3, which they interpret as being caused by a change in the spectrum of the metagalactic ionising flux, passing from a pure

power-law with an index of -1.8 at $z \lesssim 3$ to a very soft ionising flux, with a large break at the He II edge ($B \equiv J(4\ Ryd)/J(4.5\ Ryd) \geq 20$) at $z \gtrsim 3$. This could be a result of the transition from galaxy-dominated to quasar-dominated ionising flux, but, given the rapidity of the transition, it appears to be more likely that we are seeing the redshift at which He II ionises completely to He III. This result has been confirmed by the observation of highly patchy He II Lyα absorption at this redshift (Reimers et al. 1997). Savaglio et al. (1997) investigated three optically thin metal systems with $N_{\rm HI} \sim 10^{15}$ cm^{-2} at redshifts $z \geq 3.5$ and found that [Si/C] ratios lower than 10 times the solar value can be obtained only assuming a large jump in the spectrum of the ionising UV background beyond the He II edge ($J(1\ Ryd)/J(4\ Ryd) \gtrsim 1000$). This result is suggestive of a possible increase of the stellar ionising emissivity over the declining quasar one for $z > 3$.

We tackled this issue and obtained similar results for a metal system in the spectrum of PKS 2126-158, see § 3.7.3.

3.4 The Column Density Distribution

In absorption line studies, the column density distribution function (CDDF), i.e., the number of absorbers in a given column density bin occupies a similarly, central place (and provides similarly vague information) as the luminosity function in the study of galaxies. The observational determination of the CDDF relies on a patchwork of techniques, owing to the large dynamic range of observable H I absorption. We already saw (see § 2.2.1) that there are two regimes, the linear and the damping portion of the curve of growth, in which the determination of the column density is quite unambiguous. In the central part where the line starts to saturate but the damping wings are not yet visible, the column densities are relatively difficult to measure exactly by any means. The situation is complicated further by the appearance of noticeable multi-component structures, blending among these components can mimick large Doppler parameter / high column density lines. In the case of metal lines the effect is less severe thanks to the smaller thermal width of the lines, and, in general, it is possible to fit simultaneously different ions at the same redshifts, by making the hypothesis that they present the same velocity structure. For Lyα lines simultaneous fits to the higher order Lyman lines (which

are less saturated owing to their lower oscillator strengths) can only help to some degree.

Carswell et al. (1984) found that the number of absorbers $\mathcal{N}$ per unit H I column density interval can be parametrized as

$$\frac{d\mathcal{N}}{dN} \propto N^{-\beta}, \quad \beta = 1.7 \pm 0.1 \quad 13 < \log N_{\rm HI} < 16. \tag{3.8}$$

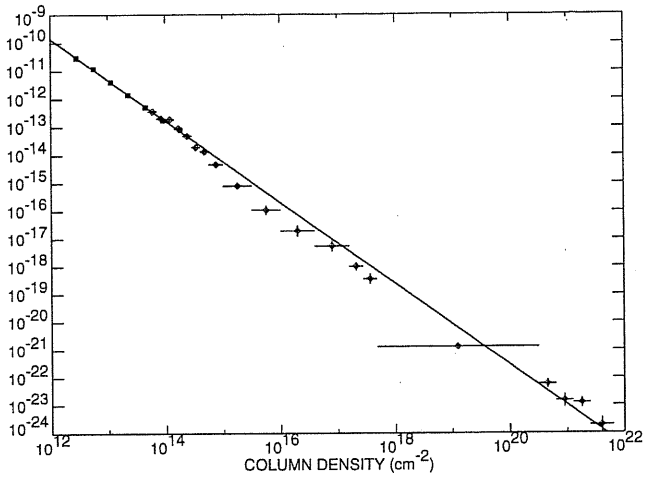

Figure 3.2: HI column density distribution at $z \sim 2.8$ from Hu et al. (1995) and Petitjean et al. (1993). The ordinate is given by the differential column density distribution $f(N) = m/(\Delta N \Delta X)$, where m is the number of clouds with column density between N and $N + \Delta N$ and ΔX is the sum of the observed absorption lengths towards each QSO. $f(N)$ is roughly a factor of 3.8 less than $(d\mathcal{N}/dN\, dz)$. The deficit of lines at $N \sim 10^{15}$ cm^{-2} is clearly visible. Reproduced from Hu et al. (1995)

Tytler (1987a) found that a single power law with slope $\beta = 1.5$ could fit well the whole range of observable column densities between $10^{13} < N_{\rm HI} < 10^{22}$ cm^{-2}. He suggested that this finding could be understood if Lyα and metal-line systems belong to the same population. Petitjean et al. (1993) considered a new compilation of high column densities H I lines, and they excluded a single power law fit at the 99 % level. There is a clear flattening at $N_{\rm HI} \simeq 10^{15} - 10^{16}$ cm^{-2} that they ascribed to the transition between Lyα and metal systems. Keck spectra (Hu et al. 1995; Lu et al. 1996; Kirkman & Tytler 1997) appear to show that the power law extends over ten orders of magnitude in column density from 10^{12}

to 10^{22} cm^{-2} (if the confusion correction made at the low column density end is justified, see Fig. 3.2). One of the more recent measurements (Hu et al.1995) gives

$$\frac{d\mathcal{N}}{dN} = 4.9 \times 10^7 N^{-1.46}, \quad 12.3 < \log N_{\rm HI} < 14.5 \tag{3.9}$$

Giallongo et al. (1996) using Eq. (3.7) for a sample of high resolution lines in the redshift range $z = 2.8 - 4.1$, found a break in the column density distribution at $N_{\rm HI} = 10^{14}$ with a flat slope, $\beta_f \simeq 1.4$, for the lower column densities and a steep one, $\beta_s \simeq 1.8$, for the higher column densities. In the high column density range (beyond $\log N_{\rm HI} \sim 16$) $d\mathcal{N}/dN$ flattens and damped systems are more abundant than they should be judging from an extrapolation of the lower column density power law. Weak evolution of $d\mathcal{N}/dN$ may occur in the sense that this turnover moves slightly with redshift (Carswell et al. 1987; Kim et al. 1997) but the evidence is currently not overwhelming, given that the dip in $d\mathcal{N}/dN$ occurs in a column density range, where the determination of $N_{\rm HI}$ is least certain (see above).

The CDDF for metal lines, represented by C IV doublets, have been computed by Petitjean & Bergeron (1994). They found that a power law of the form of Eq. (3.8) with a slope $\beta = 1.6 \pm 0.1$ gives an acceptable fit of the data in the range $13.4 < \log N_{\rm CIV} < 14.8$.

3.5 The Width of Absorption Line

By measuring the line widths we may hope to gain insights into the temperature and kinematics of absorbers. High resolution spectra showed that many low column density ($N_{\rm HI} < 10^{15}$ cm^{-2}) Lyα lines do indeed show widths $b \sim 10 - 45$ km s^{-1} that are consistent with photoionisation temperatures, though some lines appear to be as wide as 100 km s^{-1}. Median Doppler parameters are around $b_{\rm med} \sim 30 - 35$ km s^{-1}, with a largely intrinsic scatter about the mean with standard deviation ~ 15 km s^{-1} (Atwood, Baldwin & Carswell 1985; Carswell et al. 1991; Rauch et al. 1992; Giallongo et al. 1996). The Doppler parameters may decrease with increasing redshift. Williger et al. (1994) have found an excess of lines with Doppler parameter lower than $b \sim 20$ km s^{-1} at $z \approx 4$.

Occasionally, a correlation between Doppler parameter and column density has been noted (Carswell et al. 1984; Hunstead et al. 1988). However, the reality of this effect has been subject to a debate which culminated in the so-called "$b - N$ controversy" (Pettini et al. 1990; Carswell et al. 1991; Peacock 1991). A reasonable explanation was obtained by simulations which pointed out how the presence of noise in a spectrum can distort weak line profiles and make difficult the detection of weak broad lines (Rauch et al. 1992, 1993). We investigated this topic and verified the absence of correlation between Doppler parameter and column density with higher resolution data (see § 3.7.2).

Data taken with the Keck telescope's HIRES instrument with a much higher signal-to-noise (S/N) ratio have basically confirmed the earlier 4m results. Hu et al. (1995) found the Doppler parameter distribution at $z \sim 3$ to be well represented by a Gaussian with a mean of 28 km s^{-1} and width $\sigma(b) = 10$ km s^{-1}, truncated below a cutoff $b_c = 20$ km s^{-1}, corresponding to a temperature $T = 24\,000$ K for thermal broadening. Kim et al. (1997) found a decrease of the median Doppler parameter with increasing redshift more pronounced for relatively strong lines ($13.8 < \log N_{\rm HI} < 16.0$), changing from 41 km s^{-1} ($\bar{z} \sim 2.3$; Kim et al. 1997) to 31 km s^{-1} ($\bar{z} \sim 3.7$; Lu et al.1996), with lower cutoffs dropping from 24 to 15 km s^{-1} over the same redshift range. Kirkman and Tytler (1997) obtain similar results for the Doppler parameters in their Keck data at even better S/N ratios, but they question the significance of the change with redshift, and find a lower, mean b of 23 km s^{-1} ($\bar{z} \sim 2.7$) with a lower cutoff $b_c = 14$ km s^{-1} at $\log N_{\rm HI} = 12.5$. However, at $\log N_{\rm HI} = 13.8$ they have $b_c = 19$ km s^{-1}, very close to the result of Kim et al. 1997, so the analyses may well be consistent. It is conceivable that the discrepancies at lower column densities arise once more from the noise bias discussed above which may affect any dataset as long as there continues to be a supply of weaker and weaker lines crossing the detection threshold with increasing S/N ratio. A dataset spanning a large redshift range, with – most importantly – a homogeneous S/N ratio would be desirable.

Analogous detailed analysis have not been conducted for metal absorption systems. Petitjean & Bergeron (1994) found a distribution of line widths for their sample of 87 C IV subcomponents spreading out to large b values. The fact that

some likely single lines have b values in excess of 20 km s^{-1} could imply that the individual clouds have turbulent or internal large scale motions. This because such broadening cannot be only thermal since it corresponds to temperatures larger than 3×10^5 K at which carbon is mostly in the C IV form.

We investigated the distribution of both Lyα and C IV Doppler parameters in the spectrum of PKS 2126-158 and in an enlarged sample of 71 C IV subcomponents respectively. Results are reported in § 3.7.

3.6 Metal Enrichment in the High Redshift Lyα Forest

Since the historic paper by Sargent et al. (1980) metal systems and Lyα forest systems were supposed to arise from two different populations of absorbers. The main reasons for this division were the absence of clustering and the pristine composition shown by Lyα clouds. Tytler (1987a) put forward the hypothesis that all absorptions were due to a single population, based on the single power law fit of the column density distribution function in the column density range $13 < \log N_{\rm HI} < 22$ (see above). We proposed, at least, a continuity of properties between populations after detecting a significant signal in the two point correlation function for Lyα lines with $\log N_{\rm HI} \geq 13.8$ (see next chapter).

On the other hand, intensive searches have been conducted for weak metal lines in low column density absorbers. Norris et al. (1983), Williger et al. (1989), and Lu (1991) used a "shift and stack" method to search for various metals in the low column density forest. To maximise the S/N ratio the spectra are shifted to the rest frame indicated by each Lyα, and added, and the expected positions of metal lines are searched for a signal. A tentative detection of C IV, corresponding to a carbon depletion relative to solar of [C/H] ≈ -3.1 was made by Lu et al. (1991). Reliable results have been obtained with Keck spectra showing that C IV associated lines can be found in 75 % of Lyα systems with $N_{\rm HI} > 3 \times 10^{14}$ cm^{-2} and more than 90 % of those with $N_{\rm HI} > 1.6 \times 10^{15}$ cm^{-2}, corresponding to a typical metallicity of $Z \sim 10^{-2} Z\odot$ (Cowie et al. 1995; Tytler et al. 1995; Songaila & Cowie 1996). Unfortunately, the detection threshold for C IV is close

enough to make it hard to determine whether the decreasing rate of detections is a genuine turnover to primordial composition below a few $\times 10^{14}$ cm^{-2}, or just a selection effect. It is interesting that in the CDM models a column density contour of 10^{14} cm^{-2} at redshift 3 delineates the transition between a continuous filamentary structure in the universe, with typical widths of less than 100 kpc, and the voids. Galaxies would have to spill metals only within the filaments to create the widespread metal-enrichment observed, and they could still have left most of the volume of the universe pristine. Thus a drop in metallicity at a few $\times 10^{14}$ cm^{-2} would not come unexpected. A relatively uniform metal abundance across the whole column density range could be an indication of an earlier phase of nucleosynthesis.

3.7 The Absorption Spectrum of the QSO PKS 2126-158

In this section I present the reduction process and the results from the analysis of the high resolution spectrum of the QSO PKS 2126–158 ($z_{\rm em} = 3.27$), (based on the paper by D'Odorico et al. 1998a). This object is one of the brightest quasars known ($m_V \simeq 17.1$) and for this reason its spectrum has been studied in many previous works (Jauncey et al. 1978; Young et al. 1979; Sargent & Boksenberg 1983; Meyer & York 1987; Sargent et al. 1988a; Sargent et al. 1989; Sargent et al. 1990; Giallongo et al. 1993, hereafter GCFT).

3.7.1 Data Reduction

Our spectrum is a sum of 8 echelle and 8 long slit observations obtained at ESO (La Silla), with the NTT telescope and the EMMI instrument (D'Odorico 1990). The spectra have been treated with the ESO reduction package, MIDAS.

The final echelle spectrum has a resolution $R \simeq 27000$. The S/N ratio per resolution element at the continuum level is doubled compared with the previous spectrum by GCFT: it ranges from S/N $\simeq$ 6 to 44 in the interval $\lambda\lambda 438 - 519$ nm, while beyond the Lyα emission it has a roughly constant value of $\simeq$ 30. The

resolution of the long slit spectrum is $R \sim 8500$.

The determination of the continuum in the QSO spectrum is a critical step because it affects the measurement of the absorption line parameters. While the continuum redward of Lyα emission can be drawn without difficulty, the high line density in the Lyman forest complicates the operation (Young et al. 1979; Carswell et al. 1982). We fulfilled this task with the help of a procedure, allowing to select automatically and in a reproducible way the regions of the spectrum free of strong absorption lines or artificial peaks (e.g. due to cosmic rays), i.e. where the RMS fluctuation about the mean becomes consistent with noise statistics. The continuum level has been estimated by spline-fitting these regions with quadratic polynomials.

The lines have been fitted with Voigt profiles convolved with the instrumental spread function, making use of a minimization method of χ^2. This step has been performed within the MIDAS package with the programme FITLYMAN (Fontana & Ballester 1995). The number of components of each absorption feature is assumed to be the minimum required to give a reduced $\chi^2 < 1$ (corresponding to a confidence level $P \gtrsim 50\%$).

All the lines shortward of the Lyα emission not identified as due to metals have been fitted as Lyα and Lyβ. For a further control, we used the blue, lower-resolution spectra to search for Lyβ, Lyγ and Lyδ lines in correspondence to the stronger Lyα lines ($\log N_{\rm HI} \geq 14$). In few cases, we could ascertain that the absorption was not Lyα. Such lines are listed as unidentified and they probably belong to metal systems still to be recognized.

3.7.2 Lyman Alpha Statistics

We obtained a statistically well-defined sample of Lyα lines in the region between the Lyβ and Lyα emissions ($\lambda\lambda 438.0 - 519.1$ nm). The Lyα lines affected by the proximity effect ($8h^{-1}$ Mpc from the QSO; see Bajtlik, Duncan & Ostriker 1988; Lu et al. 1991) and those associated with metal systems (indicated as MLyα) are excluded from the sample.

The distribution of lines in the $b - \log N$ plane is shown in Fig. 3.3. The lack of lines at the top left corner can be ascribed to an observational bias: the line

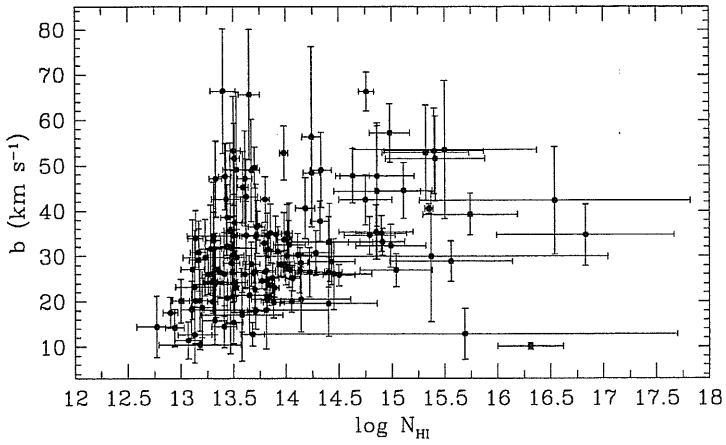

Figure 3.3: Plot of the Doppler parameter b vs. the logarithmic column density $\log N_{\rm HI}$ for the Lyα line sample

selection criterion is based on the depth of the centre of the line an thus it tends to miss lines with low column density and large Doppler width.

The dataset can be considered virtually complete (for typical b values) for $\log N_{\rm HI} \gtrsim 13.3$.

A considerable fraction of the lines with $\log N_{\rm HI} \gtrsim 14.5$ belongs to complex saturated systems. As a consequence, the deblending choices and therefore the fitting parameters may be not always unique. A considerable improvement can be obtained when the simultaneous fit of the saturated Lyα and the corresponding Lyβ line is possible. Unfortunately, due to the high density of Lyα lines at these redshifts, Lyβ absorptions with an uncontaminated profile are rare and the uncertainty for most of the lines in the right region of the diagram 3.3 cannot be removed.

Simulations carried out by Fontana and Ballester (1995) show that, for isolated, unsaturated lines at S/N $\gtrsim 8$, the parameters given by the fitting procedure are quite close to the "true" value, with a small and symmetric scatter around it. At column densities larger than $10^{14.5}$ cm^{-2} the Lyα lines are saturated and b and N correlate strongly, increasing the uncertainties.

By comparing our list of lines with that published by GCFT, we have verified on real data how the $b - \log N$ diagram changes when the average S/N ratio is

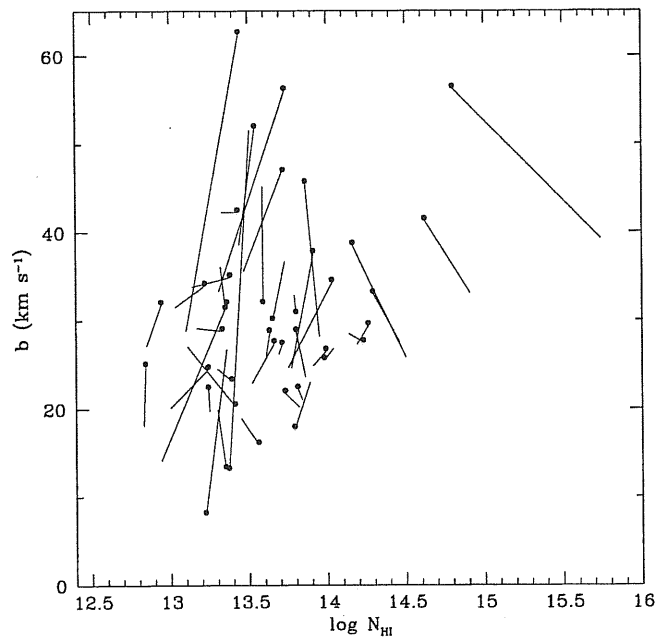

Figure 3.4: Migration diagram illustrating the effect of improving the S/N ratio on the determination of the absorption lines parameters. The points correspond to the GCFT determination

almost doubled, checking the trends expected on the basis of simulations.

To carry out a meaningful comparison, we have considered only the isolated Lyα lines in common between the two line lists, at wavelengths $\lambda > 475$ nm. In this range the S/N per resolution element is $\gtrsim 8$.

Figure 3.4 shows the individual trajectory of each absorption line in the plane $b - \log N$. GCFT parameter values are indicated by the solid black circles and our values correspond to the end of the adjoining line.

Saturated lines tend to move keeping approximately constant the equivalent width value, as observed in the simulations (Fontana & Ballester 1995). Often they correspond to complex features. If the improved S/N is not sufficient for a proper deblending, but the better definition of the wings forces a fit with a lower b parameter, then they move toward higher column densities.

Lines with low column density in regions with low S/N ratio have poorly defined profiles and are very sensitive to the continuum level. The S/N increase,

together with a choice of a slightly lower continuum with respect to GCFT, has induced a migration from high b values and small column densities toward smaller Doppler widths.

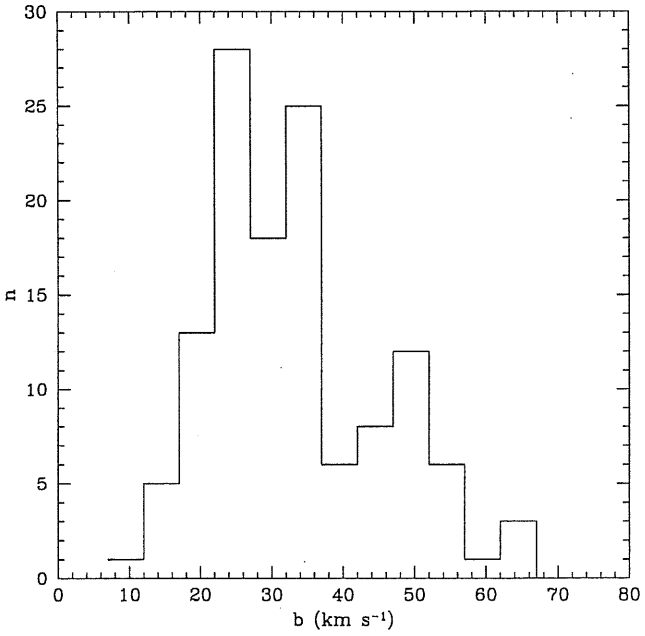

Figure 3.5: Doppler parameter distribution for the Lyα line sample, out of 8 Mpc from the quasar PKS 2126–158

The Doppler width distribution of the complete ($\log N_{\rm HI} \geq 13.3$) sample of Ly$\alpha$ lines is shown in Fig. 3.5.

Making the assumption that line broadening is due exclusively to thermal motion, the relation between Doppler parameter and temperature, $b^2 = 2\,k\,T/m$, (where k is the Boltzmann constant and m is the hydrogen mass) implies that the peak value $b \simeq 25$ km s^{-1} corresponds to a cloud temperature of $T \sim 38000$ K.

The Lyα sample contains 21 lines, i.e. $\simeq 14.5$ %, with $10 \leq b \leq 20$ km s^{-1}, and 12 lines, i.e. $\simeq 8.3$ %, with $15 \leq b \leq 20$ km s^{-1}. Such percentages are almost halved compared with those found in the previous work by GCFT. Nonetheless, it has to be noted that the peak value of the b distribution is roughly the same for the two samples.

As we already mentioned in § 3.5, the nature of the distribution of Lyα clouds Doppler parameters has led in the past to controversies. Pettini et al. (1990) found Lyα lines with a median Doppler parameter of 17 km s^{-1} and a strong intrinsic correlation between Doppler width and column density. From these results a scenario of very cool, dense and practically neutral clouds emerged, in contrast with previous models. Starting from data with similar resolution, Carswell et al. (1991) following different selection and analysis criteria obtained significantly larger average and median Doppler parameters and, above all, no $b - N$ correlation. Our median b value is intermediate between the results of Pettini et al. and that of Carswell et al. and it agrees with Keck result at very high resolution (Hu at al. 1995; Lu et al. 1996).

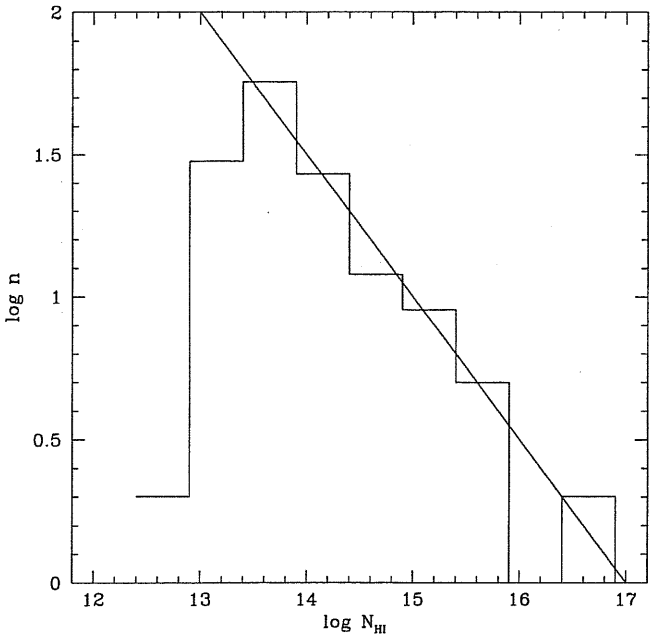

Figure 3.6: Column density distribution of the Lyα lines out of 8 Mpc from the quasar PKS 2126–158. The overplotted solid line represents a power-law distribution, $\propto N^{-\beta}$, with $\beta = 1.5$ (Hu et al. 1995; Giallongo et al. 1996)

The column density distribution of our Lyα line sample is shown in Fig. 3.6. As shown in Fig. 3.3, for values $\log N_{\rm HI} \lesssim 13.3$ a selection bias is expected: only lines with small Doppler parameters are detectable. This is confirmed by the drop

of the column density distribution below this $\log N_{\rm HI}$ value. The shape of the distribution is in agreement with the power-law fit $N^{-\beta}$, with $\beta \simeq 1.5$, obtained in the same column density range in other works (Hu et al. 1995; Giallongo et al. 1996).

The number density of lines per unit redshift, in the wavelength interval $\lambda\lambda 438-506.5$ nm for Lyα absorptions with $\log N_{\rm HI} \geq 14$, is in good agreement with the cosmological redshift distribution (Giallongo et al. 1996).

3.7.3 Metal Systems and the UV Background

As a first guess, heavy-element systems have been identified on the basis of the line list presented by GCFT. Eventually, other element lines have been added to the previously found systems, making use of a list of the lines most frequently observed in QSO absorption spectra, derived from Morton (1991) and Verner et al. (1996).

As for the Lyα lines, the operation of fitting has been carried out within the context FITLYMAN of the reduction package MIDAS. In the case of the complex profiles observed for the heavy-element lines, a minimum number of components needed to obtain a $\chi^2 < 1$ and a good fit was deduced from the C IV profiles (or, in one case at low redshift, from the Mg II one).

Then an identification programme, based on the method of Young et al. (1979), has been applied to the newly observed lines that seemed not to belong to existing systems.

Besides the systems previously found at lower resolutions (Young et al. 1979; Sargent et al. 1988a; GCFT), we show two new systems at redshifts $z_{\rm abs} = 2.5537$ and $z_{\rm abs} = 3.2165$ identified on the basis of the CIV doublet, whose column densities and equivalent widths were too weak to be detected in previous observations. Six of the twelve systems show a multicomponent substructure, with a velocity extent up to ~ 350 km s^{-1}. We have not adopted any velocity window, or other strict rules for the classification into systems versus sub-systems. We have simply considered as one system those groups of lines for which a simultaneous fit was required because of the superposition of the various line profiles.

For the systems at $z_{\rm abs} = 2.8195$ (I) and $z_{\rm abs} = 2.9072$ (J) the reliability of

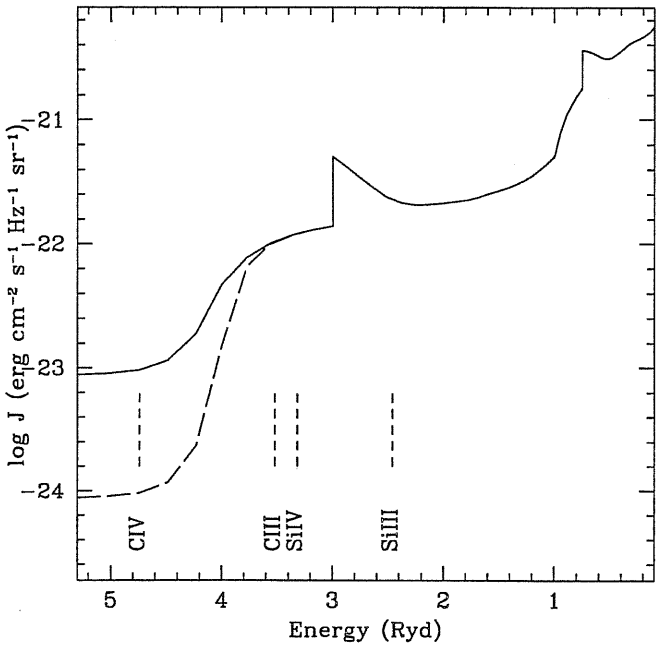

Figure 3.7: The shape of the UV background at the epoch $z = 2.9$. Continuous line: as computed by Haardt and Madau (1996); dashed line: with the jump at the He II edge, $B \equiv J(4\ Ryd)/J(4.5\ Ryd)$, increased by a factor of 10. The ionisation potentials of a few ions are shown

the fit of the Lyman series allows a measure of the metallicity. In the Lyman limit system at $z_{\rm abs} = 2.9675$ (K) it is possible to put an upper limit to it on the assumption of $\log {\rm N_{HI}} \geq 17$. We have used the standard photoionisation code CLOUDY (Ferland 1996), following the same approach described in detail in Savaglio et al. (1997).

In the analysis of the system J, if the standard UV background at $z = 2.9$ (Haardt & Madau 1996) is adopted and the size of the cloud is assumed not smaller than 20 Kpc, we end up with an unplausibly high overdensity of silicon with respect to carbon, $\gtrsim 50$. To obtain a more realistic [Si/C] $\lesssim 1$ the cloud size should decrease substantially below 10 kpc. Incongruities of this type are not uncommon, as reported by Songaila and Cowie (1996) and Savaglio et al. (1997). Several processes may increase the Si IV/C IV ratio: a photoionising background dominated by local sources, non-equilibrium temperatures and non-uniform radiation fields

(Giroux & Shull 1997). In Savaglio et al. (1997) and here we have explored the possibility of an enhanced break at 4 Ryd in the metagalactic background radiation, as could be originated by a contribution of primeval galaxies in addition to the standard QSO background. With a jump, $B \equiv J(4\ Ryd)/J(4.5\ Ryd)$, increased of a factor 10 at 4 Ryd (Fig. 3.7), the observed line ratios turn out to be compatible, assuming a cloud size of 30 kpc, with a metallicity -2.5 dex solar and an overabundance of silicon with respect to carbon of a factor 5-6.

The analysis of systems I and K, under the assumption of the same UV background with an enhanced jump at 4 Ryd and cloud sizes $\gtrsim 15$ kpc, provides metallicities of $\simeq -2.2$ and $\lesssim -2.7$, respectively. System I, again, requires an overabundance of silicon with respect to carbon of a factor 4 that would become much larger if the standard UV background is assumed.

3.7.4 Statistics of C IV Systems

In order to investigate the statistical properties of C IV systems, an enlarged sample has been created by merging the data of PKS 2126–158 with those recently obtained at similar resolution for Q0000–26 (Savaglio et al. 1997) and Q0055–269 (Cristiani et al. 1995). All the spectra have been analysed in a homogeneous way.

The distribution of C IV column densities is shown in Fig. 3.8. It can be described as a single power-law distribution, $d\mathcal{N}/dN \propto N^{-\beta}$, with $\beta = 1.64$ (in agreement with Petitjean & Bergeron 1994), down to a completeness limit $\log N_{\rm CIV} \sim 13.3$.

The distribution of the Doppler parameters for the same C IV lines has a mode of $b \simeq 14$ km s^{-1} and a standard deviation $\sigma_b \simeq 8$ km s^{-1}. A simple derivation of the expected $< b_{\rm CIV} >$ from the corresponding $< b_{\rm Ly\alpha} >$, with the hypothesis of pure thermal broadening with a single temperature would provide $< b_{\rm CIV} > \simeq 7$ km s^{-1}. This indicates either that turbulent motions are present or some of the lines are blends of subcomponents, which would be separated only by observations with higher spectral resolution and S/N.

Figure 3.9 shows the equivalent widths, estimated from the b and $N_{\rm CIV}$ values derived for each subcomponent by profile fitting, versus the column density. It

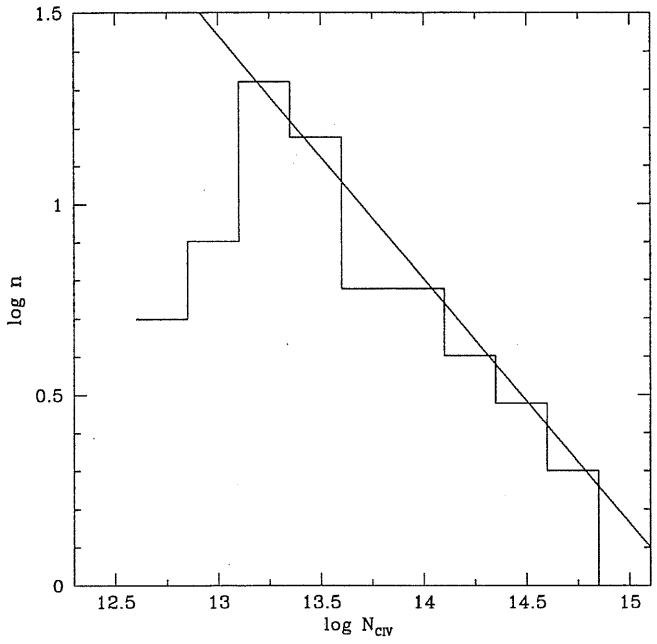

Figure 3.8: Histogram of the logarithmic number of C IV absorptions as a function of the column density. The solid line represents a power-law: $N_{\rm CIV}^{-\beta}$, $\beta = 1.64$ (Petitjean & Bergeron 1994)

can be seen that most of the lines are on the linear part of the curve of growth, at least for $N_{\rm CIV} < 10^{14}$ cm^{-2}. Curves of growth have been drawn for values of the Doppler parameter $b_{\rm CIV} = 6, 14, 22$ km s^{-1}.

3.8 Conclusions

In this chapter, I have presented some of the results obtained from the analysis of QSO absorption spectra. I have discussed not only the most recent observational results, but also how they have evolved in the course of the years as higher resolution and S/N spectra increased the available dataset.

In the case of the Lyα forest, a valuable improvement has been registered with the advent of the HIRES spectrograph at the Keck telescope. In particular, the detection of weak C IV absorptions associated to Lyα forest lines with column density as low as 5×10^{14} cm^{-2} has marked a turn in the picture of low column

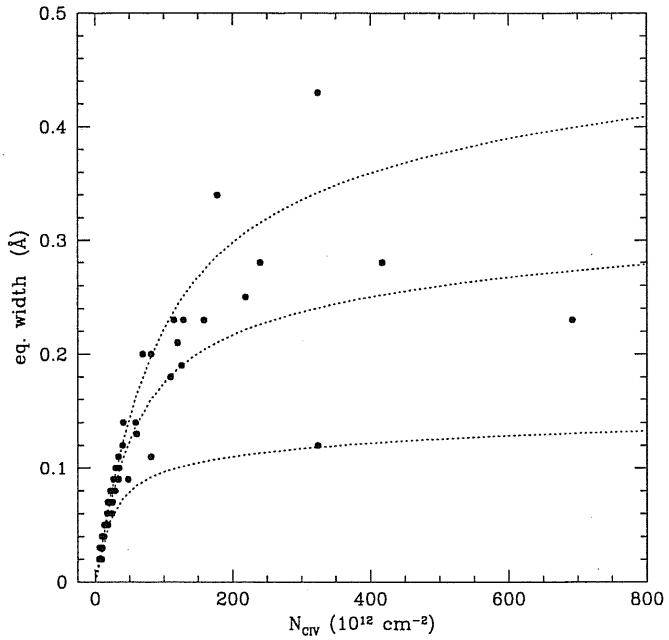

Figure 3.9: Rest equivalent width of single C IV system components as a function of the column density (in units of 10^{12} cm^{-2}). The overplotted curves of growth have been computed for values $b = 6, 14, 22$ km s^{-1} of the Doppler parameter (Spitzer 1978; Press & Rybicki 1993)

density Lyα absorbers. Another fundamental piece of information in this context, obtained with the 4m class ESO New Technology Telescope, has been given by our detection of a significant clustering signal for the Lyα forest lines at slightly lower column densities, as described in the next chapter.

Metal systems are less abundant than Lyα forest lines so it is harder to compile statistically significant sample of data.

We performed the reduction and studied the distribution of line parameters of a high-resolution spectrum of the quasar PKS 2126–158 ($z_{\rm em} = 3.27$) observed in the context of the ESO Key Program on QSO absorption lines. The analysis of the Lyα forest shows that:

• The peak value of the Doppler parameter distribution is $\sim$ 25 km s^{-1}. This value is large enough to be in agreement with a standard ionisation model, but not too high ($b \gtrsim 30$) to make the Lyα clouds contribution to baryonic matter exceed

the standard nucleosynthesis value. This because b can be related to the ionising parameter (Donahue & Shull 1991) and the barion density can be written as a function of the average ionisation fraction (Hu et al. 1995).

• The plot of the statistical sample of Lyα lines in the $b - \log N$ plane shows no significant correlation between the two parameters.

• The column density distribution shows a cutoff, due to incompleteness and blending, for $\log N_{\rm HI} \lesssim 13.3$, for higher values it is well described by a power-law $d\mathcal{N}/dN \propto N^{-\beta}$ with $\beta = 1.5$ (Hu et al. 1995; Giallongo et al. 1996).

• The number density of lines per unit redshift is in agreement with a standard power-law evolution of the type $d\mathcal{N}/dz \propto (1+z)^{\gamma}$ with $\gamma = 2.7$ (Giallongo et al. 1996).

From the analysis of metal absorptions the following results have been obtained:

• Two new C IV absorption systems have been detected at $z = 3.2165$ and $z = 2.5537$.

• A mean metallicity of ~ -2.5 dex solar has been found using the metal systems at $z = 2.8195$, $z = 2.9072$ and $z = 2.9675$. In order to make the column densities of the intervening systems compatible with realistic assumptions about the cloud sizes and the [Si/C] ratios, it is necessary to assume an increase of a factor 10 in the jump at 4 Ryd of the standard spectrum of the ionising UV background (Haardt & Madau 1996).

• Merging our data on PKS 2126–158 with those obtained at comparable resolution for the two quasars Q0055–26 and Q0000–26, has provided a relative large sample (71 doublets) of C IV absorptions, complete down to a column density $\log N_{\rm CIV} \simeq 13.3$. The C IV column density distribution is well fitted by a single power-law, with $\beta = 1.64$. The mode of the Doppler parameter distribution is $b_{\rm CIV} \simeq 14$ km s^{-1}.

Chapter 4

The Clustering Properties of Absorbers

In this chapter, I presént a detailed description of the statistical techniques and the results of the study of the spatial distribution properties of QSO absorbers.

In the first part, I focus on metal absorption systems, and in particular on C IV systems: they were found to auto-correlate since the first observations and for this reason they were associated to the galactic population. The study of the clustering properties of the Lyα lines and its evolution across the years are described in Section § 4.4. The last section is dedicated to a new technique that recovers the underlying density field from the transmitted flux. Once it will become easily and reliably implementable, this algorithm will represent a powerful tool to constrain cosmological model of structure formation.

4.1 Introduction

The evolution with redshift of the spatial distribution of matter gives a fundamental hint on the nature of the processes at the origin of the formation of galactic structures.

In the last two decades, redshift surveys provided a wealth of information about local galaxies. They revealed a galaxy distribution consisting of a web of thin, large walls outlining the edges of voids tens of Mps in diameter, with no

strong evidence of inhomogeneities on larger scales. The clustering of galaxies is found to increase with luminosity and to depend on morphology and color, with early type and red galaxies more clustered than spiral, irregular and blue galaxies, respectively. This is due in part to segregation effects that have to be understood on the basis of a convincing structure formation theory. In general, it is accepted that a power-law reasonably represents the behaviour of the two-point correlation function (see definition below) for the observed galaxies in the range of scales $0.1\ h^{-1}$ Mpc $\lesssim r \lesssim 10\ h^{-1}$ Mpc,

$$\xi_{gg}(r) = \left(\frac{r}{r_g}\right)^{-\gamma}, \tag{4.1}$$

where the exponent $\gamma \simeq 1.8$ and the correlation length $r_g \simeq 5\ h^{-1}$ Mpc.

As previous studies were confined to the nearby universe, nowadays new observational resources permit to extend the correlation analysis to samples as deep as $z \sim 1.5$ (e.g. the Canada-France Redshift Survey, Le Fèvre et al. 1996; the Canadian Network for Observational Cosmology 2, Carlberg et al. 1998). This is the best way to constrain different cosmological models, since the prediction for the clustering properties of galaxies at low redshift are quite similar, while they strongly differentiate at high redshift.

To go at even higher redshifts other kinds of objects have to be considered. Quasars have been the first class of sources used to obtain direct information about clustering at high redshift. They have some appealing properties when compared to galaxies as cosmological probes of the intermediate-linear regime of clustering: they have a rather flat redshift distribution, their point-like images are less prone to the surface-brightness biases typical of galaxies and they sparse-sample the environment. The standard objection with regard to the information content in surveys of QSOs and radio galaxies is that, since they are exotic objects, one risks learning possibly more about the formation and evolution of super-massive black holes in galactic nuclei, which is something interesting *per se*, than about cosmology.

An exciting progress in this field has been achieved with the recent compilation of the sample of the so called Lyman Break Galaxies (LBGs), selected photometrically at $z \sim 3$ (Steidel & Hamilton 1993). The clustering strength of this

galaxy population is comparable to that of present-day spiral galaxies (Steidel et al. 1998; Giavalisco et al. 1998; Adelberger et al. 1998) and it is as large or larger than any measured in recent intermediate-redshift galaxy samples ($0.3 \lesssim z \lesssim 1$). By comparing the observed correlation length of LBGs to that of the mass predicted from CDM theory a linear bias of $b \sim 4.5$ (for $q_0 = 0.5$) is inferred. The strong clustering and the bias are consistent with biased galaxy formation theories and provide additional evidence that these systems are associated with massive dark matter halos.

Thus, galaxies are exotic objects too, in the sense that they are also biased tracers of mass.

The absorption spectra of QSOs, as already said, are sensitive probe of the gaseous baryonic component of the universe, thus we can investigate the spatial distribution properties of normal matter at redshifts larger than ~ 1 by properly measuring and interpreting the clustering properties of absorption lines.

4.2 The Two Point Correlation Function: General Definition

To study the nature of QSO absorbers and their relationship with local cosmological objects, one of the first ideas was to compare their spatial distribution with the distribution of galaxies, for obvious reasons the first estimator considered was the two point correlation function (TPCF).

The definition of the TPCF, ξ, is related to the joint probability

$$\delta^{(2)} P = n^2 \delta V_1 \, \delta V_2 \, [1 + \xi(r_{12})] \tag{4.2}$$

of finding an object in the volume element δV_1 and another one in δV_2, at separation r_{12}. In Eq. (4.2) the factorization of the n^2 term (n being the object mean number density) makes $\xi(r)$ a dimensionless quantity and the total probability turns out to be normalized to the square of the total number of objects in the distribution. According to its definition, the value of the correlation function is a measure of the non-random behaviour of the distribution and, for an isotropic clustering, depends only on the modulus of the separation vector $\mathbf{r}_{12}$. In partic-

ular, object positions are said to be correlated if $\xi(r) > 0$ and anticorrelated if $-1 \leq \xi(r) < 0$, while a Poissonian distribution is characterized by $\xi(r) = 0$ at any separation.

The TPCF is known to be a satisfactory estimator when used to investigate weak clustering on scales considerably smaller than the total interval covered by the data, which is the first domain explored in the case of QSO absorption lines. The binning, intrinsic to this method, causes a loss of information, but the ease in visualizing its results and in including observational effects in the computing codes have made the TPCF one of the favorite statistical estimators in cosmology.

4.3 Spatial Distribution of Heavy Element Lines

4.3.1 The Measure of Clustering along a Single Line of Sight

Here, I make an historical excursus of the results on this issue, by presenting also the different algorithms adopted to estimate the TPCF. The original work in this section has been published in D'Odorico et al. (1998a).

Metal line absorption systems, in particular systems identified by the C IV $\lambda\lambda$ 154.8, 155.0 nm doublet, were found to show fine splitting among components on a scale of $\sim$ 150 km s^{-1}, as soon as sufficient resolution spectra became available (Boksenberg & Sargent 1975). Sargent and collaborators (1980), having a statistically significant sample of absorption lines, developed the practical tools to estimate the TPCF.

Basically, the cosmological distance between two absorbers at redshifts z_1 and z_2 has to be computed. The matter is complicated by the change of distance scale in the universe with time. It is then appropriate to calculate separations in comoving coordinates. The comoving distance of an object at redshift z measured at the present epoch is:

$$S_0(z) = \frac{c}{H_0} \int_1^{1+z} du[2\sigma_0 u^3 + (1 + q_0 - 3\sigma_0)u^2 + \sigma_0 - q_0]^{-1/2} \qquad (4.3)$$

where $\sigma_0 = \Omega_0/2$. Then at any earlier epoch when the scale factor is a, the distance is

$$S(z) = S_0(z)(a/a_0) = S_0(z)/(1+z). \tag{4.4}$$

For cosmological models with $\Lambda = 0$ so that $\sigma_0 = q_0$, and $q_0 = 1/2$, Eq. (4.3) becomes:

$$S_0(z) = \frac{2c}{H_0}[1 - (1+z)^{-1/2}]. \tag{4.5}$$

The splittings of two lines at redshift z_1, z_2 can be described as:

1. The actual distance between the absorbers were they to survive to the present epoch fixed in cosmological coordinates. This is given by the difference in the S_0 coordinate of Eq. (4.3) which is, for small redshift differences $\Delta z = |z_1 - z_2|$ and in units of the present Hubble radius ($c = H_0$),

$$s_0 = |S_0(z_1) - S_0(z_2)| = \Delta z(1+z)^{-1}(1+2q_0z)^{-1/2}, \tag{4.6}$$

where $z = (z_1 + z_2)/2$.

2. The separation in distance of the two absorbers measured at the cosmic time when the photons we receive now were passing through them. Note that this distance s is measured in Hubble radii and that the Hubble constant at redshift z is $H(z) = H_0(1+z)(1+2q_0z)^{1/2}$. Thus

$$s = s_0(1+z)^{-1}H(z)/H_0 = s_0(1+2q_0z)^{1/2}. \tag{4.7}$$

The velocity splittings due to the Hubble flow between two such absorbers measured at the present epoch and at the cosmic time when the photons were passing through the absorbers are: $v_0 = c\,s_0$ and $v = c\,s$ respectively.

Given a redshift range $[Z^a, Z^b]$, it is possible to define a distance range $[S_0^a, S_0^b]$. For each line, the distance S_0 can be evaluated and then the separation $s_0 = S_0^i - S_0^j$ for each line pair. With L lines, $L(L-1)/2$ values of s_0 are obtained, which are cast into bins of size Δs_0. Then to allow for the fact that a finite span of data necessarily permits more small separations than do large ones, the number of separations s_0 can be divided by the response function:

$$\psi(s_0) = (S_0^b - S_0^a - s_0)/(S_0^b - S_0^a). \tag{4.8}$$

With this correction, a Poissonian distribution of cloud separations will then appear flat, with an average of $m_b = L(L-1)\Delta s_0/(S_0^b - S_0^a)$ line pairs in a bin. If we find $m(s_0)$ line pairs in a bin, then approximately, $\xi(s_0) = m(s_0)/m_b - 1$.

Applying this statistics to a sample of three high resolution QSO spectra, Sargent et al. (1980) found a massive peak on a velocity scale $90 - 180$ km s^{-1}, they interpreted this as possibly due to multiple clouds within a single galaxy halo, rather than to galaxy-galaxy correlations.

To further investigate this issue, first a sample of 33 QSOs spectra (Young, Sargent & Boksenberg 1982a) and then a larger one of 55 QSOs spectra (Sargent, Boksenberg & Steidel 1988) have been considered. The sample is formed by around 200 C IV redshifts ranging from 1.27 to 3.53 with a median $z \simeq 1.87$.

The TPCF is generally flat, as would be expected for absorbers randomly distributed along the line of sight. In addition, there is a significant clustering of the redshifts on velocity scales in the range $200 - 600$ km s^{-1} in the rest frame of the absorbers. The mean amplitude of the correlation function over this range is 11.5 ± 1.3.

Considerations on the rotational amplitudes and random motions observed in luminous galaxies suggests that it is unlikely that the observed C IV splittings are due to motions within galaxies; it is more likely that we are observing the tip of the galaxian TPCF. A comparison of the observed C IV TPCF with the simplest theory of the evolution of hierarchical clustering shows that the model is in tolerable agreement with the observations.

Heisler, Hogan & White (1989) reconsidered the sample of 55 QSOs spectra of Sargent, Boksenberg and Steidel (1988) and analysed the clustering of C IV absorptions adopting a new estimate for the redshift space TPCF,

$$\xi(v, \Delta v) = \frac{n_{\rm obs}(v, \Delta v)}{n_{\rm exp}(v, \Delta v)} - 1, \tag{4.9}$$

where $n_{\rm obs}$ is the observed number of line pairs with velocity separations between v and $v + \Delta v$ and $n_{\rm exp}$ is the number of pairs expected in the same velocity interval from a random distribution in redshift. First, the set of bins in v is decided requiring approximately equal numbers of expected random pairs in each bin once statistics were accumulated for the sample as a whole. The following steps are

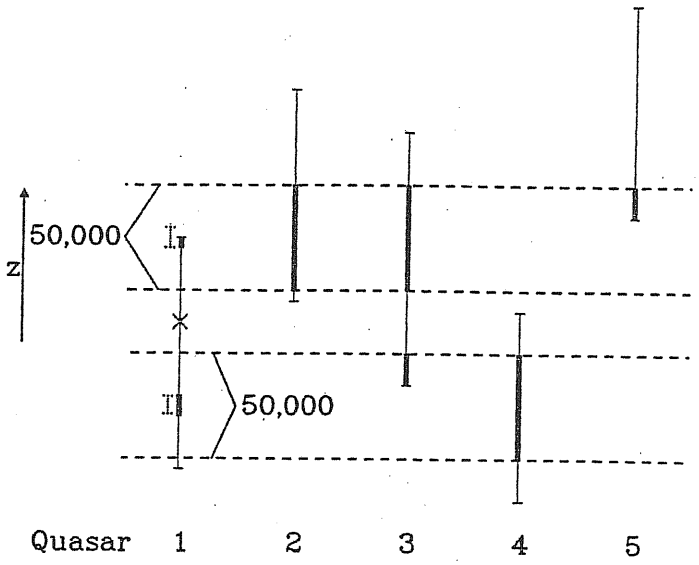

Figure 4.1: Pictorial explanation of the counting algorithm described in the text. Starting from a particular reference system in quasar 1, marked with an (x), they count the number of systems within the small dotted bins that are placed equidistantly in velocity above and below the reference system. Below the reference system the full width of the bin is accessible; above, the upper redshift limit for quasar 1 lies within the bin, and thus the accessible range is narrower. The expected number of random systems is found by centering over each of the dotted bins a larger interval of 50 000 km s^{-1}; these intervals are shown by dashed lines. Reproduced from Heisler, Hogan & White (1989)

then followed to estimate $\xi(v, \Delta v)$ in each bin:

1. For each absorption system in each QSO, the range of redshifts whose velocity separations from the system lie within the current bin and for which C IV systems could have been detected is determined.
2. The count of observed pairs within the current bin is augmented by the number of C IV systems found in the varius redshift intervals.
3. The random density of absorbers at the position of each redshift interval is estimated by dividing the number of absorption systems within $\pm 25,000$ km s^{-1} of its center in the 54 *other* QSOs by the total pathlength in this redshift range for which C IV lines could have been detected in them (see Fig. 4.1). This density

is then multiplied by the length of the redshift interval under consideration to estimate the expected number of random systems it should contain.

They clearly detect correlations in the lowest velocity bin ($v < 1000$ km s^{-1}). A significant signal is also shown in the range 1000 km s^{-1} $< v <$ 10,000 km s^{-1}, even if it can be due to the presence of a large cluster of systems in just one of the QSO of the sample.

Petitjean and Bergeron (1990, 1994) investigated the clustering properties of Mg II and C IV doublets at average redshifts $z \sim 1$ and $z = 2.65$ respectively, on scales smaller than 1000 km s^{-1}, following Sargent et al. (1980). They found that the distribution of the velocity separations Δv normalized for the width of the bin, is well fitted by a sum of two Gaussians distributions,

$$f(\Delta v) = \sqrt{\frac{2}{\pi}} \left[\frac{A}{\sigma_a} \exp\left(-\frac{\Delta v^2}{2\sigma_a^2} \right) + \frac{1-A}{\sigma_b} \exp\left(-\frac{\Delta v^2}{2\sigma_b^2} \right) \right]. \qquad (4.10)$$

where A weights the two Gaussians. The fitted velocity dispersions for the Mg II and C IV sample are $\sigma_a = 80$ km s^{-1} and $\sigma_b = 390$ km s^{-1}, $\sigma_a = 109$ km s^{-1} and $\sigma_b = 525$ km s^{-1}, for the Mg II and C IV respectively.

They suggested that the small velocity scale may reflect relative motion of clouds within the same galactic halo, while the large velocity separations could correspond to the clustering of several haloes. Another interpretation (Haehnelt, Steinmetz & Rauch 1996; Rauch et al. 1996) depicts C IV as groups of protogalactic clumps ($M_{\rm baryons} \sim 10^9\ M_\odot$) with small internal velocity dispersion, embedded in flattened large-scale "walls" expanding with the Hubble flow.

Works by Womble, Sargent and Lyons (1996) and Songaila and Cowie (1996) analysed clustering of lower column density samples detecting a lower signal on the same scales studied by Petitjean and Bergeron (1994).

Quashnock, Vanden Berk and York (1996) analysed line of sight correlations of C IV and Mg II absorption line systems on large scales, using an extensive catalog of 2200 heavy-element absorption line systems in over 500 QSO spectra (York et al. 1991). These authors found that the absorbers are clustered, like galaxies, on comoving scales of $50 - 100\ h^{-1}$ Mpc.

This high redshift (z from 1.5 to 3.5) superclustering of absorbers is on the same comoving scale as that traced by the voids and walls of galaxy redshift surveys in the local universe. It thus appears that the absorbers are tracing the same structure as that traced by galaxies on very large scales, and that they are effective probes of very large scale structure in the universe. Indeed, superclustering has been detected by other studies, in particular using close multiple lines of sight (Dinshaw & Impey 1996; Williger et al. 1996).

Quashnock and Vanden Berk (1998) considered the same large compilation of heavy element absorption systems to study the clustering on comoving scales r from 1 to 16 h^{-1} Mpc. They applied a likelyhood method that does not require the binning of the data but the introduction of a model function for the TPCF and they found that, for the C IV sample as a whole, the absorber line of sight correlation function is well fitted by a power law of the form $\xi(r) = (r/r_0)^{-\gamma}$, with maximum likelyhood values of $\gamma = 1.75^{+0.50}_{-0.70}$ and comoving $r_0 = 3.4^{+0.7}_{-1.0}$ h^{-1} Mpc. The clustering of absorbers at high redshift is thus of a form that is consistent with that found for galaxies and clusters, at low redshifts, on megaparsec scale. They also found that the amplitude of the correlation function, characterized by the correlation length r_0, is growing with decreasing redshift.

We faced this issue considering three QSO spectra observed at high resolution in the context of an ESO Key Programme devoted to the study of QSO absorption systems at high redshift: PKS 2126–158 (D'Odorico et al. 1998a), Q0000–26 (Savaglio et al. 1997) and Q0055–269 (Cristiani et al. 1995). (See Tab. 4.1 for the characteristics of these objects.)

To compute the TPCF of the sample of 71 C IV doublets, we adopted the definition given in Eq. (4.9). The number of expected pairs in a random distribution, $n_{\rm exp}$, is obtained as the average of the results of 5000 numerical simulations of the observed number of redshifts, trying to account for all the relevant cosmological and observational effects. In particular the set of redshifts is randomly generated in the same redshift interval as the data according to the cosmological distribution $\propto (1+z)^{\gamma}$, where γ has been taken equal to -1.2 (Sargent, Boksenberg & Steidel 1988). Observed components with a velocity splitting $\Delta v < 30$ km s^{-1} have been merged, while simulated ones have been excluded in the estimate of $n_{\rm exp}$, because

of the intrinsic line blending due to the typical widths of the C IV lines.

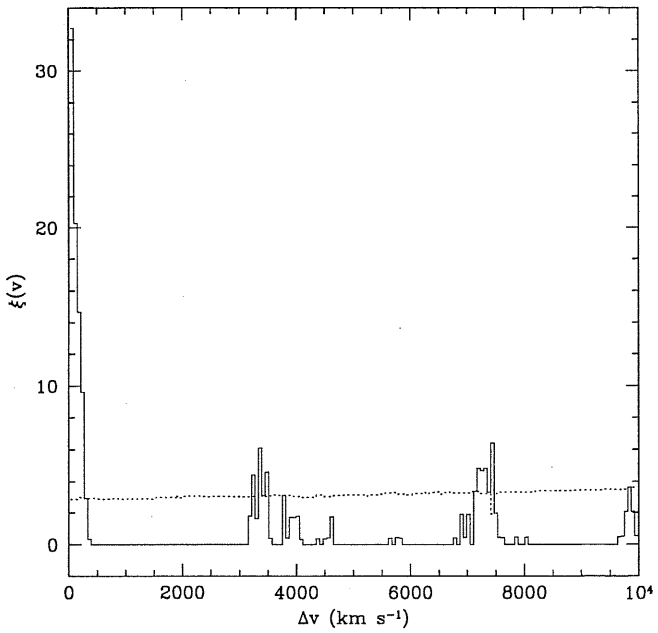

Figure 4.2: Two point correlation function at large velocity separations for the 71 C IV systems in our sample. The dotted line shows the 95 % confidence level

The TPCF for the C IV clouds is presented in Fig. 4.2 with 60 km s^{-1} bins. A strong clustering signal is detected at small velocity separations ($\Delta v < 1000$ km s^{-1}). At larger scales no significant signal is found, in particular the peaks observed in the TPCF at $\sim$ 3360 km s^{-1}, $\sim$ 7140 km s^{-1} and at $\sim$ 9840 km s^{-1} are aliases corresponding to the coupling of the low Δv (high frequency) power with the window function.

The present data have adequate resolution to allow a further investigation of the distribution on scales smaller than 1000 km s^{-1}.

The TPCF for velocity separations in the range 30 km s^{-1} $\leq \Delta v \leq$ 1000 km s^{-1} is shown in Fig. 4.3. In this velocity interval 48 pairs are observed, while $\sim$ 10 are predicted for a homogeneous distribution.

Our results are in agreement with those previously described, carried out at comparable resolution (Petitjean and Bergeron 1994; Womble, Sargent & Lyons 1996; Songaila & Cowie 1996).

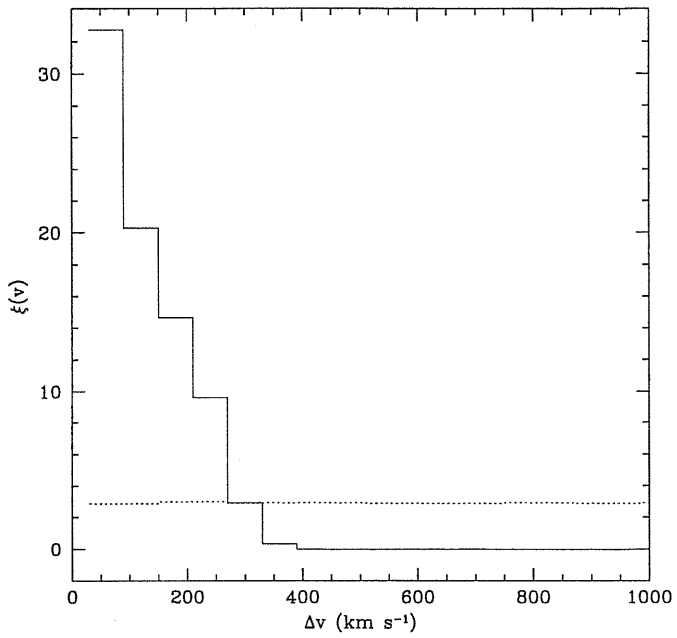

Figure 4.3: Zoomed two point correlation function at small velocity separations for the C IV systems in our sample. The dotted line shows the 95 % confidence level

The velocity scale at which the maximum clustering signal is observed is comparable with the extension of the complex metal absorption features in the spectrum. As already noticed by Petitjean and Bergeron (1994), this could be an evidence of the fact that what we are measuring is not the clustering of galaxy-like objects but of gas clouds within the same galactic halo.

If the resolution of the present data is degraded to the level typical of the older investigations (Sargent, Boksenberg & Steidel 1988; Heisler, Hogan & White 1989; e.g. a resolution of 77 km s^{-1}), a compatible result is obtained, with $\xi(0 < \Delta v < 1000 \text{ km s}^{-1}) \simeq 2.0 \pm 0.9$.

The correlation on scales larger than 1000 km s^{-1} reported by Heisler, Hogan and White (1989) is not reproduced. It has to be noted, however, that it appears to be the result of the inclusion in their sample of one "exceptional" object, Q0237-233.

The analysis of the clustering properties by means of the TPCF does not take

into account the information on the equivalent width or the column density of the lines. We investigated this aspect in the following section.

Correlations with the Column Density

It has long been assessed that a correlation exists between the equivalent width of metal absorption systems and their number of components (Wolfe 1986; Petitjean & Bergeron 1990, 1994).

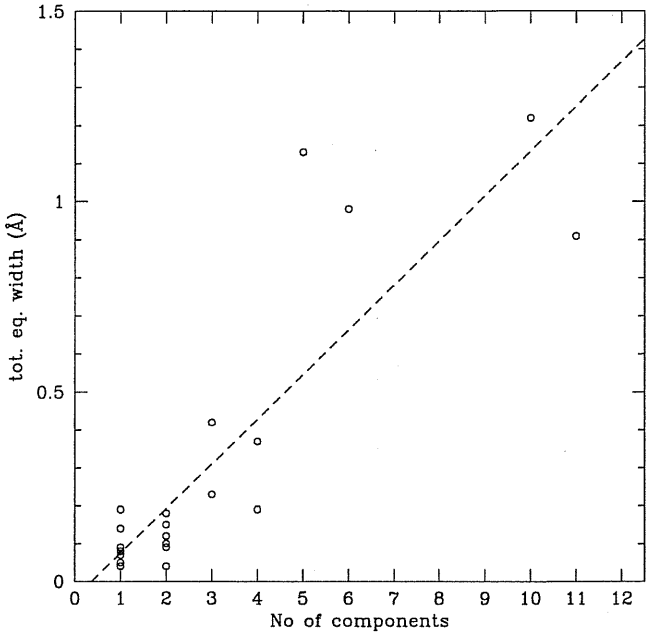

Figure 4.4: Total equivalent width of C IV systems versus the corresponding number of components. The dashed line is the linear correlation best fit for all points

In Fig. 4.4 the total equivalent width is plotted versus the number of components for the C IV absorption systems. The dashed line represents the best linear fit for all the points.

The number of components, however, is highly dependent on the spectral resolution and on subjective taste in the fitting procedure. To overcome this problem, we examined the maximum velocity separation in a given system versus the total equivalent width (Fig. 4.5). A correlation is apparent and the dashed line is, again,

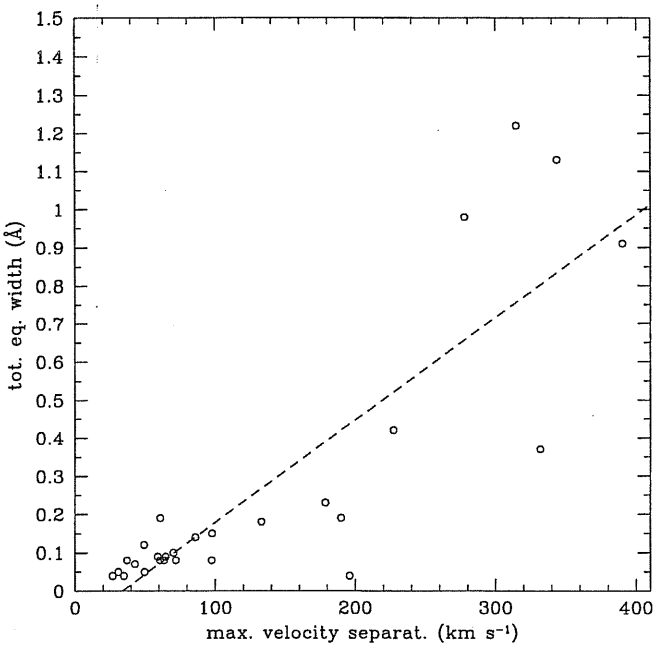

Figure 4.5: Total equivalent width of C IV systems versus the corresponding maximum velocity separation. The dashed line is the linear correlation best fit

the best linear fit.

The observed trends suggest that the clustering amplitude of the C IV lines could be a function of the column density.

In order to further investigate this issue, in Fig. 4.6 the amplitude of the TPCF (in the bin $30 < \Delta v < 90$ km s^{-1}) is plotted as a function of the median value of the C IV column density. The upper right point represents the result obtained in this work for the C IV absorptions in our sample. The lower left point comes from the lower column density sample obtained by Songaila and Cowie (1996) for the QSO 1422+231. The TPCF has been computed, also for the latter data, according to the procedure described previously. Lines of lower column density show indeed a smaller amplitude of the TPCF at $30 < \Delta v < 90$ km s^{-1}.

The correlation of the clustering amplitude with column density is analogous to what has been observed for Lyα lines (Cristiani et al. 1997; see Fig. 4.9) and it is qualitatively consistent with a picture of gravitationally induced correlations.

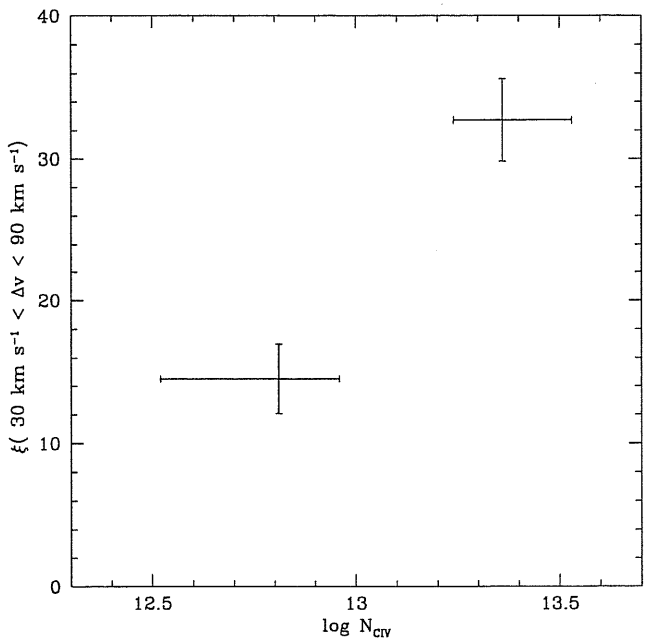

Figure 4.6: Amplitude of the clustering of C IV lines as a function of the column density. The upper right point has been obtained for our sample of C IV lines. The lower left point has been derived from the data of the QSO 1422+231 (Songaila & Cowie 1996). The vertical bars represent 1σ poissonian uncertainties in the determination of the TPCF while the horizontal ones show the 68% confidence interval of the column densities

In summary, the TPCF estimates obtained for heavy element systems along single lines of sight shows a strong signal for velocity splittings $\Delta v \lesssim 600$ km s^{-1} and a dependence of the clustering amplitude with the redshift (Quashnock & Vanden Berk 1998) and with the column density of the absorbers (D'Odorico et al. 1998a; Quashnock & Vanden Berk 1998). These evidences have convinced researchers that metal line systems are related to galaxies and/or to galaxy clusters, this has been proved directly in the case of Mg II absorptions (Bergeron & Boissé 1991; Steidel, Dickinson & Persson 1994; Churchill, Steidel & Vogt 1996).

4.3.2 The Estimate of the TPCF across Multiple Lines of Sight

The TPCF measured along the line of sight in the velocity space presents two major drawbacks: the measured signal can be due in part to peculiar velocities

within the absorbers rather than to their position in space and there is no straightforward way to compare the resulting one-dimensional correlation function with the three-dimensional correlation computed for low-z galaxy samples.

Indeed, as said above, many papers have modelled QSO absorption line clustering in terms of spatial separations indicated by their relative velocities, whereas for highly overdense structures, large differences between velocity clustering and spatial clustering become apparent (e.g. Kaiser 1987).

A possible way to overcome these difficulties is to study the clustering of absorbers in adjacent lines of sight that are close enough together so that the angular separation between them is less than or comparable to the velocity scales where clustering is seen or sought in single sightlines. Purely internal velocity splittings cannot shift an absorption in one QSO if the objects sits only in front of another, and even if splittings increase the number of cross-correlation pairs between two sightlines, they also increase the background density of pairs, so that both effects cancel in a properly normalized TPCF.

Furthermore, very high resolution spectra of close pairs or groups of QSOs allow to study clustering of absorbers in small redshift bins in order to have reliable information on the redshift evolution of the correlation properties of absorbers.

The correlation of absorption lines observed in paired lines of sight can also be interpreted as an indication of the size of the underlying absorbers, as discussed in the following chapter.

Multiple sightlines are also highly valuable in probing the large scale structure. In particular, multiple, well-sampled sightlines can provide a stronger test for voids by placing more absorbers in a void-sized volume than could possibly be obtained along a single sightline.

Williger et al. (1996) explored the spatial clustering of CIV absorbers between adjacent lines of sight for a sample of 25 QSOs within an ~ 1 deg^2 region of the South Galactic Pole having redshifts $1.5 \lesssim z_{\rm em} \lesssim 3.4$. The pair separation l_{12} for systems at redshifts z_1, z_2 and angular separation α are calculated by the law of cosines

$$l_{12} = \frac{(S_1^2 + S_2^2 - 2S_1S_2\cos\alpha)^{1/2}}{1+z}, \qquad (4.11)$$

where $z = (z_1 + z_2)/2$, and S is the proper distance given by Eq. (4.5).

They simulated a sinthetic data set extracting randomly the same number of lines, in the same redshift range as the observed ones from the expected C IV absorber redshift density function $d\mathcal{N}/dz = (d\mathcal{N}/dz)_0(1+z)^\gamma$ using the parameter found by Steidel (1990). The TPCF in three dimensions is computed using the estimator of Davis and Peebles (1983),

$$\xi(l) + 1 = \frac{N_c^2}{N_c(N_c-1)/2} \frac{DD(l)}{DR(l)}, \tag{4.12}$$

where $DD(l)$ is the number of observed pairs with separation l while $DR(l)$ refers to the number of cross-correlated observed-synthetic C IV pairs with separation l. The normalization is due to the fact that the total number of pairs in the synthetic-observed cross-correlation is larger than that in the observed-observed auto-correlation (N_c is the number of C IV absorptions).

They find evidence at the $3\,\sigma$ level for groups of C IV absorbers on the $5 - 10\ h^{-1}$ Mpc proper distance scale at $2.3 \lesssim z \lesssim 2.5$, separated by a $15 - 25\ h^{-1}$ proper Mpc gap, but because of the exiguity of the sample (they find 22 C IV at a resolution of 0.2 nm) they could not obtain a significant TPCF.

Crotts, Burles, and Tytler (1997) studied the cross-correlation properties of the C IV absorption systems in the QSO triplet 1623+27 ($z_{\rm em} \simeq 2.5$). At $z = 2.15$ the transverse separations between the three QSOs correspond to velocities in the Hubble flow of 286 to 399 km s^{-1}, and to proper separations of 0.51 to 0.71 h^{-1} Mpc.

Sightline cross-correlation for all C IV systems results in a TPCF $\xi(\Delta v < 500\ \text{km s}^{-1}) = 0.90$, a significantly weaker signal when compared with the clustering in velocity on corresponding scales along single line of sight.

The inconsistency of the TPCF implies that single sightline correlation functions cannot be used in a straightforward way to study the spatial clustering of absorbers on velocity scales of several hundred km s^{-1}. This may be due either to internal velocities within absorbers that are caused by nongravitational processes, or by motion within gravitational potentials that have separated from the Hubble flow. In either case, most of the line of sight correlational power is due to behavior not described by the absorber positions alone, but by some peculiar motion.

The results in this field are still preliminary: close QSO pair spectra observed at high redshift are needed in order to have larger samples. Furthermore, a direct comparison of the amplitude of the one-dimensional TPCF along the single lines of sight with that of the three-dimensional TPCF between different lines of sight, even if on the same velocity scale, could be misleading because the two functions are measuring different quantities and first a reliable relationship between them should be computed.

4.4 The Clustering Properties of Lyα Absorbers

4.4.1 The unidimensional TPCF of Lyα lines

The original work in this section has been published in Cristiani et al. (1997).

At a variance with what said for the metal line systems, the search for clustering of the Lyα lines has provided along the years results that have been defined disappointing or even controversial.

Systematic studies of the distribution of redshifts in the QSO Lyα forest began in the early 1980's with the work by Sargent et al. (1980). We already described the statistical tools introduced by Sargent and collaborators. They found that the TPCF of the Lyα absorption lines is flat to a level of $\sim$ 15 % on all scales from 300 to 30,000 km s^{-1}. On a finer scale the TPCF is observed to be flat down to the resolution limit 150 km s^{-1}.

Sargent, Young and Schneider (1982) considered the clustering of Lyα lines in a close pair of high redshift QSOs. The separation between the two lines of sight is 173″, that corresponds at a redshift $z = 2.5$ to a proper separation of $\simeq 670\ h^{-1}$ kpc. The autocorrelation function is seen to be flat on all velocity scales from 138 km s^{-1} to 12,500 km s^{-1}. They also measured the cross-correlation function between the two lines of sight , applying a variation of the response function (4.8) to normalize the number of oberved pairs. If $[S^A_{\rm min}, S^A_{\rm max}]$ and $[S^B_{\rm min}, S^B_{\rm max}]$ are the distance ranges in the two lines of sight, the number of pairs with separation s_0 is normalized by dividing for the function $\psi(s_0) = F(s_0)/F(0)$, where

$$F(s_0) = \mathrm{Min}[S^A_{\mathrm{max}}, S^B_{\mathrm{max}} + s_0] - \mathrm{Max}[S^A_{\mathrm{min}}, S^B_{\mathrm{min}} + s_0] + \mathrm{Min}[S^A_{\mathrm{max}} + s_0, S^B_{\mathrm{max}}] - \mathrm{Max}[S^A_{\mathrm{min}} + s_0, S^B_{\mathrm{min}}]. \quad (4.13)$$

The cross-correlation is also flat on the same scales.

Almost all the subsequent results have failed to detect any significant correlation on velocity scales $300 < \Delta v < 30000$ km s^{-1} (Bechtold 1987; Webb & Barcons 1991).

On smaller scales ($\Delta v = 50-300$ km s^{-1}) there have been indications of weak clustering (Webb 1987; Rauch et al. 1992; Chernomordik 1995; Cristiani et al. 1995), together with relevant non-detections (Pettini et al. 1990; Stengler-Larrea & Webb 1993).

In fact, the absence of power in the TPCF has been claimed as a striking characteristics of the Lyα forest and has been used as a basic argument to develop a scenario of the Lyα absorbers as a totally distinct population with respect to metal systems and therefore galaxies.

We investigated this issue taking advantage of a large dataset obtained in the framework of an ESO Key Programme devoted to the study of QSO absorption systems at high redshifts. The sample is formed by spectra of six QSOs, at resolutions between 9 and 14 km s^{-1}, with emission redshifts ranging from 3.27 to 4.12. The list of objects is given in Tab. 4.1 together with the main characteristics of the sample of Lyα absorption lines. This constitutes a unique database, especially at high redshift, where the density of lines provides particular sensitivity to any clustering signal.

We have complemented our data with other spectra available in the literature with similar resolution and redshift range, obtaining a final sample of 15 QSOs (in the following *extended sample*, see Tab. 4.1).

All the identified heavy element systems present in the Lyα forest have been removed from the extended sample, which consists of almost 1600 Lyα lines with observed column density $\log N_{\mathrm{HI}} \geq 13$.

Lines affected by the proximity effect, i.e. within a proper distance of 10 Mpc from the emission redshift of each QSO, have been excluded from the clustering

Table 4.1: QSO Spectra from the ESO KP

QSO Name	z_{em}	FWHM (km s^{-1})	Mag	No of lines	Sample Limit $\log(N_{\rm HI})$
$2126-15$	3.27	11	V=17.3	106	13.3
$2355+01$	3.39	9	V=17.5	76	13.3
$0055-26$	3.67	14	V=17.5	187	13.3
$1208+10$	3.82	9	V=17.5	66	13.3
$1108-07$	3.95	9	R=18.1	38	13.3
$0000-26$	4.12	12	V=17.5	181	13.3

Other QSO Spectra from the literature

QSO Name	z_{em}	FWHM (km s^{-1})	Mag (V)	No of lines	Sample Limit $\log(N_{\rm HI})$
$1331+17^a$	2.10	18	16.9	55	13.0
$1101-26^b$	2.15	9	16.0	37	13.3
$2206-19^c$	2.56	6	17.3	47	13.3
$1700+64^d$	2.72	15	16.1	73	13.3
$1946+76^e$	3.02	10	15.8	72	13.3
$0636+68^f$	3.17	8	16.6	172	13.0
$0302-00^f$	3.29	8	17.6	151	13.0
$0956+12^f$	3.30	8	17.6	151	13.0
$0014+81^f$	3.41	8	16.5	164	13.0

References:

[a] Kulkarni et al. 1996

[b] Carswell et al. 1991

[c] Rauch et al. 1993

[d] Rodríguez-Pascual et al. 1995

[e] Fan and Tytler 1994

[f] Hu et al. 1995

analysis.

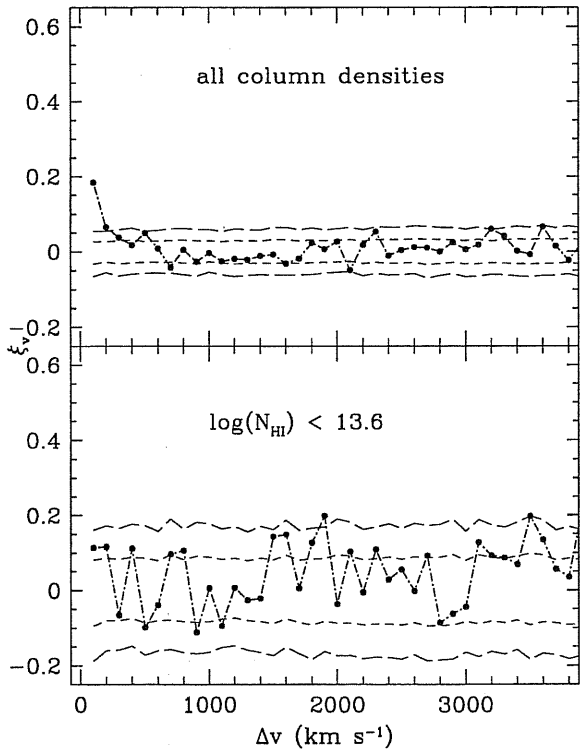

Figure 4.7: Two-point correlation function in the velocity space. Upper panel: for the complete sample of Lyα lines. Lower panel: for lines with column densities $< 10^{13.6}$ cm^{-2}. The short-dashed and long-dashed lines represent the 1σ and 2σ confidence limits for a Poissonian process

We adopted the TPCF estimator given by Eq. (4.9). In our line sample n_{exp} is obtained averaging 1000 numerical simulations of the observed number of redshifts, trying to account for all the relevant cosmological and observational effects. In particular the set of redshifts is randomly generated in the same redshift interval as the data according to the cosmological distribution $\propto (1+z)^{\gamma}$, where the best value of $\gamma = 2.65$ has been derived from a maximum likelihood analysis of the same database (Giallongo et al. 1996). The results are not sensitive to the value of γ adopted and even a flat distribution (i.e. $\gamma = 0$) gives values of ξ that differ typically by less than 0.02. Incomplete wavelength coverage due to gaps

in the spectrum or line blanketing of weak lines due to strong complexes is also accounted for. Lines with too small velocity splittings, compared with the finite resolution or the intrinsic blending due to the typical line widths - the so called "line-blanketing" effect (Giallongo et al. 1996) -, are excluded in the estimate of n_{exp}.

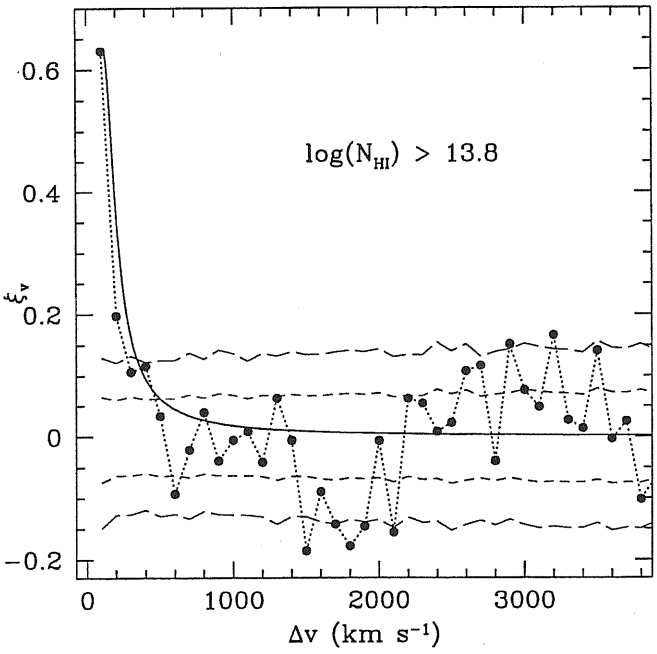

Figure 4.8: Two-point correlation function in the velocity space for lines with column densities $> 10^{13.8}$ cm^{-2}. Confidence limits as in Fig. 4.7. The continuous line shows the model of Heisler, Hogan and White, Eq. (4.14), with $\gamma = 1.77$, $\sigma = 50$ km s^{-1}, $r_{cl} = 100h^{-1}$ kpc and $r_0 = 130h^{-1}$ kpc at $z = 3$

The resulting correlation function for the full *extended sample* of Lyα lines is shown in the upper panel of Fig. 4.7. A weak but significant signal is present with $\xi \simeq 0.2$ in the 100 km s^{-1} bin: 739 pairs are observed while only 624 are expected for a random distribution, a 4.6 σ deviation from poissonianity.

As in Cristiani et al. (1995), we have explored the variations of the clustering as a function of the column density. In the bottom panel of Fig. 4.7 the correlation function for lines with $\log N_{\rm HI} \leq 13.6$ is shown. All the evidence for clustering has disappeared. On the contrary, for lines with $\log N_{\rm HI} \gtrsim 13.8$ (Fig. 4.8), the cor-

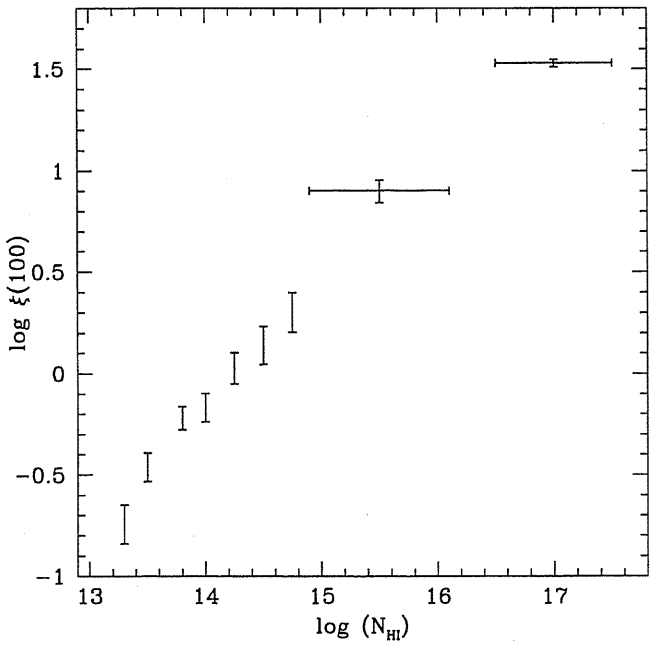

Figure 4.9: Evolution of the amplitude of the TPCF as a function of the column density threshold for the sample of the Lyα lines. The two points in the upper-right side of the picture show the correlation of the C IV metal systems (see text)

relation function at $\Delta v = 100$ km s^{-1} shows a remarkable increase in amplitude ($\xi \simeq 0.6$) and significance: 234 pairs are observed while only 145 are expected for a random distribution, a more than $7\,\sigma$ deviation from poissonianity. No relevant feature other than the peak at small velocity separations is observed. In particular the evidence for anti-clustering on scales $\sim 600 - 1000$ km s^{-1}, previously suggested by Meiksin and Bouchet (1995) on the basis of data of Q0055–259 and Q0014+813, is not confirmed. It should be noted that the assessment of the significance of a deviation from poissonianity is conceptually different when a velocity bin *given a priori* (as in the analysis of small-scale clustering) is analysed, with respect to the case in which a deviation in *any* of the bins is searched for. While in the former situation the confidence levels shown in Figg. 4.7, 4.8, 4.10 (obtained by comparing the number of pairs observed in each bin with the distributions expected from simulations) are perfectly valid, in the latter case the significance of a positive or negative deviation in *any* of the bins is considerably less than what a

naive comparison would at first glance suggest.

Figure 4.9 shows the variation of the amplitude of the TPCF as a function of the column density threshold for the sample of the Lyα lines. On the same plot are reported for comparison the TPCF's derived for high and low column density C IV metal systems from the works by Petitjean and Bergeron (1994) and by Songaila and Cowie (1996), respectively. A reference value of $\log N_{\rm HI} \sim 17 \pm 0.5$ has been assigned to high column density CIV metal systems on the basis of observations of 0420-388 (Atwood, Baldwin & Carswell 1985; Petitjean & Bergeron 1994). For the low column density C IV systems a value of $\log N_{\rm HI} \sim 15.5 \pm 0.6$ is derived from Songaila and Cowie (1996).

An extrapolation of the increasing amplitude trend observed for the TPCF of the Lyα lines would easily intercept the corresponding estimates derived from the CIV metal systems. We stress also the similarity between the shapes of the TPCF's of Lyα and C IV systems, when observed at comparable resolution (see Fig. 4.3) and the fact that a trend of increasing correlation with increasing column density is observed also in C IV systems (see e.g. D'Odorico et al. 1998a, Quashnock & Vanden Berk 1998), for which the strong systems appear usually in clumps of many components, while the weak doublets (generally optically thin at the Lyman limit) are typically made of an isolated, single component (Bergeron 1995).

Fernandez-Soto et al. (1996) have investigated the clustering properties of the Lyα clouds on the basis of the corresponding CIV absorptions, suggesting that CIV may resolve better the small-scale velocity structure that cannot be fully traced by Lyα lines. As a consequence, the estimates of the TPCF of the Lyα absorbers, although the effects of the non-negligible width of the lines (the "line-blanketing") are considered in our simulations, should be regarded as a lower limit to the real clustering amplitude. However, it is not straightforward to translate the properties of the C IV absorbers in the corresponding ones of the Lyα absorbers, since observations (e.g. Prochaska & Wolfe 1996) show that the velocity structures of high and low-ionisation species are often different. Since the underestimation of the Lyα TPCF would be more severe at larger column densities, the trend of an increasing correlation with increasing column density appears to be real, from the lowest column densities up to those corresponding to the strongest metal systems.

In Fig. 4.10 we study the evolution of the TPCF with the redshift for the sub-

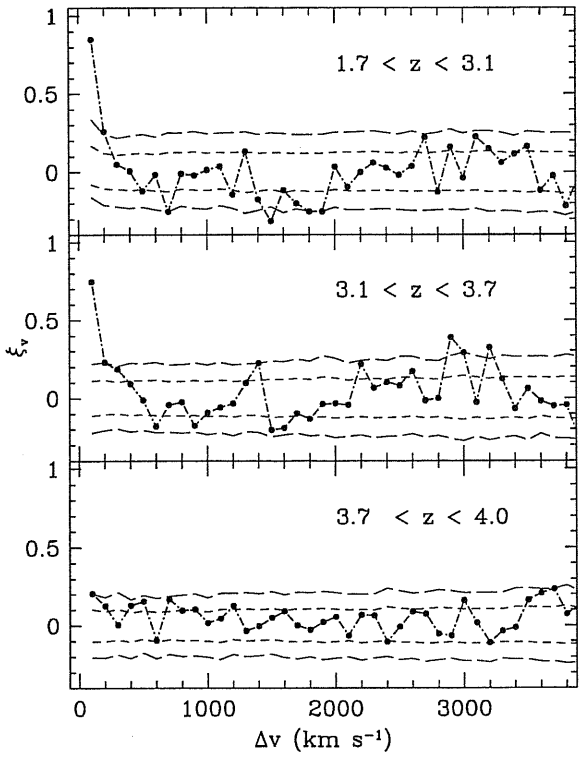

Figure 4.10: Evolution of the two-point correlation function with redshift for the sample of the Lyα lines with column densities above $\log N_{\rm HI} = 13.8$. The three redshift intervals $1.7 < z < 3.1$, $3.1 < z < 3.7$, $3.7 < z < 4.0$ (upper, middle and lower panel, respectively) have been chosen in a way that the confidence limits, estimated as in Figg. 4.7-4.8, turn out to be roughly the same in all three ranges

sample of Lyα lines with column densities $\log N_{\rm HI} > 13.8$. The amplitude of the correlation at 100 km s^{-1} decreases with increasing redshift from 0.85 ± 0.14 at $1.7 < z < 3.1$, to 0.74 ± 0.14 at $3.1 < z < 3.7$ and 0.21 ± 0.14 at $3.7 < z < 4.0$. This trend persists at lower redshifts, as confirmed by the measure of the Lyα TPCF by Ulmer (1996), $\xi = 1.8^{+1.6}_{-1.2}$, on scales of $200-500$ km s^{-1} at $0 < z < 1.3$ and by the results on the Lyα forest lines in the spectrum of the Hubble Deep Field South QSO J2233-606 for which $\xi_{100} \simeq 1.3$ at $< z >= 1.7$ (Savaglio et al. 1999).

The observed clustering properties are qualitatively consistent with a scenario of gravitationally induced correlations. The dependence of the clustering ampli-

tude on the column density, is easily explained by models involving biasing for the formation of structures in the universe. Objects associated with the stronger potential wells are expected to be more clustered and the H I column density is naturally related to the depth of the well (and to the associated mass).

The trend of increasing correlation with decreasing redshift (Fig. 4.10) is also a strong prediction of any model of structure formation based on gravitational instability. On the contrary, theories of explosive structure formation (Vishniac, Ostriker & Bertschinger 1985; Weinberg, Ostriker & Dekel 1989) expect a velocity correlation either unchanging or diminishing with time.

Numerical simulations in the framework of a CDM model, including photoionisation and cooling of the baryonic component, have reproduced to a remarkable extent the observed correlation function and also its dependence on the column density of the lines (Mücket et al. 1996).

In this way the observed clustering properties (in particular the correlation between clustering amplitude and column density) become part of a new continuity scenario that rests on two other pieces of evidence: the metallicities of the order 10^{-2} observed for Lyα clouds with $\log N_{\rm HI} > 14$ (Cowie et al. 1995; Tytler et al. 1995) and the volume density and cosmological evolution of the same clouds found to be similar to those of the damped systems (Giallongo et al. 1996). All together this suggests a physical association between the Lyα clouds with $\log N_{\rm HI} > 14$ and the halos of protogalactic systems. At low redshift a considerable fraction of the Lyα lines has been observed to be associated with luminous galaxies and the local large-scale structure (Lanzetta et al. 1995; Le Brun, Bergeron & Boissé 1996).

The typical column density below which no clustering of the Lyα lines is observed ($\log N_{\rm HI} \sim 13.6$, Fig. 4.7) corresponds to the position of the break in their column density distribution (Giallongo et al. 1996; Hu et al. 1995), which has been identified with the transition from a variety of systems in various stages of gravitational infall and collapse (or even under-densities) to gas associated with star forming galaxies (Zhang, Anninos & Norman 1996; Mücket et al. 1996).

Heisler, Hogan and White (1989) presented a relationship to derive the three-dimensional spatial autocorrelation function of the absorbers ξ_r, from the velocity correlation ξ_v, assuming a Gaussian distribution of the peculiar motions of the

absorbers with respect to the Hubble flow and a power-law form for the spatial correlation function, as observed for galaxies.

$$\xi_v = \int_0^\infty H dr\, \xi(r)\, P(v \mid r) \propto \int_{r_{cl}}^\infty \frac{H dr}{\sigma_v} \left(\frac{r}{r_o}\right)^{-\gamma} \times \left\{ \exp\left[-\frac{(Hr - v)^2}{2\sigma_v^2}\right] + \exp\left[-\frac{(Hr + v)^2}{2\sigma_v^2}\right] \right\} \quad (4.14)$$

At small velocity splittings, the correlation scale r_o depends mainly on the cloud sizes r_{cl} and on the velocity dispersion σ_v assumed. The very fact that a significant correlation in velocity is observed results in an upper limit on the cloud proper sizes of the order of 100 h^{-1} kpc. We will see in the next chapter that large sizes for the Lyα absorbers are obtained from the analysis of QSO pairs spectra (e.g. D'Odorico et al. 1998b; Crotts & Fang 1998). Furthermore, the trend observed in Fig. 4.10 is consistent, in the framework of a simple hierarchical clustering model predicting that r_o scales with the redshift according to the relation $r_o(z) = r_o(0)(1+z)^{-5/3}$, with large cloud sizes ($r_{cl} \lesssim 100\ h^{-1}$ kpc) and low velocity dispersions (few tens km s^{-1}).

Following Cristiani et al. (1995), we assume an index $\gamma = 1.77$ for the power-law spatial correlation function as commonly measured for galaxies, relatively low velocity dispersions (< 100 km s^{-1}) as suggested by observations of quasar pairs (e.g. Dinshaw et al. 1994) and cloud sizes of 100 h^{-1} kpc, this implies, to reproduce the observed correlation in the velocity space, a correlation scale at $z \sim 3$ of $100 - 150\ h^{-1}$ comoving kpc. This amplitude extrapolates, according to the $(1+z)^{-5/3}$ scaling, to $1 - 1.5\ h^{-1}$ Mpc at the present epoch, to be compared with the values 5 and $3.5h^{-1}$ Mpc observed for present day optical (de Lapparent, Geller & Huchra 1988) and IRAS (Fisher et al. 1994) galaxies, respectively.

A spatial correlation function with a correlation scale $r_o(z=0) \sim 1.5$ Mpc, convolved with a velocity dispersion $\sigma_v \sim 50$ km s^{-1}, would result in a correlation function in the velocity space non dissimilar with respect to what is observed for C IV systems (Womble, Sargent & Lyons 1996; Songaila & Cowie 1996) if a cloud radius $\sim 7.5\ h^{-1}$ kpc (proper) is assumed. Such a value appears considerably smaller than what is commonly inferred for galactic halos giving origin to metal-rich absorption systems ($r \sim 40h^{-1}$ kpc, Bergeron 1995), but certainly at

least part of the small-scale correlation observed for C IV systems is due to the motion of clouds of size $\ll 40$ kpc within individual galaxy halos. The observed correlations in the velocity space might then arise from the same spatial correlation function both for Lyα and metal systems, with different degrees of amplitude suppression due to the different sizes of the regions giving origin to the Lyα and metal absorption, a well-known phenomenon occurring when the cloud sizes are non-negligible with respect to the correlation scale (Bajtlik 1995).

4.4.2 The Cross-Correlation of Lyα Absorbers

In § 4.3.2, we discussed the drawbacks of the TPCF computed along the line of sight.

Sargent, Young and Schneider (1982), as already said, found no significant correlation between the Lyα absorption lines in a QSO pair of separation $\sim 173''$. Shaver and Robertson (1983) obtained spectra of the QSO pair Q0307-195A,B (separated by $58''$): they found a marginal tendency among the Lyα lines in the two spectra to correlate with each other at small velocity separations ($\lesssim 500$ km s^{-1}).

Crotts (1989) considered three high redshift QSOs ($< z_{\rm em} >= 2.5$) with mutual angular separations $127''$, $147''$ and $177''$. A cross-correlation of the Lyα forest samples in each of the three QSOs with samples from the other two reveals that Lyα absorbers with equivalent width $w_r > 0.04$ nm cluster in space on scales less than $\sim 700\ h^{-1}$ kpc for velocity splittings $\Delta v < 100$ km s^{-1}. Lyα absorbers with $w_r > 0.08$ nm may cluster up to $\Delta v = 400$ km s^{-1}. Neither clustering signal is seen for lines with 0.02 nm $< w_r <$ 0.04 nm.

In Crotts and Fang (1998), they investigated the same issue on new, higher resolution spectra of the same triplet. They again found a clustering signal for the Lyα sample with $w_r > 0.04$ nm and $< z >= 2.25$ for velocity separations $\Delta v < 50$ km s^{-1}, significative at the $\sim 2.5\,\sigma$ level. There is a weak signal for the sample with $w_r > 0.08$ nm on scales $\Delta v < 600$ km s^{-1}.

4.4.3 Lyα Lines Tracing the Large Scale Structure of the Universe

The TPCF is known to be inefficient for detecting structures on scales comparable with the dimension of the sample (e.g. Pando & Fang 1996). For this reason to investigate the clustering properties of the Lyα absorbers on scales of several comoving megaparsecs, searching for over- and under-densities of lines significantly deviating from a Poissonian distribution, the void probability function and, more generally, neighbor statistics (Ostriker, Bajtlik & Duncan 1988; Liu & Jones 1990; Fang 1991; Meiksin & Bouchet 1995) have been used, invariably revealing the non-Poissonian (clustered) nature of the distribution of clouds in velocity space (Ostriker, Bajtlik & Duncan 1988; Bi, Börner & Chu 1989; Liu & Jones 1990; Babul 1991).

Structure has been detected on many different scales: Fang (1991) used a Kolmogorov-Smirnov test for nearest neighbor intervals to detect a signal on scales $30 - 50\ h^{-1}$ Mpc. Mo et al. (1992), from an analysis of the extrema in the slope of the TPCF saw structure at 60 and 130 h^{-1} Mpc. Meiksin & Bouchet (1995) found an anti-correlation in the TPCF around $3 - 6\ h^{-1}$ Mpc. Pando & Fang (1996), applying the wavelet transform, found clusters $\sim 20\ h^{-1}$ Mpc in size in the Lyα forest. The physical interpretation of the various results is not entirely obvious. There are intrinsic scales in the spectra (like the quasi-periodic change in S/N ratio caused by the sensitivity maxima of the orders in an echelle spectrum) which are of similar magnitude ($\sim$ 5000 km s^{-1}, or $\sim 25\ h^{-1}$ Mpc comoving at $z \sim 3$) as some of the above detections.

Voids in the Lyα forest provide a test for models of the large-scale structure in the Universe. However, lack of Lyα forest absorption lines may not necessarily mean that a particular volume of space contains no hydrogen. We observe the clouds only through Lyα absorption, i.e. the presence of H I. If the clouds in one area were for any reason more ionised than average, more of the hydrogen would be H II, and we could observe a void there. Thus, the absence of voids can also be considered as a test of the homogeneity of the UV ionising flux.

The probability function for a Poissonian gap of size Δz in a spectrum with absorption line density $d\mathcal{N}/dz$ is simply

$$P(\Delta z) = \exp\left[-\left(\frac{d\mathcal{N}}{dz}\right)\Delta z\right]. \tag{4.15}$$

Carswell and Rees (1987) concluded that voids with sizes like those in the local universe ($\sim 50\ h^{-1}$ Mpc comoving) cannot fill more than 5 % of the volume at $< z >\sim 3.2$. This result was confirmed by work by Duncan, Ostriker and Bajtlik (1989), based on a larger dataset. A similar conclusion was reached by Pierre, Shaver and Iovino (1988), who found that the Lyα absorbing gas cannot exhibit a void structure that is similar to that of low redshift galaxies, without producing strong clustering inconsistent with the observations.

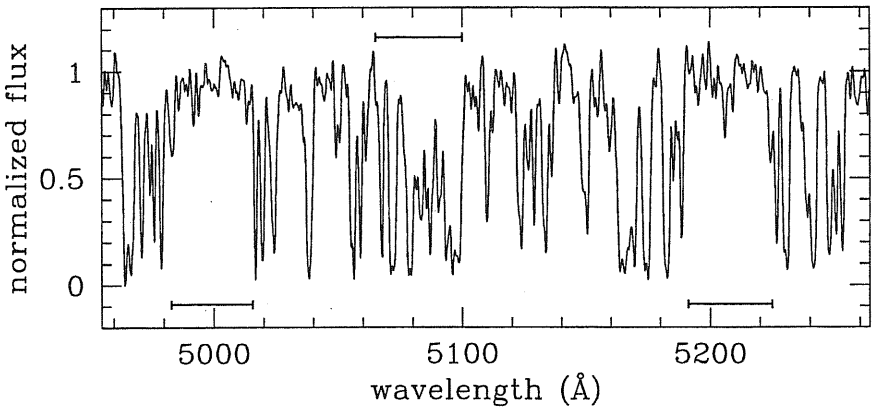

Figure 4.11: Over and under-densities of lines at a scale of 20 Mpc in the spectrum of Q0055 − 26. The positions of two voids, at 500 and 520 nm respectively (see Tab. 4.2), are shown together with and over-density at 508 nm. The relatively low resolution of the plot does not allow to fully visualize the significance of the over-density: 15 lines with $\log N_{\rm HI} \geq 13.3$ are observed while only 7 are expected

Nevertheless, individual large gaps have been found. Crotts (1987), discovered an 43 h^{-1} Mpc gap towards Q0420-388. This result was variously contested and confirmed in a dispute about significance levels (Ostriker, Bajtlik & Duncan 1988; Crotts 1989; Duncan, Ostriker & Bajtlik 1989; Bi, Börner & Chu 1991, Rauch et al. 1992), fuelled, among other things by the lack of a universally adopted definition of the term "void" which takes into account that a void may be void of lines only down to a certain detection threshold. It turned out that Crotts'

gap, though not entirely empty, is a region of significantly lower line density. Dobrzycki and Bechtold (1991b) found another void of size $\sim$ 32 h^{-1} Mpc in the spectrum of Q0302–003, and Cristiani et al. (1995) discovered a significant pair of smaller voids towards Q0055–269.

High resolution data, less affected by blending effects, are ideal also for the study of voids, but great care has to be taken in the statistical approach to avoid the pitfalls of "a posteriori statistics". Detections of voids in individual cases are interesting but a more general approach, assessing more quantitatively how common is the phenomenon, is to be preferred. Differences in the redshifts and S/N (and consequently in the line density) can make difficult a global and uniform usage of extended samples of data.

We have designed the following procedure for the extended sample described in the previous section: for each of the objects listed in Tab. 4.1 and for lines with $\log N_{\rm HI} \geq 13.3$ we have estimated through Monte Carlo simulations the void size Δr (in comoving coordinates) for which the probability to find at least one void $\geq \Delta r$ is 0.05. The objects for which the density of lines and the absence of spectral gaps allow a meaningful analysis are listed in Tab. 4.2. Then we have searched each spectrum for voids of dimensions equal or larger than Δr. In 3 out of 9 cases at least one void was detected (two voids for $0055-26$). The binomial probability corresponding to such occurrences is $8 \cdot 10^{-3}$.

The four voids reported in Tab. 4.2 have been found in the spectra of Q0636+68 at $434-436.2$ nm (16.1 Mpc size), Q0055–26 at $498-501.7$ and $518.8-522.4$ nm (22.3 and 19.8 Mpc, respectively) and Q0000–26 at $562.7-564.4$ nm (8.5 Mpc). The well known void at $505-510$ nm in the spectrum of Q0302–00 (Dobrzycki & Bechtold 1991b) has not been found because the spectrum of Hu et al. (1995) does not cover this region of the Lyα forest.

From the published spectra and from Fig. 4.11 it is apparent that the regions corresponding to the voids are not completely devoid of lines: weak absorptions are observed within the voids. This agrees with low-redshift observations (Shull, Stocke & Penton 1996) that have shown that in the local Universe voids are not entirely devoid of matter.

Even if underdense regions are statistically significant in our sample, the filling

Table 4.2: QSO Spectra used in the search for voids

QSO Name	z_{em}	minimum void size Δr (Mpc)	Number of detected voids
$0636+68$	3.17	15.9	1
$2126-15$	3.27	24.7	0
$0302-00$	3.29	17.5	0
$0956+12$	3.30	16.9	0
$2355+01$	3.39	13.1	0
$0014+81$	3.41	15.8	0
$0055-26$	3.67	14.9	2
$1208+10$	3.82	8.5	0
$0000-26$	4.12	6.9	1

factor is rather low: the voids reported in Tab. 4.2 cover only about 2% of the available line-of-sight path-length, confirming what already noticed by Carswell and Rees (1987).

Over- and Under-Densities of Lines

Voids are just an extreme case of spectral regions showing an under-density of lines. Various theoretical reasons prompt to tackle the issue of over- and under-densities of lines from a more general point of view: the typical signature of a "proximity effect" due to a foreground quasar is a lack of weak lines in the Lyman forest, rather than lack of lines in general (Dobrzycky & Bechtold 1991a); the relative filling factor of under-densities and over-densities may provide a constraint on the theories of structure formation.

We have analysed the spectra of the quasars reported in Tab. 4.1 with a counts-in-cells technique searching for over- and under-densities of lines with $\log N_{\rm HI} \geq 13.3$ on scales from 10 to 80 Mpc and comparing the observed counts with Monte Carlo simulations in order to assess the significance of the deviations. On smaller

Table 4.3: Over-densities of lines found in QSO spectra

Scale	10 Mpc	20 Mpc	30 Mpc	40 Mpc	80 Mpc
QSO Name	over-density center (nm)				
$1331+17$	349.4	349.3	349.0	349.0	-
	360.5	360.9			
$1101-26$	374.2	-	374.3	-	-
$2206-19$	-	-	-	-	-
$1700+64$	-	-	-	-	-
$1946+76$	442.7	443.5	-	-	-
$0636+68$	-	-	-	-	-
$2126-15$	-	-	-	-	-
$0302-00$	-	-	469.6	-	-
$0956+12$	-	-	-	-	-
$2355+01$	-	521.7	-	-	-
$0014+81$	-	-	-	-	-
$0055-26$	508.5	508.0	-	-	-
$1208+10$	562.1	-	564.5	-	*
$1108-07$	-	-	-	583.9	*
$0000-26$	-	-	-	-	-

scales the shot noise is too large, on larger scales the "integral constraint", forcing the simulated number of lines to be equal to the observed one, prevents from the possibility of detecting any deviation. The threshold to define significant a deviation (for example in excess) in a given spectral interval has been set in a way that at a given scale for a given quasar there is a 0.05 probability of observing at least one deviation of this type on the whole spectrum for a locally Poissonian distribution. In Tab. 4.3 we report the results of the analysis. An asterisk denotes that a meaningful search at that particular scale has not been possible for the given object.

Table 4.4: Under-densities of lines found in QSO spectra

Scale	10 Mpc	20 Mpc	30 Mpc	40 Mpc	80 Mpc
QSO Name	under-density center (nm)				
$1331+17$	*	*	370.7	-	-
$1101-26$	*	*	367.9	-	367.0
$2206-19$	*	423.8	-	-	-
$1700+64$	*	-	-	-	-
$1946+76$	*	-	-	479.9	-
$0636+68$	*	-	-	-	-
$2126-15$	*	-	-	-	-
$0302-00$	*	-	-	-	-
$0956+12$	*	-	-	-	-
$2355+01$	*	-	-	-	-
$0014+81$	*	-	-	-	-
$0055-26$	526.3	500.1	500.7	498.8	-
		520.8	540.6		
$1208+10$	-	-	-	-	-
$1108-07$	-	-	-	-	*
$0000-26$	563.7	-	-	-	-

5 QSOs out of 15 show at least one over-density in their spectrum at 10 Mpc scales, corresponding to a binomial probability of $6 \cdot 10^{-4}$ of being drawn from a poissonian distribution of lines. 4 QSOs show at least one over-density of lines at 20 and 30 Mpc scales, corresponding to a binomial probability of $5 \cdot 10^{-3}$. At larger scales the number of significant over-densities decreases. None is observed at 80 Mpc. Over-densities detected at one scale tend to persist at larger scales, analogously to what observed in the wavelet identification of structures from galaxy counts (Slezak, Bijaoui & Mars 1990).

Under-densities appear to be roughly as common as over-densities, once the lack of sensitivity at lower redshifts and smaller scales, due to the low density of lines, is taken into account. In practice in Tab. 4.4 the voids of Tab. 4.2 are recovered, except for the void of Q0636+68, the least significant case.

The existence of a roughly equal number of over and under-densities on scales $10 - 40$ Mpc is easily understandable in terms of the linear theory of the evolution of the perturbations, that is a plausible approximation at such relatively high redshifts: gravity has not yet had time to give a significant skewness to the (under)over-density distribution. Besides, almost any hierarchical clustering scenario would expect that at $z \sim 2-4$ gravity has not yet had enough time to transfer power on 80 Mpc scales and give origin to significant over or under-densities.

4.5 Recovering the Mass Power Spectrum from the Lyα Forest

The analysis of clustering of QSO absorption lines by means of the TPCF was aimed to the comparison of the results with the analogous measure for the local galaxies. If, more in general, we want to use QSO absorption spectra to test the predictions of theoretical models of cosmic structure formation, we have to take into account the selection function and the biases that allow to relate the TPCF of the discrete population of absorbers to that of the underlying mass distribution, or to its Fourier transform, the power spectrum $P(k)$.

Since now, most efforts to measure $P(k)$ have focused on galaxy redshift surveys. This route to the primordial power spectrum has several obstacles, includ-

ing shot noise stemming from the discrete nature of the galaxy distribution and nonlinear gravitational evolution of $P(k)$, which is important over much of the accessible range of scales. The most difficult problem to overcome is the uncertain relation between the galaxy and mass distributions, usually parametrized in terms of a (possibly scale dependent) "bias factor" between the galaxy and mass power spectra. Furthermore, galaxy redshift surveys primarily probe structure at a single epoch, redshift zero.

As I will describe in § 6.2, simulations sustained a relatively new paradigm for the Lyα forest already investigated using analytical calculations (e.g. Bond, Szalay & Silk 1988; Mc Gill 1990; Bi, Börner & Chu 1992). The forest is now generally believed to be due to absorption by large-scale neutral hydrogen density fluctuations of moderate amplitude ($\delta \lesssim 10$) in a warm photoionised intergalactic medium (IGM). Important results of these simulations are a tight correlation between the H I and the dark matter distribution (on scales larger than the Jeans length of the IGM), $\rho_b \simeq \rho$, and a simple temperature-density relation for the gas (e.g. Hui & Gnedin 1997; Haehnelt & Steinmetz 1998), $T = T_0 \, (\rho_b/\bar{\rho}_b)^{\alpha}$. The parameters T_0 and α depend on the reionisation history of the universe and on the spectral shape of the UV background. They are expected to lie in the ranges $4000 \lesssim T_0 \lesssim 15,000$ K and $0.3 \lesssim \alpha \lesssim 0.6$ (Hui & Gnedin 1997). For most of the gas in this density regime, the Lyα optical depth is proportional to the neutral hydrogen density (Gunn & Peterson 1965), which since the gas is in photoionisation equilibrium is proportional to the density times the recombination rate:

$$\tau \propto \rho_b^2 \, T^{-0.7} = \mathcal{A} \, (\rho_b/\bar{\rho}_b)^{\beta}, \tag{4.16}$$

$$\mathcal{A} = 0.433 \left(\frac{1+z}{3.5}\right)^6 \left(\frac{\Omega_b h^2}{0.02}\right)^2 \left(\frac{T_0}{6000\mathrm{K}}\right)^{-0.7} \times \left(\frac{h}{0.65}\right)^{-1} \left(\frac{H(z)/H_0}{3.68}\right)^{-1} \left(\frac{\Gamma}{1.5 \times 10^{-12}\mathrm{s}^{-1}}\right)^{-1}, \tag{4.17}$$

with $\beta \equiv 2 - 0.7\alpha$, varying in the range $1.6 - 1.8$. Γ is the H I photoionisation rate. Because Eq. (4.16) describes the analog of Gunn-Peterson absorption for a non-uniform, photoionised medium (ignoring the effect of peculiar velocities), it

has been called the Fluctuating Gunn-Peterson Approximation (FGPA, see Rauch et al. 1997; Weinberg et al. 1998; § 6.3).

The transmitted flux in a QSO spectrum, $F = e^{-\tau}$, is monotonically related to ρ in this approximation. Because the relation between τ and ρ is fairly simple, one can extract information about the underlying mass density field from the observed flux distribution.

One approach would be to invert the above relations, deriving τ from F and ρ from τ. However, the mapping between ρ and F is nonlinear (en exponential of a power law), and it depends on the unknown constant of proportionality in the $\tau - \rho$ relation. Also, it is difficult to measure τ accurately in saturated regions, where $F \simeq 0$.

Croft et al. (1998) proposed a method to recover the density field along the line of sight using "Gaussianization", a technique introduced by Weinberg (1992) as a tool for recovering primordial density fluctuations from the observed galaxy distribution. In the hypothesis that the primordial density field has a Gaussian probability distribution function (PDF), the flux in a QSO spectrum is monotonically mapped back to a Gaussian density field, by sorting the pixels so that they have the same rank order but a Gaussian PDF.

Then the one-dimensional power spectrum of the approximated initial density field along the line of sight, $P_{1D}(k) =< \delta^2(k) >$, can be estimated. $\delta^2(k)$ is computed using a fast Fourier transform and is averaged over bins in k.

In general, the three-dimensional power spectrum of some random field is related to its one-dimensional projection through the following integral (Kaiser & Peacock 1991):

$$P_{1D}(k) = \frac{1}{2\pi} \int_k^\infty P(y)\, y\, dy. \tag{4.18}$$

In order to recover $P(k)$ from $P_{1D}(k)$, one must invert Eq. (4.18),

$$P(k) = -\frac{2\pi}{k} \frac{d}{dk} P_{1D}(k) \tag{4.19}$$

To normalize the spectrum, Croft and collaborators used the derived $P(k)$ to set up initial conditions for particle-mesh N-body simulations. They evolved these simulations to the observed redshift and required them to reproduce the

three-dimensional power spectrum of the (non-Gaussianized) flux on large scales. The value of the uncertain parameter $\mathcal{A}$ is determined in the normalizing simulations by matching an independent observation, the effective mean optical depth $\bar{\tau}_{\rm eff} \equiv -\ln < e^{-\tau} >$.This observational determination of $\mathcal{A}$ removes any dependence of the derived $P(k)$ on unknown "bias factors".

The test of the method over artificial QSO spectra created from realistic hydrodynamic cosmological simulations shows that it successfully recovers both the shape and the amplitude of the mass power spectrum on scales from 2 h^{-1} to 12 h^{-1} Mpc. The main uncertainties are due to continuum fitting, that limits the largest scale at which $P(k)$ can be measured to $k \sim 2{\times}10^{-3}$ (km s^{-1})$^{-1}$, the possible inhomogeneity of the UV background intensity (and hence the photoionisation rate) and to the assumption that primordial fluctuations are Gaussians. While the errors that affect the determination of $P(k)$ are mainly "cosmic variance" errors that can be limited taking as many independent sightlines as possible.

An application of this method to 19 QSO spectra (Croft et al. 1999) indicates that the slope and amplitude of the derived $P(k)$ are consistent with the predictions of some scale-invariant, COBE-normalized CDM models (e.g. an open model with $\Omega_0 = 0.4$, $h = 0.65$, $\sigma_8 = 0.75$) and inconsistent with others (the COBE CDM model, with $\Omega_0 = 1$, $h = 0.5$, $\sigma_8 = 1.2$). A comparison of $P(k)$ to the measured clustering of Lyman Break Galaxies implies that the latter are a highly biased population, with bias factor $b \sim 2 - 5$.

Equation (4.19) assumes that the distribution of matter is isotropic with respect to the line of sight. Redshift-space distortions mean that this is not strictly true (Kaiser 1987), and these distortions must be taken into account for a truly accurate inversion of one-dimensional clustering.

Hui (1999) considered that redshift distortions imply that P is in general a function of the wavevector along the line of sight, $k_{||}$, as well as of the magnitude of the three-dimensional wavevector, k, in which case differentiation of P_{1D} alone would not recover the true shape of the three-dimensional real space mass power spectrum. Equation (4.18) becomes

$$P_{1D}(k_{||}) = \frac{1}{2\pi} \int_{k_{||}}^{\infty} P(k_{||}, k)\, k\, dk, \qquad (4.20)$$

and the effect of redshift space distortions on the power spectrum, at both small and large scales, can be described by

$$P(k_{||}, k) = W(k_{||}/k, k)\, P(k) \tag{4.21}$$

where $P(k)$ is the isotropic power spectrum in the absence of peculiar motion, and W is a suitable distortion kernel. Finally, the one-dimensional redshift-space transmission power spectrum can also be related to the three-dimensional real-space mass power spectrum by an effective kernel, $W^{F\rho}$:

$$P^F_{1D}(k_{||}) = \frac{1}{2\pi} \int_{k_{||}}^{\infty} W^{F\rho}(k_{||}/k, k)\, P^{\rho}(k)\, k\, dk. \tag{4.22}$$

The special case considered by Croft et al. (1998) corresponds to $W^{F\rho}$ =const.. Hui finds that in general a simple differentiation method tends to make the inverted three-dimensional real-space power spectrum flatter than it really is on large scales. He carried out a linear perturbative calculation (i.e. assuming $\delta_F \ll 1$) of W,

$$W(k_{||}/k, k) = [1 + \beta_f k^2_{||}/k^2 + \Delta_b]^2 \exp[-k^2_{||}/k^{s2}_{||}]; \tag{4.23}$$

$$\beta_f = \frac{f_\Omega}{2 - 0.7\alpha}, \quad \Delta_b = \frac{\alpha}{2.8\alpha - 8} k^2_{||} b^2_{\rm th} \tag{4.24}$$

where $f_\Omega = d\ln D/d\ln a$, with D being the linear growth factor and a the expansion scale factor, $b_{\rm th}$ is the thermal broadening width, and α is the exponential in the temperature-density relation. The smoothing scale $k^s_{||}$ is associated to the thermal broadening scale as well as spectral resolution.

A promising application of this machinery is a version of the Alcock and Paczyński test (1979, AP) proposed by Hui, Stebbins and Burles (1999). The idea of the AP test is that an object placed at cosmological distance have a definite relationship between its angular and redshift extents, which is cosmology-dependent. This relation is particularly sensitive to the presence of a component of the cosmological energy contents which has negative pressure, let us call it Q. In the case of the Lyα forest, the "object" to use is the TPCF or its Fourier counterpart: the one-dimensional correlation along the line of sight is compared to the cross-correlation

between two close-by lines of sight. More precisely, given the observed transmission power spectrum, one can invert Eq. (4.22) to obtain the mass power spectrum, and then use the following equation to predict the cross-spectrum P_X^F between two lines of sight with separation θ

$$P_X^F(k_{||},\theta) = \frac{1}{2\pi}\int_{k_{||}}^{\infty} W(k_{||}/k,k)\, P^\rho(k)\, J_0[ku(\theta)]k\, dk \qquad (4.25)$$

where J_0 is the spherical Bessel function and $u(\theta)$ is the transverse velocity separation for the given θ. This prediction is cosmological dependent mainly through the parameter $u(\theta)$ and also through the distortion kernel $W(k_{||}/k,k)$. A comparison of the predicted with the observed cross spectra, provides a version of the AP test: assuming the wrong cosmology will result in a wrong prediction for the cross spectrum.

Also McDonald and Miralda-Escudé (1999) pointed out the effects of peculiar velocities in the recovery of the mass power spectrum and applied the technique to an AP test.

Both groups of research showed that at least ~ 25 pairs of QOSs with angular separations $< 3'$ are needed to reach a 4σ level of discrimination between different cosmological models.

Nusser and Haehnelt (1999) propose to use an analytical model of the IGM for a direct inversion of the absorption features in QSO spectra. The advantage of this approach is that no particular cosmological model has to be assumed.

The optical depth in redshift space due to resonant Lyα scattering is related to the H I density along the line of sight in real space by

$$\tau(w) = \sigma_0 \frac{c}{H(z)}\int_{-\infty}^{\infty} n_{HI}(x)\, \mathcal{H}[w - x - v_p(x), b(x)]dx, \qquad (4.26)$$

where w is the redshift space coordinate, σ_0 is the effective cross section for resonant line scattering, $H(z)$ is the Hubble constant at redshift z, x is the real space coordinate (in km s^{-1}), $\mathcal{H}$ is the Voigt profile normalized such that $\int \mathcal{H}\, dx = 1$, , v_p is the line of sight peculiar velocity and $b(x)$ is the Doppler parameter due to thermal/turbulent broadening. Assuming that the IGM that produces the absorption features in the Lyα forest is highly ionised and in photoionisation equilibrium

($n_{HI} \propto n_H^2 J^{-1}$), and that the gas density traces the fluctuations in the DM density on scales larger than the Jeans length:

$$n_{HI} = \bar{n}_{HI} \left(\frac{n_H(x)}{\bar{n}_H} \right)^{\beta} = \bar{n}_{HI} \left(\rho(x)/\bar{\rho} \right) \tag{4.27}$$

where the possible range for the parameter β is $1.56 \lesssim \beta \lesssim 2$ (Hui & Gnedin 1997) depending on the reionisation history and $\bar{n}_{HI}$ is the H I density at mean DM density. On scales larger than the Jeans length, Eq. (4.26) can be written as

$$\tau(w) = \mathcal{A}(z) \int_{-\infty}^{\infty} \left(\frac{\rho(x)}{\bar{\rho}} \right)^{\beta} \mathcal{H}[w - x - v_p(x), b(x)] dx, \tag{4.28}$$

where $\mathcal{A}$ is given in Eq. (4.17). The method presented by Nusser and Haehnelt directly recover the optical depth and the corresponding mass and velocity fields in the line of sight from the normalized flux, $F = e^{-\tau}$, using a Lucy-type iterative scheme (Lucy 1974). They showed that the density field can be successfully recovered below a threshold value above which the corresponding flux saturates. The inversion becomes increasingly more difficult at higher densities due to shell crossing and shock heating of the gas.

They further used mock spectra extracted from an N-body simulation to show that the moments of the mass density can be reasonably well estimated by fitting a log-normal PDF to the low density tail of the PDF of the recovered density.

This approach, even if more direct, is less indicated to estimate the correlation function and the power spectrum of the recovered density field than the method presented by Croft et al. (1998). The latter is more robust and more broadly applicable for several reasons:

1. it is relatively insensitive to what is occurring in saturated regions, which cannot be inverted directly from observations of Lyα alone, and which in any case are less likely to obey FGPA;

2. it relies mainly on large scale clustering information; it therefore does not require data that fully resolve all Lyα features;

3. the use of simulations in the normalizing procedure provides a convenient way to estimate the unknown parameter $\mathcal{A}$.

On the other hand, it will be affected by peculiar motions, and it is not able to

recover high density regions. The entity of the resulting errors has not yet been investigated.

4.6 Conclusions

The amplitude of the clustering of QSO absorbers gives a hint both on the nature of these objects and on the distribution of "normal" matter at high redshifts. In the case of QSO spectra, we are not biased by selection effects due to luminosity and we trace neutral hydrogen down to slightly overdense or even underdense regions.

The clustering properties of Lyα absorbers were a strongly debated topic across the years. We formed a sample of 15 QSOs observed at high resolution (between 9 km s^{-1} and 14 km s^{-1}), for a total of 1600 lines and we computed the TPCF along the line of sight. We assessed a $\sim 5\,\sigma$ deviation from poissonianity in the amplitude of the TPCF in the bin $50 \leq \Delta v \leq 150$ km s^{-1} for the lines with $\log N_{\rm HI} > 13.8$.

This result together with the Keck detection of C IV absorptions associated to Lyα lines with $\log N_{\rm HI} \sim 14$ changed the interpretation of the nature of these absorbers. They were no longer thought as completely separated from the heavy element absorbers and, consequently, from the galactic population.

Now, thanks to the scenario depicted by semi-analytical and numerical simulations, these absorption can be interpreted as due to the outer regions of galactic halos or to the filamentary structure which links them.

As regards metal absorption systems, the presence of clustering was confirmed since the first observations. We computed the TPCF of 71 C IV doublets from a subsample of the QSOs considered in the Lyα study. Due to the high resolution of the spectra, we could verify that there is a clustering signal up to $\Delta v \simeq 300$ km s^{-1}, in agreement with previous study at lower resolution.

In the last couple of years new approaches have been proposed to study the clustering properties of absorbers, to relate them more effectively to the analogous measure for galaxies and for the matter distribution.

The first preliminary results of the cross-correlation between multiple lines of sight and of the recovering of the power spectrum of the density have been

obtained recently.

In order to increase the significance and verify the viability of the obtained results, more and more QSO pairs and close groups have to be observed at intermediate to high resolution. The present instrumentation capabilities restrict to a few suitable candidates the choice for this kind of investigation.

I am part of a project for measuring the cross-correlation of C IV absorbers at $z \sim 2$ among multiple lines of sight. The observations at intermediate resolution of a group of 8 QSOs and two QSO pairs, spanning proper spatial scales $1-10\ h^{-1}$ Mpc, will be carried out in September and October 1999. This will supply the largest dataset ever obtained to study cross-correlation of C IV absorptions.

In the near future, a clear improvement in the number of known quasars will be given by the SLOAN survey, while the performance of the very high resolution UVES spectrograph at the UT2 ESO telescope will allow multi-object spectroscopy in 1 square degree fields.

Chapter 5

The Size and Geometry of Lyα Absorbers

In this chapter, I deal with the problem of the estimate of the size of QSO absorbers. The transverse dimension is a fundamental piece of information to distinguish between competing theoretical models for the absorbers and, together with the intensity of the ionising background, allows to compute the mass of these objects.

The paper by D'Odorico et al. (1998b) contains the original work described here.

5.1 Introduction

The lack of two-dimensional information is one of the main shortcomings of high redshift QSO spectroscopy. This problem makes it hard to understand the geometry of the absorbers, and to disentangle positions in velocity and real space. An intuitive way to put at least limits on the spatial dimension across the sky is through observations of common absorption systems in spatially separated lines of sight (multiple images of gravitationally lensed QSOs, or QSO pairs).

The ray path separation of gravitationally lensed QSOs usually spans subgalactic scales (< 100 kpc). The first studies of these spectra have put lower limits on the extension of the clouds of $\sim 5 - 30\ h^{-1}$ kpc (the range is due to the

uncertainties in the determination of the redshift of the lensing galaxies) simply considering the separation between the lines of sight at the redshift of the common absorption (Weymann & Foltz 1983; Foltz et al. 1984)

Smette and collaborators (1992, 1995) have used Monte Carlo simulations to determine the size of the clouds by comparing the observed ratio of the number of anticoincidences (lines that are present only in one spectrum and not in the other) and the total number of lines in the sample, to the corresponding ratio obtained by randomly distributing - over idealised identical, spherical clouds - pairs of lines of sight separated by the same linear distance as between the light-paths towards the two considered images. They derived a $2\,\sigma$ lower limit of $50\ h^{-1}$ kpc for the cloud diameter at $z \approx 2$. The estimated size, together with the strong correlation of equivalent widths and the small velocity differences of coincident lines, are hardly compatible with the relatively small clouds envisaged in the pressure confined model and in the mini-halo model (see § 6.1), nor they are consistent with the possibility that the absorption of a significant fraction of the systems could arise in the virialised regions of galaxies, or could exhibit the small scale variations typical of the interstellar medium.

Further results again pointing at very large sizes have been obtained for QSO pairs with line of sight separations of a few arcminutes or less.

Shaver and Robertson (1983) saw common absorption over $\sim 380\ h^{-1}$ kpc at $z \sim 2$ in metal absorption systems (Q0307-195A,B). Francis and Hewett (1993) found coincident damped Lyα absorption over $\sim 3\ h^{-1}$ Mpc across the sky.

Considering the firm lower limits from gravitational lensing there was reason to expect large sizes for the low column density forest as well, and such evidence was eventually found: Bechtold and collaborators (1994) and Dinshaw and collaborators (1994) obtained indipendently optical MMT data of the unusually close QSO pair Q1343+266A,B (also known as Q1343+264A,B; with a separation of 9.5 arcseconds corresponding to $\sim 40\ h^{-1}$ kpc at $z \sim 2$). The former obtained an estimate of the size of Lyα absorbers with the application of Bayes's Theorem (see below), while the latter followed the approach outlined by Smette et al. (1992); both of them obtained radii of $\sim 100\ h^{-1}$ kpc at redshifts just below 2.

Even larger sizes appear to occur at somewhat lower redshift. Dinshaw and collaborators (1995) deduced a most probable diameter of $360\ h^{-1}$ kpc for $0.5 <$

$z < 0.9$ from HST FOS spectra of the Q0107-25A,B pair (angular separation 1′.29). Dinshaw and collaborators (1994, 1995) observed also that coincident absorptions in the two lines of sight show small velocity differences, $\lesssim 100$ km s^{-1}, with respect to the velocity dispersions produced on the same spatial scales for clouds in the discs or virialised halos of normal galaxies (of the order of $\sim 200 - 400$ km s^{-1}). Additional observations of the spectra of the Q0107-25A,B pair indicate, using a simple maximum likelihood analysis (McGill 1990), a characteristic radius of 505 h^{-1} kpc assuming spherical absorbers and of 715 h^{-1} kpc in the case of disc-like geometry (Dinshaw et al. 1997, D97).

Fang and collaborators (1996, FDCB) developped a procedure based on Bayes's Theorem that points to a median diameter of $\sim 0.5\ h^{-1}$ Mpc for Q1343+266A,B. Furthermore, they compiled a sample with the closely separated QSO pairs observed at high resolution available in the literature and found a trend of larger inferred cloud radius with larger proper separation between the two sightlines. This could be due to nonuniform cloud sizes or to clustered clouds. Crotts and Fang (1998, CF98) studied the geometry of the absorbers using new spectra of the triplet Q1623+2651, the strong-lined absorbers seem consistent with flat, nearly round shapes expanding with the Hubble flow.

Stimulated by these last two papers and having at our disposal new observations of the QSO pair Q0307-195A,B, we focussed on this subject and critically reformulated the statistical machinery adopted up to that moment in order to obtain a reliable estimate of the radius of simply idealised absorbers.

5.2 Bayesian Statistics for Cloud Size Determination

As already said, a maximum likelihood technique to obtain an estimate of the absorber size was first proposed by McGill (1990) and later on it was improved by FDCB. We introduced some modifications to the method, in particular, to deal more reliably with the case of low numbers of coincidences and/or anticoincidences. In the following, we give a description of the procedure and we report the results obtained for a compilation of all the QSO pairs observed at high resolution

available in the literature at the time of our writing.

The spectra of observed QSO pairs show a number of *hits*, $\mathcal{N}_{\rm h}^o$, and of *misses*, $\mathcal{N}_{\rm m}^o$. A *hit* occurs when an absorption line above a given threshold in rest equivalent width, w_0, appears in both QSO spectra, with a velocity difference less than a limit, Δv. A *miss* occurs when a line is seen in any of the QSO spectra, but no line above the w_0 threshold is seen in the other.

The main idea applied by FDCB is to assume that the observed number of hits and misses is produced by a binomial random process; if the probability for a hit is given by Ψ, the likelihood function of this binomial process is

$$\mathcal{L}(\mathcal{N}_{\rm h}^o, \mathcal{N}_{\rm m}^o|\Psi) = \Psi^{\mathcal{N}_{\rm h}^o}\,(1-\Psi)^{\mathcal{N}_{\rm m}^o}. \tag{5.1}$$

Ψ is the probability that both ray paths intersect the cloud, given that at least one ray path does. Ψ is related to the probability Φ, that one given line of sight intersects the cloud given that the other adjacent ray path already does, by the formula:

$$\Psi = \Phi/(2-\Phi). \tag{5.2}$$

if the two ray paths have approximately the same propensities for detectable intersections with Lyα clouds. In practice, line detection thresholds can vary from one spectrum to the other and/or across a single spectrum because of the intrinsic magnitude difference of QSOs, their intensity variations with wavelength, and the spectrograph response function. Thus, given a line in spectrum A, the chance of detecting a corresponding line in spectrum B, Φ_A, is different from the chance Φ_B for detecting a line in spectrum A, given a line in B. That is, the probability Φ is line of sight dependent because of different intrinsic signal-to-noise ratio in the two-component spectra. The simplistic hypothesis we make for the nature of the Lyα absorbers do not justify such a detailed and elaborate theoretical fits of the data, thus in the following we will assume $\Phi \equiv \Phi_A = \Phi_B$.

The functional form of the probability Φ depends on the assumptions for the geometrical shape of the absorbers. We assumed two simple geometries:
1. Single radius, spherical clouds. Defining $X = S/(2\,R)$, where S is the proper separation of ray paths and R is the cloud radius, it is found (McGill 1990):

$$\Phi = (2/\pi)\,[\arccos X - X(1-X^2)^{1/2}] \quad \text{for } X < 1, \tag{5.3}$$

and $\Phi = 0$ otherwise.

2. Circular discs, with a given radius R, much larger then their thickness, and an observed inclination angle θ. The probability that one ray path intersects the disc within the angle $\theta \rightarrow \theta + d\theta$, given that the other ray path already does, is (McGill 1990):

$$\begin{aligned}\Phi(\theta) &= (\cos\theta/\pi)\left\{\arccos\left(\frac{X}{\cos\theta}\right) - \left(\frac{X}{\cos\theta}\right)\right. \\ &\quad \left.\times\left[1-\left(\frac{X}{\cos\theta}\right)^2\right]^{1/2}\right\} \quad \text{for } X < \cos\theta,\end{aligned} \tag{5.4}$$

and $\Phi = 0$ otherwise. By integrating over θ for randomly oriented discs, the probability is then given by:

$$\Phi = \int_{-\pi/2}^{\pi/2} \Phi(\theta)\, d\theta. \tag{5.5}$$

In practice, from the observations the numbers of hits and misses, $\mathcal{N}_{\rm h}^o$ and $\mathcal{N}_{\rm m}^o$, is obtained. Before using those numbers in the determination of the size, it is necessary to correct them for the number of accidental hits due to random velocity matches.

Two different procedures have been devised to compute the value of $\mathcal{N}_{\rm h}^o$: in the first, the velocity separation between all the line pairs is computed and, when a given line combines to give more than one hit, only the one with the smallest velocity separation is taken into account, and the others are neglected. If, for example, lines 1 and 2 of spectrum A combine with lines 1 and 2 of spectrum B forming the couples 1-1, 1-2 and 2-2, and the couple 1-2 has the smallest velocity separation, we count just one hit. The second method takes the minimum value between the "A vs. B" hits and the "B vs. A" hits (where A and B are the lines of sight to the QSOs in the pair).

The two methods turned out to yield the same result in all the investigated cases.

The number of observed misses is given by the sum of the two $\mathcal{N}_{\rm m}^o$'s, "A vs. B" and "B vs. A".

To take into account the possibility of *accidental hits* due to random velocity matches, we have used simulations of Poissonian distributions of absorption lines in the two QSO lines of sight, with a number of lines in each spectrum equal to the observed one. For each simulation we have computed:

- the number of accidental hits, $\mathcal{N}_{\rm r}$, applying the same procedures used for the real spectra;
- the corrected number of hits, $\mathcal{N}_{\rm h}$, (misses, $\mathcal{N}_{\rm m}$) by subtracting $\mathcal{N}_{\rm r}$ from (adding $2\,\mathcal{N}_{\rm r}$ to) the observed value.

Eventually, the probability distribution of the number of corrected hits $\mathcal{C}(\mathcal{N}_{\rm h})$ was evaluated, for $\mathcal{N}_{\rm h}$ varying from $\mathcal{N}_{\rm h}^o$ to 0, by dividing the number of simulations that gave a certain value for $\mathcal{N}_{\rm r}$ for the total number of simulations. The probability $\mathcal{C}(0)$ takes into account all the occurrences with $\mathcal{N}_{\rm r} \geq \mathcal{N}_{\rm h}^o$.

For each pair of values of corrected hits and misses, given the expressions for the probability distributions Φ and Ψ, it is possible to apply Bayes's Theorem (Press 1989, FDCB) to yield the a posteriori probability density for the radius R,

$$\mathcal{P}(R, \mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}) = \frac{\mathcal{L}(\mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}|\Psi)\, f(R)}{\int_0^\infty \mathcal{L}(\mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}|\Psi)\, f(R)\, dR}. \tag{5.6}$$

At variance with FDCB, we have chosen as a prior distribution, $f(R)$, a uniform distribution for R between 0 and $R_{\rm max}$, and not a uniform distribution of $\Psi(R)$.

It is worth stressing that a uniform prior distribution for the probability Ψ introduces a cutoff at high and low R that forces the 95 % confidence intervals to be narrow also in situations with a very low number of hits or misses. In particular, a spurious dependence of R on the proper separation S is induced in the case of very few hits, which is the common situation at large separations, because the probability $\Psi(R)$ actually depends on S. Since there is no particular justification to assume a uniform $\Psi(R)$, we have chosen a uniform distribution for R which does not introduce biases in the results.

Equation (5.6) then becomes

$$\mathcal{P}(R, \mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}) = \frac{\mathcal{L}(\mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}|\Psi)}{\int_0^{R_{\rm max}} \mathcal{L}(\mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}|\Psi)\, dR}. \tag{5.7}$$

We have adopted as $R_{\rm max}$ a value of 3 h^{-1} Mpc which represent the *mean free path* between two absorbers along the line of sight, as obtained from the line distribution $\partial^2\mathcal{N}/\partial z\, \partial N_{\rm HI}$ (Giallongo et al. 1996) at a redshift $z \sim 2$, in the column density range $14 \leq \log N_{\rm HI} \leq 17$.

In the case of discs, the maximum value is increased by a factor $\sqrt{2}$, yielding a value $R_{\rm max} \simeq 4\ h^{-1}$ Mpc.

It has to be noted that the value of this maximum radius does not influence in practice the resulting modal radii and confidence intervals, because the probability distributions for R usually go to zero well before this upper limit. $R_{\rm max}$ plays a significant role only if the number of misses is extremely low (1 or 0).

In the case of no hits, recipe (5.7) gives an incorrect result, in fact, the probability Φ becomes zero for $R < S/2$ by definition. In this extreme occurrence we have then to write:

$$\mathcal{P}(R, \mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}) = A \times \begin{cases} 1 & \text{if } R < S/2 \\ \mathcal{L}(\mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}|\Psi) & \text{if } R \geq S/2 \end{cases} \tag{5.8}$$

where, A is the proper normalisation factor which makes $\int_0^{R_{\rm max}} \mathcal{P}(R, \mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}) = 1$.

The global $\mathcal{P}(R)$ for a given QSO pair is finally obtained by the sum:

$$\mathcal{P}(R) = \sum \mathcal{P}(R, \mathcal{N}_{\rm h}, \mathcal{N}_{\rm m}) \times \mathcal{C}(\mathcal{N}_{\rm h}), \tag{5.9}$$

for $\mathcal{N}_{\rm h}$ which goes from $\mathcal{N}_{\rm h}^o$ to 0.

5.3 Results and Discussion

We observed the QSO pair Q0307-195A,B (already studied by Shaver & Robertson 1983) at a resolution of 0.07 and 0.05 nm in the UV and blue range respectively. Both QSOs are about of 19th magnitude, their redshifts are $z_{\rm em}({\rm A}) = 2.1439 \pm 0.0003$ and $z_{\rm em}({\rm B}) = 2.1217 \pm 0.0003$ (SR83), and they are separated by 56 arcseconds on the plane of the sky.

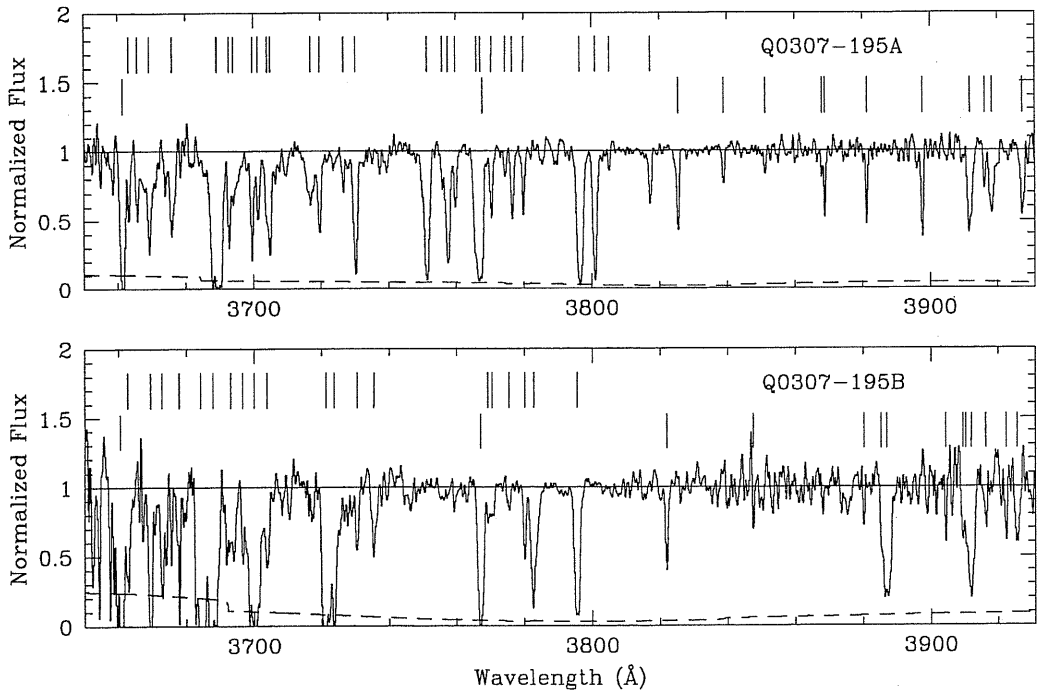

Figure 5.1: Spectra of Q0307-195A,B obtained with the NTT EMMI instrument as a function of vacuum heliocentric wavelength. The dashed line shows the 1σ error in the flux. Tick marks indicate absorption lines detected at or above the 3.5σ confidence level, upper and lower tick marks are referred to Lyα and metal absorption lines respectively

The normalised spectra are shown in Figs. 5.1, 5.2.

In order to investigate the possible correlations of the radius R with the line of sight separation and with the redshift, we have applied our analysis to 7 other data samples of comparable quality, published in the literature: Q1343+2640A,B (Crotts et al. 1994), Q1026-0045A,B (Petitjean et al. 1998), Q0107-025A,B (D97), Q1517+2356,1517+2357 (CF98), Q1623+2651A,1623+2653,1623+2651B (CF98). The QSO pair Q1343+2640A,B has been observed also by Dinshaw et al. (1994), we have chosen to use Crotts et al. (1994) line list because of the larger wavelength coverage and larger number of identified lines.

The relevant information are displayed in Tab. 5.2. Columns 6 and 7 report the observed number of hits and misses, respectively.

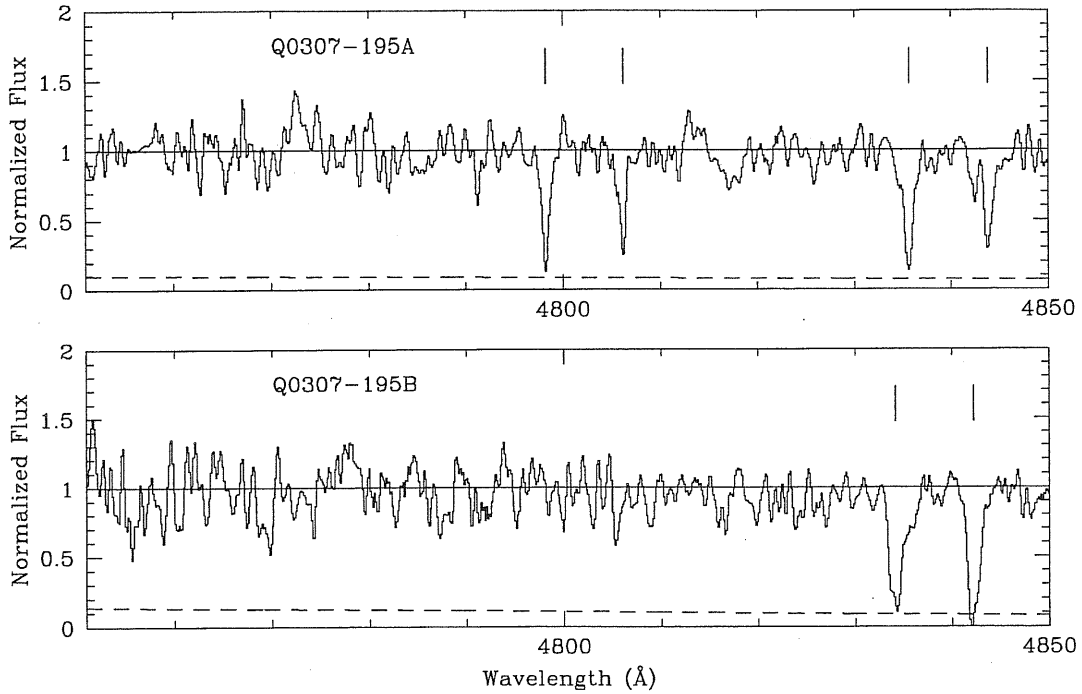

Figure 5.2: Same as Fig. 5.1, for the C IV region

In order to avoid spurious effects due to the "proximity effect", only the Lyα lines with velocity separation greater than 5000 km s^{-1} from the lower QSO redshift in the pair were considered in the sample.

A value $\Delta v = 200$ km s^{-1} has been adopted as the most suitable velocity window to define hits and misses, being significantly larger than the resolution ($\sim$ 40 km s^{-1}) and close to the clustering scale observed for Lyα absorbers (Cristiani et al. 1997). As a consequence, all the Lyα lines in the sample whose separation was less than 200 km s^{-1} have been merged into a single line with wavelength given by the average of the wavelengths and equivalent width equal to the sum of the equivalent widths of the "parent" lines.

In Tab. 5.1, the hits observed in the spectra of Q0307-195A,B are reported together with the velocity differences between the lines.

The most probable number of accidental hits for each QSO pair is: 0 for 1343+2640A,B; 0 for 1026-0045A,B (for both $w_0 = 0.2$ and $w_0 = 0.3$); 2

Table 5.1: Lyα lines that give rise to hits in the spectrum of Q0307-195A,B

No.	λ_A (Obs.) (Å)	W_A (rest) (Å)	λ_B (Obs.) (Å)	W_B (rest) (Å)	$v_A - v_B$ (km s^{-1})
1*	3669.36	0.33	3669.45	0.54	-7±9
2	3676.08	0.32	3677.81	0.27	-141±8
3*	3689.15	1.34	3687.80	1.19	110±9
4	3694.16	0.23	3693.10	0.54	86±16
5	3699.64	0.26	3700.04	1.25	-32±4
6	3704.01	0.20	3703.81	0.28	16±13
6	3704.98	0.22	3703.81	0.28	95±8

* Hit in the case $w_o \geq 0.3$

for 0307-195A,B ($w_0 = 0.3$); 6 for 0307-195A,B ($w_0 = 0.2$); 0 for 0107-025A,B; 1 for 1517+2356/57; 2 for 1623+2651A,B; 2 for 1623+2651A/53 and 5 for 1623+2653,51B.

In the actual computation of the cloud sizes the average value of the proper separation in the Lyα redshift range has been adopted (see Tab. 5.2).

In Fig. 5.3 the probability distributions given by Eq. (5.9) for all the QSO pairs presented in the paper are plotted for the two cases of spherical and disc geometry.

In Figs. 5.4 and 5.5 the values estimated for the modal radii are plotted as a function of S. We consider that if the *typical* radius $R_{50} \leq S/2$, where R_{50} is defined as the radius at which the probability equals the 50 %,

$$\int_0^{R_{50}} \mathcal{P}(R) = 0.5, \tag{5.10}$$

only an upper limit can be set to the dimension of the absorber. This is the case for Q0307-195A,B (with $w_0 \geq 0.03$ nm) for which $R_{50} \simeq 90\ h^{-1}$ kpc $< 114\ h^{-1}$ kpc for the spherical model and $R_{50} \simeq 107\ h^{-1}$ kpc $< 114\ h^{-1}$ kpc for the disc model. The dashed line represents the most probable modal radius, $R = 362\ h^{-1}$ kpc, with 95% confidence limits $298 < R < 426\ h^{-1}$ kpc (spherical absorbers) and $R = 412\ h^{-1}$ kpc, with 95% confidence limits $333 < R < 514\ h^{-1}$ kpc (disc). These values correspond to the peak of the total probability density obtained as the product of the probability densities of the considered QSO pairs.

The quality of the present data on the pair Q0307-195A,B and of the data by Petitjean et al. (1998) for the pair Q1026-0045A,B allows defining a complete

Table 5.2: Parameters of the QSO pairs forming our enlarged sample and Lyα cloud radius estimates in the hypothesis of spherical and disc geometry

							Spheres		*Discs*	
QSO Pair	Angular Separ. (arcsec)	Lyα z Range	S, Proper Separation (h^{-1} kpc)	W_0 cutoff (Å)	$\mathcal{N}_h^o$	$\mathcal{N}_m^o$	95% Confid. Interval in R (h^{-1} kpc)	Modal Radius (h^{-1} kpc)	95% Confid. Interval in R (h^{-1} kpc)	Modal Radius (h^{-1} kpc)
1343+2640A/B[1]	9.5	1.756-1.979	39-40	0.4	6	1	50-2705	80	70-3000	117
1026-0045A/B[2]	36	0.833-1.398	149-155	0.2	5	15	98-268	137	117-385	179
				0.3	4	12	98-305	139	118-443	183
0307-195A/B[3]	56	2.011-2.070	228-229	0.2	6	7	10-294	146	11-406	171
				0.3	2	9	< 269		< 365	
0107-025A/B[4]	86	0.481-0.892	301-361	0.3	5	6	291-1776	475	379-2636	663
1517+2356/57[5]	102	1.390-1.787	427-438	0.4	3	21	34-460	277	39-618	330
1623+2651A/B[5]	127	2.025-2.409	497-519	0.4	5	20	99-604	348	113-828	425
1623+2651A/53[5]	147	1.971-2.409	575-604	0.4	7	18	356-883	481	414-1247	619
1623+2653/51B[5]	177	2.025-2.468	687-723	0.4	5	32	27-600	418	30-773	482

[1] Crotts et al. 1994
[2] Petitjean et al. 1998
[3] This paper
[4] Dinshaw et al. 1997
[5] Crotts & Fang 1997

sample of lines down to an equivalent width $w_0 = 0.02$ nm. The resulting number of hits and misses, the radius and confidence interval are reported in Tab. 5.2. The corresponding values of the global modal radius and 95 % confidence interval are $R = 350\ h^{-1}$ kpc and $276 < R < 417\ h^{-1}$ kpc (sphere), $R = 408\ h^{-1}$ kpc and $325 < R < 501\ h^{-1}$ kpc (disc).

The extended confidence intervals in Figs. 5.4, 5.5 decrease the evidence of a correlation between the separation of the lines of sight and the expected radius of the Lyα absorbers, with respect to FDCB and CF98 determinations. However, if we compute a linear fit of R versus S, adopting as an error on the radius the minimum difference between the radius itself and the extrema of the confidence interval, we find a correlation with a slope 0.7 ± 0.2 in the case of the sphere model and 0.8 ± 0.2 in the case of the disc model. Such correlation would not be expected if the hypothesis of unclustered, single sized clouds was realistic, and it can be ascribed to cloud clustering, non-spherical shape and/or a range of cloud sizes spanning the observed range in S.

We performed another test to investigate the nature of the cloud population.

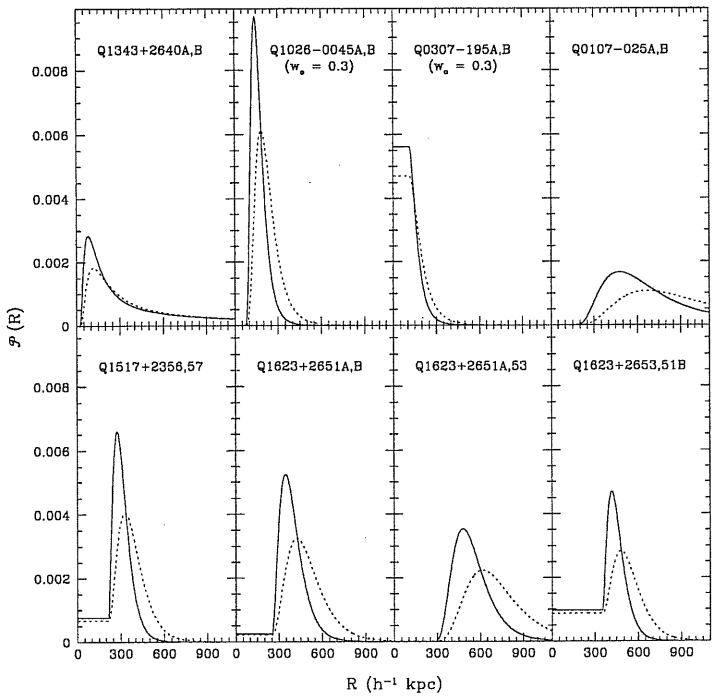

Figure 5.3: Probability distribution $\mathcal{P}(R)$ as given by eq. 5.9 for spherical absorbers (*solid curve*) and randomly inclined discs (*dotted curve*) as functions of cloud radius for all the QSO pairs considered in the paper. The plateau at separation below $S/2$ arises because there is a non-vanishing probability that the number of corrected hits be 0 ($\mathcal{C}(0) \neq 0$)

We divided the sample of Lyα lines ordered in equivalent width in two halves, the central equivalent width results to be $w_{\rm half} = 0.053$ nm (that for $b = 30$ km s^{-1} corresponds to $\log N_{\rm HI} \simeq 15.6$). The two new samples have been used again to compute the modal radii of Lyα absorbers for the two simple geometries: sphere and disc.

The upper sample gives a best estimate radius $R = 360\ h^{-1}$ kpc, with 95% confidence limits $273 < R < 442\ h^{-1}$ kpc (sphere) and $R = 414\ h^{-1}$ kpc, with 95% confidence limits $317 < R < 538\ h^{-1}$ kpc (disc). These values are quite similar to the ones obtained with the complete sample, apart from an increase of the 95 % confidence interval due to the fact that the number of hits and misses has decreased. On the other hand, the low equivalent width sample is characterized

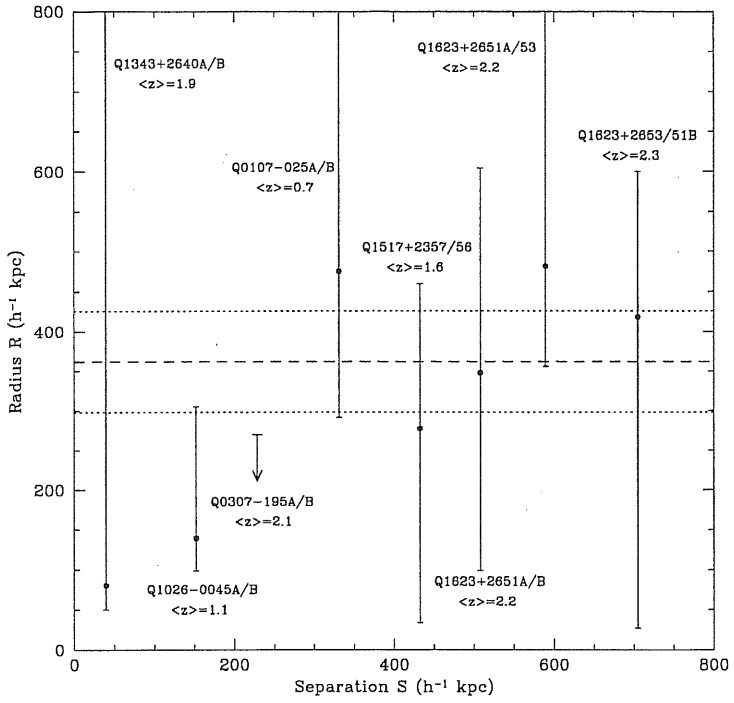

Figure 5.4: Inferred cloud radius R (and 95 % confidence interval) from a model assuming single-radius, unclustered spherical clouds, as a function of QSO pair sightline separation S. The dashed line shows the most probable value for R and the dotted lines the 95 % confidence interval (see text)

by the fact that almost all the pairs present no observed hits.

This last result is at a variance with the findings of the simulations of Cen and Simcoe (1997, see below), in which the number of hits increases with decreasing column density threshold. The overall result points against the picture of a population of spherical (or disc-like) single-radius clouds, from which we would have expected a decrease of the size with increasing equivalent width threshold (in the hypothesis, suggested by observations, that the column density in the cloud falls off gradually from a central peak).

A plot of the modal radius as a function of the average redshift of the Lyα forest of the corresponding QSO pair is shown in Figs. 5.6, 5.7. No correlation can be inferred between the two, the probability that the observed poor χ^2 has occured

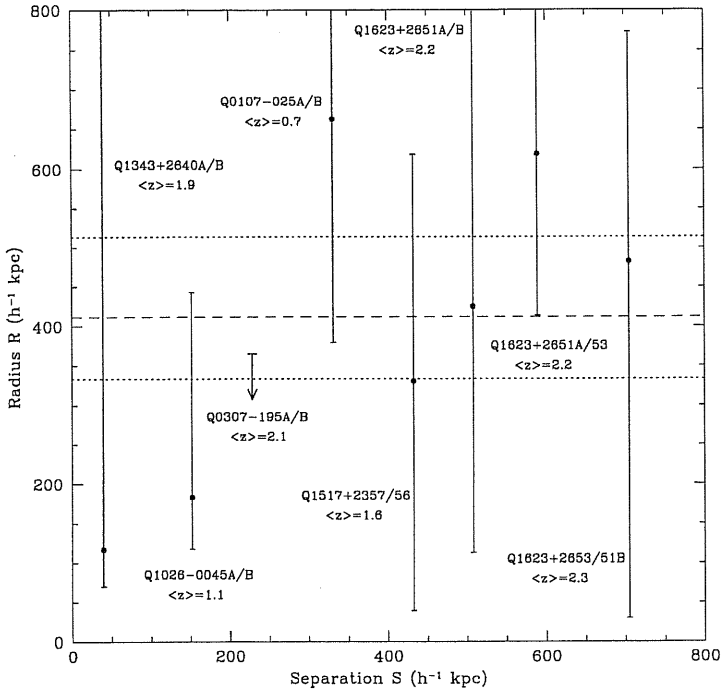

Figure 5.5: Same as Fig. 5.4 but for single-radius, unclustered thin discs

by chance (i.e., the goodness-of-fit parameter) is 0.004 and 0.06 for the sphere and disc model respectively. Thus, our compilation of data does not confirm the trend of incresing dimension with decreasing redshift found by Dinshaw et al. (1998).

Given the large transverse size determined at $z \sim 2$ by our analysis, a constraint on the geometry of the absorbers can be put by the following argument. With the presently inferred sizes and the typical UV background conditions at $z \sim 2$ (Giallongo et al. 1996), optically thin absorbers of $\log N_{\rm HI} \sim 14 - 15$ are expected to show a remarkably small fraction of neutral Hydrogen, roughly $10^{-5.5} - 10^{-5}$. In this way the mean intergalactic density contributed by these clouds (Hu et al. 1995), especially in the spherical case, is close to conflict with the baryon limit from nucleosynthesis $\Omega_b \leq 0.015h^{-2}$ (Walker et al. 1991). This potential incongruity can however be alleviated if flattened geometries are assumed.

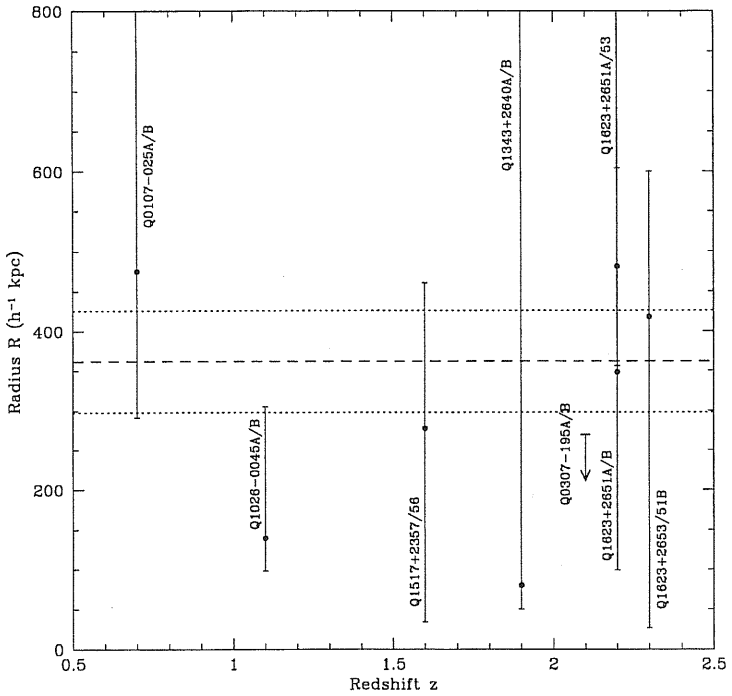

Figure 5.6: Inferred cloud radius R and 95 % confidence interval for spherical clouds (see Fig. 5.4) as a function of the average redshift of the Lyα forest in the QSO pair spectra. The dashed line shows the most probable value for R and the dotted lines the 95 % confidence interval (see text)

5.4 Towards More Realistic Models

The approach described in the previous section, although customarily used to determine the size of the Lyα absorbers, is expected to give just a rough idea of the dimension of these structures, due to the simplistic hypotheses made on their geometry.

In particular, the information on the equivalent width of the lines are not taken into account, since no density distribution is assumed inside the absorbers. Besides, all the clouds are assumed to have the same radius.

In the literature (Smette et al. 1992, 1995; Charlton et al. 1997; Dinshaw et al. 1997) this problem has been tackled adopting a statistical technique that utilizes the information about the equivalent width of the lines.

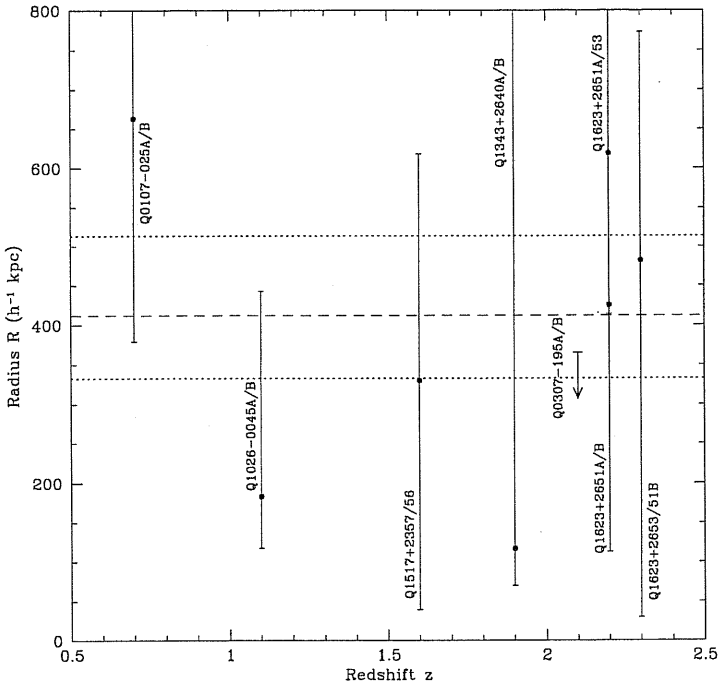

Figure 5.7: Same as Fig. 5.6 but for single-radius, unclustered thin discs

Dinshaw et al. (1997) adopted a power-law column density profile for the absorbers, $N(r) = N_{\rm lim}(r/R_0)^{-\gamma}$, where $N_{\rm lim} = 1.26 \times 10^{14}$ cm^{-2} for $b = 35$ km s^{-1} is the limiting column density of the sample, R_0 is the radius of the absorber at $N = N_{\rm lim}$ and $\gamma = 4$, and three simple geometries with fixed radius: spheres, randomly inclined discs, and pseudo-filamentary structures (approximated as discs with fixed inclination $\cos i = 0.2$).

The column density distribution along paired lines of sight for each of the geometric models considered is simulated by means of the Monte Carlo techniques outlined by Smette et al. (1992). The relevant features coming out from the analysis of the plots of $\log N_A$ versus $\log N_B$ are the following (D97):

1. for both spherical and filamentary absorbers there are sharp boundaries delineating regions of significant and zero probability, this is not true for randomly inclined discs for which the column density across the disc and the radius of the disc depend on the inclination;

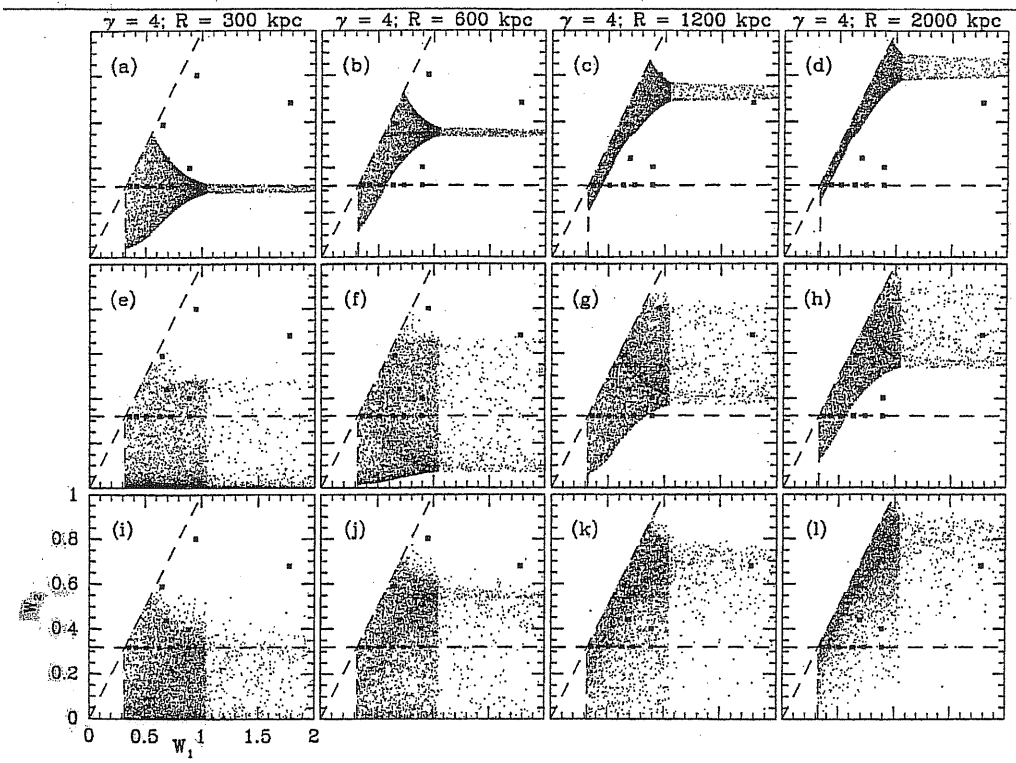

Figure 5.8: Equivalent width distributions for spherical halo models (*top row of panels*), filamentary structures (*middle row of panels*), and randomly inclined discs (*bottom row of panels*) as compute by D97. The observed and simulated equivalent widths are arranged such that $w_1 \equiv \max(w_A, w_B)$ and $w_2 \equiv \min(w_A, w_B)$. The horizontal dashed lines represent the rest equivalent width threshold ($w_0 > 0.032$ nm) of the "strong sample" of lines of Q0107-025A,B defined in D97. The points lying above the dashed lines represent hits, and those falling below the orizontal dashed line are misses. The superimposed solid squares are the equivalent widths of the observed hits and misses for the D97 strong-line sample. Reproduced from Dinshaw et al. (1997)

2. models involving spherical clouds with radii comparable to the line of sight separation cannot produce coincidences at large column densities, whereas anti-coincidences at large column densities cannot arise in large spherical absorbers;

3. for radii progressively larger than the line of sight separation, the column densities become more highly correlated;

4. for progressively larger γ (i.e., steeper drops in column density across the absorber), the cross section for detecting high column density systems increases, producing a greater spread in the probability density because a larger range of column densities is sampled.

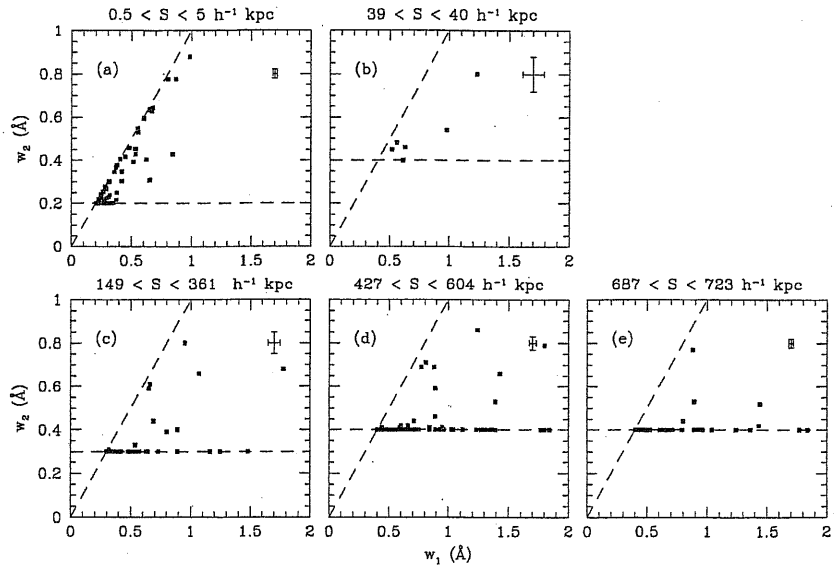

Figure 5.9: Equivalent widths of the hits and misses for the 8 QSO pairs in our sample plus the data by Smette et al. (1995), they are grouped as follows: (a) HE1104-1805; (b) Q1343+2640A,B; (c) Q1026-0045A,B Q0307-195A,B and Q0107-025A,B; (d) Q1517+2356,1517+2357 Q1623+2651A,B and Q1623+2651A,1623+2653; (e) Q1623+2653,1623+2651B. The equivalent widths are arranged such that $w_1 \equiv \max(w_A, w_B)$ and $w_2 \equiv \min(w_A, w_B)$. The horizontal dashed lines represent the rest equivalent width threshold: $w_0 > 0.02$ for (a), $w_0 > 0.03$ for (b), (c) and $w_0 > 0.04$ for (d), (e). The error bars in the upper right corners of the plots represents the average 1 σ errors for the equivalent width

The distribution of column densities can be converted into the corresponding equivalent width distribution adopting an approximation to the curve of growth, having an accuracy better than 10 % in the range of optical depths $0.1 \leq \tau \leq 40$ (Chernomordik & Ozernoy 1993):

$$w = \frac{\sqrt{2}b\lambda}{c}\left[\ln\left(1+\frac{\pi}{2}\tau^2\right)\right]^{1/2}, \tag{5.11}$$

where λ is the wavelength of the line, c is the speed of light, $\tau = 1.497 \times 10^{-2}\lambda f N/b$ is the optical depth, and $f = 0.416$ is the oscillator strength for the Lyα transition.

Figure 5.8 (D97) shows the equivalent width distributions for spherical (*top*

panels), pseudo-filamentary (*middle panels*), and disc-like (*bottom panels*) absorbers with superimposed the equivalent width of the observed coincidences and anticoincidences for the sample of lines of Q0107-025A,B with $w > 0.032$ nm.

In Fig. 5.9, we plotted the observed equivalent widths of the hits and misses for an enlarged sample of QSO pairs formed by our previously described 8 cases plus the lensed QSO HE 1104-1805A,B observed by Smette et al. (1995). The pairs are grouped according to their typical proper separation. It should be noted that it is now certain that the double QSO HE 1104-1805A,B is actually a gravitational lens system, as the lensing galaxy has been detected (Courbin et al. 1998; Remy et al. 1998). The exact value of the separation between the two light beams depends on the still badly determined lens redshift. The values we reported in Fig. 5.9 have been estimated using Fig. 6 of the paper by Smette et al. (1995), for the redshift $z_{\rm lens} = 1.32$, favoured by Remy et al. (1998).

It is clear from the figure that the fraction of hits decreases as the proper separation increases, while, at the same time, the correlation between the equivalent widths of the absorbers in the two line of sight becomes poorer and poorer.

A comparison of Fig. 5.8 with Fig. 5.9(a) shows reasonable agreement with the prediction of a disc model with $R_0 \sim 7\,S$ that is $R_0 \sim 90\ h^{-1}$ kpc; panel (b) agrees with the case $R_0 = 2 - 3\ S$ for both filament and disc model, that is $R_0 \simeq 80 - 120\ h^{-1}$ kpc; panel (c) can be explained by filaments and disc with $R_0 = S$ which implies $R_0 \sim 220\ h^{-1}$ kpc and, finally, panels (d) and (e) agrees with filament and disc models with $R_0 \leq S$ or $R_0 \lesssim 500 - 700\ h^{-1}$ kpc. Altogether a disc geometry with $R_0 \sim 100 - 200\ h^{-1}$ kpc seems to be favored.

It should be noted, however, that the number of degrees of freedom is not the same in the three models: once a given radius is chosen in the disc idealisation the two orientation angles can vary at random, for the filament geometry one angle is fixed, and for the sphere no angular variable exists to define an orientation. In this way, it is obviously easier for the disc model to reproduce the real data that show a substantial amount of scatter due to measurement errors and possibly to a non-unique value of the cloud size.

In fact, by introducing one more degree of freedom in the spherical cloud model: i.e. allowing the radius R to vary according to one of the two distribution: (1)exponential, and (2)power law with indices $\alpha = -4, -3, ..., 5, 6$, it is found that

a power law distribution with $\alpha = 3$ provides an equally reasonable representation of the observed equivalent width distribution (D97).

5.5 Hints from Numerical Simulations

The relevance of cosmological simulations for the study of the nature of QSO absorbers will be discussed in the next chapter.

Thousands of synthetic pair spectra can be obtained from double lines of sight crossing randomly simulated cubes. These large database can be processed as the observed spectra to verify the reliability of the algorithm applied to estimate the size of the absorbers, study the dependence on transverse separation, redshift and column density threshold and work out tests sensitive to the shape and distribution of the absorbers.

On the other hand the observationally determined size of the absorbers is a key element to identify the corresponding objects in the simulations (see, e.g., Cen et al. 1994, dealing with $\sim$ 20 kpc proper diameter clouds, in contrast with Miralda-Escudé et al. 1996, dealing with clouds 100 kpc wide and 1 Mpc long) and the comparison of results from large compilations of real data with those of simulated spectra can put constraints on the cosmological model originating the simulation.

Here we review the two main theoretical works that propose observational tests involving pairs of sightlines.

Charlton et al. (1997) performed two hierarchical N-body/hydrodynamical simulations of a CDM-dominated flat universe in two boxes of comoving dimensions 3.2 and 9.6 Mpc. Sinthetic spectra were generated for variuos random lines of sight through the simulation boxes at z =3, 2 and 1, following Zhang et al. (1997) procedure. A Bayesan estimate of structure sizes from the simulated double lines of sight yields (1) larger sizes for lower $N_{\rm HI}$ contour levels, as would result from a density distribution that decreases smoothly from a central value, (2) a size that increases with the double lines of sight separation used to probe it, as obtained from observations and, (3) a decrease of the size at fixed column densities with decreasing redshift, this could be partly due to the effects of contamination by chance pairs and the increased line densities at higher redshifts.

Three diagnostic have been applied to the synthetic sample of coincidences

and anticoincidences: (1) the distribution $\log N_A$ versus $\log N_B$, (2) the $N_{\rm HI}$ distribution of anticoincidences, and (3) the distribution of velocity differences $f(\Delta V \equiv |V_A - V_B|)$. The results have been compared with Monte Carlo simulations of double lines of sight for a spherical structure with an isotropic velocity field, a constant rotation "disc", and a flattened "slab" with constant velocity flowing radially toward or away from a central point (Charlton, Churchill & Linder 1995). They show good agreement with those obtained from idealised models in which Lyα absorption is produced by a single population of discs with $N_{\rm HI}(r) \propto r^{-4}$. In particular: (1) Hits exhibit an increasing spread of column densities from the correlation line as the double lines of sight separation is increased. (2) The distribution of column densities for anticoincidences does exhibit some large values for double lines of sight with all separations. This indicates that spherical clouds are not a good description of the simulated data. (3) The velocity difference distribution for coincident lines becomes wider for more widely spaced double lines of sight, as expected for coherent motion in the plane of a disc.

A visual inspection of the structures responsible for absorption in a slice through the 3.2 Mpc box at redshifts 3, 2 and 1 shows that contours of increasing column density trace from sheetlike, through elongated, to spherical shapes. The lower density threshold which define extended structures at high redshifts, correspond to more compact objects as the universe expand to smaller redshifts, and structures at the same threshold thus appear to decrease in size. At $z = 2$ the filamentary structures appear to have a coherence length ranging from a few hundred kiloparsecs up to about 1 Mpc, with a thickness of $100 - 200$ kpc.

A more complex approach was followed by Cen and Simcoe (1997) with a hydrodynamic simulation of a spatially flat CDM universe with a non-zero cosmological constant over the redhift range $2 \leq z \leq 4$. They did not restrict their analysis to simulated spectra, but they define Lyα clouds in their simulated universe as regions with density above a chosen threshold and investigated the relationship between absorption and absorber in a closer way. They found that clouds are larger at lower $\delta\rho$, as expected, and exhibit a wide range in size, from a few to about a hundred proper kiloparsecs. The median size varies from 15 h^{-1} kpc for $\delta\rho = 30$ (corresponding to $N_{\rm HI} \sim 10^{15}$ cm^{-2}) to 35 h^{-1} kpc for $\delta\rho = 3$ (corresponding to $N_{\rm HI} \sim 10^{13}$ cm^{-2}). Furthermore, Lyα clouds display a vari-

ety of shapes ranging from quasispherical clouds to filaments and sheets. This analysis of simulated quasar double sightlines allows to distinguish the nature of coincident absorption lines. At small transverse double line of sight separations ($S < 40\ h^{-1}$ kpc proper), a hit most likely represents absorption by a common cloud. However, at larger S, a pair of coincident lines most likely samples two separate clouds which belong to the same cloud cluster. At $S = 100\ h^{-1}$ kpc, 80-90 % of hits can be explained in this fashion. At very large separations, $S > 500\ h^{-1}$ kpc, the coincident line events are entirely due to random intersections of two unrelated, uncorrelated clouds. A signature of clustering-dominated coincident lines would be that the difference in line of sight velocity between the two absorption lines should be a weak function of the separation perpendicular to the line of sight.

Cen and Simcoe infer that the estimate of the actual size of Lyα clouds from the analysis of observed coincident absorption lines in the hypothesis of simplified geometry (e.g., spheres, discs) is reasonably accurate only when the sightline separations are small ($S \leq 50\ h^{-1}$ kpc). For larger S, the inferred sizes are not related to the true sizes of the clouds and are grossly inflated.

5.6 Conclusions

The determination of the transverse size of QSO absorbers is an intriguing topic which obviously holds a big relevance in the investigation of the nature of these objects.

A natural way to put at least lower limits on the dimension is to exploit double lines of sight toward gravitationally lensed QSOs or close pairs and look for common absorptions.

To properly face this issue, we compiled a sample with all the QSO pairs observed at intermediate to high resolution available in the literature for a total of 8. We looked for hits and misses in the pair spectra and applied a Bayesian statistic to compute the probability distribution for a given radius R in the case of spherical and disc geometry (McGill 1990 and Fang et al. 1996). Our approach presents some improvements in order to deal correctly with the extreme cases of very low numbers (1 or 0) of hits or misses and avoid any spurious dependence on the separation between the lines of sight. Our results point to very large Lyα

clouds with proper radius $R \sim 350 \pm 50\ h^{-1}$ kpc. These sizes rule out the classic theoretical models for Lyα clouds pressure confined by an external medium and gravitationally confined by dark matter minihalos which predicted very small objects.

Our approach lowered the significance of the correlation radius versus separation of the lines of sight, even if we cannot state its disappearance. On the other hand, we do not find any correlation of the radius with the average redshift of the Lyα absorbers.

If we take into account the information on the equivalent width of the lines and compare the distribution w_A versus w_B with the results of Monte Carlo simulations by Dinshaw et al. (1997), a disc shape with density profile $N(r) \propto (r/R_0)^{-4}$ and $R_0 \sim 100 - 200\ h^{-1}$ kpc seems favoured.

Large database of synthetic pair spectra obtained from cosmological simulations for different line of sight separation, redshifts and column density threshold constitutes a valuable tool to test the viability of the size estimation algorithm and suggest sensitive tests to diagnosis the shape of absorbers. Cen and Simcoe (1997) suggest that the size we detect are inflated by accidental hits in double lines of sight with large transverse separation. Significative results in this sense will be obtained only when large samples of real data will be available, the study of close QSO pairs and groups is a major target of 8-10meter class telescopes.

Chapter 6

Models and Simulations of QSO Absorbers

In this chapter, I give an overview on the theoretical side of the study of QSO absorption spectra. The first part is a review of the models proposed for the intergalactic population of clouds in which Lyα forest absorptions should arise, different kind of confinement were introduced to reproduce the density and temperature properties of these objects.

The latter half of the chapter describes the "modern" approach - by semi-analytical and numerical simulations - that disclosed a new scenario for the origin of QSO absorption systems.

6.1 The first generation of models

Theoretical modelling of absorption systems can be traced back to Spitzer's (1956) prediction (expanded by Bahcall & Spitzer 1969) that normal galaxies have large gaseous halos giving rise to heavy element UV absorption lines. Bahcall and Salpeter (1965) considered groups of galaxies, Arons (1972) suggested forming low mass protogalaxies as the probable sites of Lyα absorption. Sargent and collaborators (1980) asserted that observational results indicated a different nature for metal and Lyα forest absorption systems, with the evidence for the latter pointing away from galaxies, to distinct astronomical objects, intergalactic gas clouds.

6.1.1 Lyα Clouds Confined by the Pressure of an Intercloud Medium

If the Lyα absorbers correspond to overdense clumps of gas, their persistence throughout the history of the universe must be either due to only a slow change in their properties, or to replenishment of the clouds on a shorter time scale. An apparent lack of evolution in the number density of the forest lines (later shown to be a statistical fluke) justified the treatment of the clouds as "self-contained entities in equilibrium" (Sargent et al. 1980). A two phase intergalactic medium was postulated, with the hot, tenuous intercloud medium (ICM) in pressure equilibrium with the cooler and denser Lyα clouds. The basic properties of the standard version of the pressure confinement model (Sargent et al. 1980; Ostriker & Ikeuchi 1983; Ikeuchi & Ostriker 1986) can be summerized as follows:

The Lyα absorbers are supposed to be spherical, and, since gravity is ignored, homogeneous clouds. They are in photoionisation equilibrium with an ionising UV background. The gas is heated by photoionisation, and cools via thermal bremsstrahlung, Compton cooling, and the usual recombination and collisional excitation processes. The ICM is expanding adiabatically by the cosmic expansion at all times because the high degree of ionisation does not allow for efficient photoionisation heating. The denser clouds embedded in the hot ICM start out in isothermal expansion with a temperature fixed by thermal ionisation equilibrium ($T_c \sim 3 \times 10^4 K$), until the density $n_c \propto P_{\rm ICM}/T_c \propto (1+z)^5$ has dropped sufficiently that photoheating cannot compensate for the work of expansion any longer, and the clouds begin to cool and to expand less rapidly. The sound speed drops even faster so ultimately pressure equilibrium with the ICM ceases and the clouds enter free expansion.

The available range of cloud masses is constrained by the requirements that the clouds must be small enough not to be Jeans-unstable, but large enough not to be evaporated rapidly when heated by thermal conduction from the ambient ICM (Sargent et al. 1980; Ostriker & Ikeuchi 1983). Clouds formed at $z \sim 6$ would survive down to accessible redshifts (~ 4) only if their masses range between $10^5 < M_{\rm c} < 10^{10}\ M_\odot$.

This model is self-consistent, but there are no very compelling physical rea-

sons for preferring pressure to gravitational confinement or to no confinement at all, and the possibility of self-gravitating clouds as an alternative was discussed soon (Melott 1980; Black 1981).

Nevertheless, the pressure confinement model for Lyα clouds has some appealing properties: The hot intercloud medium might be a possible source of the X-ray background. The explosion scenario of structure formation (Schwarz, Ostriker & Yahil 1975; Ostriker & Cowie 1981) provides a theory for the formation of the clouds. Finally, the model makes testable predictions, which indeed paved the way to its demise.

If Lyα clouds were originated in cosmic shocks due to large scale explosions from galaxies and QSOs, a correlation with galaxies would have been expected. In particular, the absorbers should be clustered in a manner similar to that of galaxies (Vishniac & Bust 1987). Whereas, the observed clustering signal does not show the requisite strength (Barcons & Webb 1990).

A weighty problem is represented by the reproduction of the large range of neutral hydrogen column densities observed. The following relationship is valid,

$$N_{\rm HI} \propto M_{\rm c}^{1/3} T_{\rm c}^{-29/12} J^{-1} P^{-5/3}. \tag{6.1}$$

To reproduce only the low column density systems between $13 < \log N_{\rm HI} < 16$ the mass has to vary by 9 orders of magnitude, or the radiation field by 3 orders, or the pressure by a factor of 63. To ensure cloud survival the mass range is limited to less than 4 dex (see above), and the temperature is constant; furthermore, looking for pressure related spatial correlations among the equivalent widths of Lyα forest lines excluded pressure fluctuations $\Delta P/P > 14\%$ at the 2σ level (Webb & Barcons 1991), and a similar limit must hold for the radiation field J.

To summarize, the pure pressure confinement model is unlikely to explain the Lyα forest as a whole, though it is clear that some line of sight must go through sites where gas is locally confined by external hydrostatic or ram pressure.

6.1.2 Gravitational Confinement: Selfgravity

Self-gravitating baryonic clouds were suggested by Melott (1980) as an alternative to the pressure confinement model. Black (1981) investigated in detail the phys-

ical structure of such objects. He assumed that clouds are spheres in hydrostatic equilibrium, thus they exhibit gradients of temperature, density and pressure. He found that gravitationally stable clouds with properties consistent with the observations have to be extended ($\sim$1 Mpc), and must either be truncated and bounded by the pression of an external medium, or be large enough to overlap, providing their own boundary pressure. In this model the appearance of the intergalactic medium as a forest of lines is more due to the strong internal gradients of the neutral gas density, than to a sharp transition between separate entities. The huge sizes would also have been able to reconcile a larger mass density of the intergalactic medium with the observed cloud parameters, whereas pressure confined clouds would contain only a small fraction of all baryons. When proposed, the model met with scepticism because the large sizes appeared to contradict the scant observational evidence. After the recent estimates (D'Odorico et al. 1998b) the large dimensions are no longer a problem. However, also this model meet with some difficulty in reproducing the column density distribution (Petitjean et al. 1993a).

6.1.3 Gravitational Confinement: Cold Dark Matter Minihaloes

The first theory of gravitational structure formation to address the Lyα forest phenomenon was the hot dark matter (HDM) model. Zeldovich (1970) suggested the formation of adiabatic pancakes expected to produce a primordial gas phase with the right properties for detection in H I absorption (Doroshkevich & Shandarin 1977). The predicted size for the pancakes is too large. Even after fragmentation, coherent absorption should be extending over Mpcs across the sky (Doroshkevich & Mücket 1985).

The underlying HDM structure formation scenario has become somewhat unpopular, but the physical idea of Lyα absorbers as flattened pancakes survives into the currently favored, CDM based picture.

The properties of gas clouds under the influence of the gravitational field of dark matter have been investigated by Umemura and Ikeuchi (1985), and, more specifically in terms of the "minihalo" model by Rees (1986) and Ikeuchi (1986). In this picture, Lyα clouds are a natural by-product of the CDM structure forma-

tion scenario. Photoionised gas settles in the potential well of an isothermal CDM halo. The gas is stably confined, if the potential is sufficiently shallow to avoid gravitational collapse but deep enough to prevent the warm gas from escaping. The CDM minihalos are more compact than the self-gravitating baryonic clouds of Black (1981) because of the larger dark matter gravity. The minihalo model has the attractive feature of reproducing the overall shape of the observed column density distribution function (CDDF). If the baryon distribution is an isothermal sphere ($n_b \propto r^{-2}$), the H I density in the highly ionised region of the minihalo drops like $n_{\rm HI} \propto r^{-4}$ with radius r, and the resulting column density distribution seen by random lines of sight through a population of such halos obeys $d\mathcal{N}/dN_{\rm HI} \propto N_{\rm HI}^{-1.5}$ (Rees 1988; Milgrom 1988). The largest column densities, including damped Lyα systems, are caused by the neutral cores in the shielded centers of the clouds (Murakami & Ikeuchi 1990). Thus minihalos can produce a column density power law over almost nine decades, with the correct slope, providing a physical basis for Tytler's (1987a) suggestion of a common origin for all QSO absorbers.

A non-stationary version of the minihalo model was studied by Bond, Szalay and Silk (1988), who examined the hydrodynamics of a collapsing spherical top-hat perturbation. The absorption lines could arise in gas around the dark matter halos predicted by the CDM model. This gas could either be infalling and contracting if the halo was sufficiently massive, or expanding after having been reheated during reionisation (see also Murakami & Ikeuchi 1993).

Charlton et al. (1993,1994) modelled the Lyα clouds as slabs of hydrogen with an ionising extragalactic radiation field incident from both sides. The equilibrium configuration is determined by the combined action of selfgravity, the pull of CDM gravity and an external pressure. It was found that the change of slope in the column density distribution function (near $\log N_{\rm HI} \sim 15$) can be explained by a transition between pressure and gravitational confinement, in the sense that at higher column densities gravity takes over and imposes a steeper dependence of the neutral column density with total column density.

Other recent suggestions include the possible association of Lyα clouds with the debris from merging satellites (Wang 1993; Morris & van den Bergh 1994), galactic outflows (Fransson & Epstein 1982; Wang 1995) or primordial, linear

density fluctuations in the intergalactic medium (Bi 1993).

However, the degree of realism with which spherically collapsed dark matter "halos" can describe the observed properties of gas clouds is limited. Hydrodynamic simulations (see below) show that in a hierarchical universe at intermediate redshift (~ 2) most baryonic matter may not have settled in spherical, or rotationally supported virialised objects, as suggested by the word "halo". Virial radii of objects capable of stably confining H I clouds are $\sim$ 10 kpc (Rees 1986). The coherence lengths of Lyα systems from gravitational lensing constraints (Smette et al. 1992,1995) are much larger, implying that only few lines of sight ever hit the virialised region. Thus, Lyα absorbers may still be accreting matter at the epoch when we observe them. A way to test this possibility is to look for signs of non-equilibrium, especially departures from thermal line profile caused by the bulk motion of the infalling gas (Miralda-Escudé & Rees 1993). Meiksin (1994) has traced the formation and internal structure of minihalos and slabs with hydrodynamical simulations to search for such observable non-equilibrium effects. For a slab or pancake geometry, noticeable deviations from Voigt profiles are predicted, but they would be hard to detect for spherical clouds.

6.2 Simulations and the New Scenario

A theory of structure formation, in general, has to specify the spectrum of primordial density fluctuations, the nature and abundance of dark matter and the cosmological parameters, H_0, $\Omega_{\rm m}$, $\Omega_{\rm b}$, Λ. However, there is a large gap between a theory specified at this level and a set of testable predictions. The power of numerical simulations is that, starting from theoretically specified initial conditions, they evolve them in time and show whether and how a theoretical model can produce cosmological objects like galaxies.

Turning the attention to the diffuse intergalactic medium (IGM), the crucial physics that a simulation must incorporate to model this component correctly, in addition to gravity, is photoionisation by the UV background radiation.

Starting from the qualitative model described above (Ikeuchi 1986; Rees 1986; Bond, Szalay & Silk 1988), Cen et al. (1994; see also Miralda-Escudé et al. 1996) performed a high-resolution hydrodynamical Eulerian simulation in a CDM + Λ

model. The advantage of this approach is that it is not necessary to impose a priori any physical constraints on the clouds or on their environments. The remarkable result found is that the growth of structure on subgalactic scales by gravitational collapse provides a natural, ab initio explanation for the origin of the observed Lyα forest.

Afterwards, a variety of approaches has been adopted. A standard CDM model has been studied by Zhang et al. (1995, 1997) with an Eulerian code, and by Hernquist et al. (1996), with a Lagrangian, Smoothed Particle Hydrodynamics (SPH). As a general rule of thumb, the Eulerian codes work on a fixed mesh basis, thus they are capable of higher resolution in the void regions producing the lowest column density Lyα forest. Whereas in the Lagrangian codes the gas is represented by particles. The overall properties, density and pressure gradient, are computed by averaging over neighbouring particles. As a consequence, this codes are superior to describe regions where a larger dynamical range is required. Their use has been extended to study damped Lyα systems (Katz et al. 1996) and metal absorption systems (Haehnelt, Steinmetz & Rauch 1996; Rauch, Haehnelt & Steinmetz 1997).

6.2.1 The Nature of Lyα Absorbers

Inspite of some quantitative differences, all these studies agree on a generic picture of the Lyα forest (see Fig. 6.1). The high column density systems ($N_{\rm HI} > 10^{15}$ cm^{-2}) are those that most resemble the clouds envisaged by the minihalo model. They generally have roundish or elongated shapes and correspond to highly overdense structures ($\rho_{\rm b}/\bar{\rho}_{\rm b} > 10$). The intermediate range of column density absorptions (10^{13} cm^{-2} $< N_{\rm HI} < 10^{15}$ cm^{-2}) is produced by systems characterized by an assortment of scales and in various stages of gravitational infall and collapse. Column density contours of $N_{\rm HI} \sim 10^{14}$ cm^{-2} trace filamentary structures extending continously and with relatively constant thickness ($\sim 40 - 100$ kpc proper) over Mpc distances. Lower column density systems ($N_{\rm HI} \lesssim 10^{14}$ cm^{-2}) are associated with sheet-like structures, not unlike small versions (length scale $\sim$ a few hundred kpc to 1 Mpc proper) of Zeldovich pancakes. Gas accretes through weak shocks (creating a double humped temperature profile), and settles in a dense,

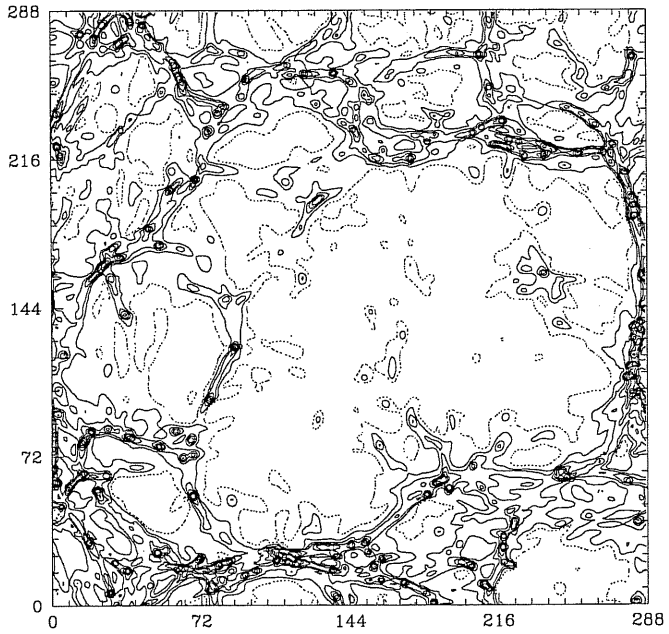

Figure 6.1: H I column density for a slice of the 10 h^{-1} Mpc (comoving) box, averaged over 48 cells in thickness, from the Λ CDM simulation by Miralda-Escudé et al. (1996). The contours are $N_{\rm HI} = 10^{13.2+0.5i}$ cm^{-2}, $i = 1$ shown dotted and $i = 2, 3, \cdot$ shown solid. The visual appearance of the low column density, sheetlike-filamentary structure has been aptly described as a "Cosmic Web" (Bond & Wadsley 1997). Reproduced from Miralda-Escudé et al. (1996)

central cooling layer, presumably to form stars. At the lowest column densities gas remains unshocked and just bounces back because of the hydrostatic pressure. The gas is partly confined by gravity and partly by ram-pressure.

Absorption features optically thin at the line center, for which $N_{\rm HI} < 10^{13}$ cm^{-2}, tend to arise in minivoids, where both the baryonic and dark matter components are underdense relative to the cosmic average density. Thus they appear not to have formed from a Jeans instability. Meiksin (1997) suggested an interpretation of these features as local maxima in an underdense region. The perturbation would grow relative to the diminishing local background and then freeze, retaining its integrity as a discrete entity.

The flow pattern of the structures is one of outflow from the voids, which

expand with a substantial fraction of the Hubble velocity. Then there is a compression into sheets at the boundaries of the voids, and a resulting flow along the sheets toward their intersections. Here the gas is cooling and collapsing, with velocity dispersions typical of dwarf galaxies at present ($v \sim 50 - 100$ km s^{-1}). These are the densest regions that could be responsible for the damped absorption systems, and could later merge to form present-day galaxies (see Katz et al. 1996).

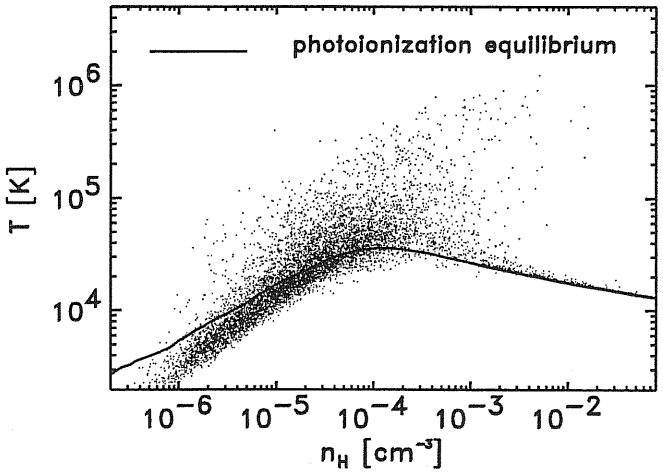

Figure 6.2: A typical density-temperature ($n-T$) scatter plot of the Lyα forest at redshift 3.1 from an SPH simulation of a standard CDM universe (Haehnelt, Rauch & Steinmetz 1996; Haehnelt, Steinmetz & Rauch 1996). Each dot represents the mean values of the total hydrogen density n and the gas temperature T along a random line of sight through the simulated box. The solid curve gives the locus of thermal photoionisation equilibrium. Around the peak of the equilibrium curve, the temperature is generally larger than the equilibrium value because of compressional and shock heating during gravitational collapse. While the modest deviation from the equilibrium curve toward smaller temperatures is due to adiabatic cooling at the lowest densities. Reproduced from Haehnelt, Steinmetz and Rauch (1996)

A typical density-temperature diagram (Figure 6.2) for random lines of sight through an SPH simulation (Haehnelt, Rauch & Steinmetz 1996) reveals significant departures from thermal photoionisation equilibrium for all but the highest density gas ($n_{\rm H} < 10^{-3}$ cm^{-3}). The temperature-density relation is generally

steeper than the equilibrium curve, because the lower density gas cools by expansion, while the gas in the density range $n_{\rm H} \sim$ a few times $10^{-5} - 10^{-3}$ cm^{-3} is heated by adiabatic compression or shock heating. Temperatures below 10^4 K occur in voids where the expansion velocity is larger.

6.2.2 The Large Baryonic Content

Regardless of the details and uncertainties in the individual simulations, it is clear that the history of galaxy formation is closely related to the Lyα absorption systems. Any gas which has not been incorporated into galaxies at high redshift should be in the Lyα clouds, and different theories on the rate at which gas can cool and form galaxies would lead to different predictions on the number of lines of different column densities.

To calculate the baryon content of Lyα clouds we need to know the ionisation correction, as most of the gas is highly ionised. For a given ionising radiation field the degree of ionisation depends on the density and thus, for a given observed column density, on the spatial extent of the gas. Deriving the mass content then requires fixing the size of the clouds either from measurement, or from theoretical prejudices. For example, the fraction of mass required to cause the observed amount of absorption can be quite large for gravitationally confined, extended, baryonic clouds (Black 1981). In contrast, the small baryon content expected if the Lyα were caused by pressure confined clouds (Sargent et al.1980) is largely a result of the small cloud sizes adopted. By varying the size and the filling factor of the absorbers, the Lyα forest can be made to contain anything from a negligible fraction up to virtually all of the baryons, and still be consistent with the observations (Meiksin & Madau 1993). Specifically for the CDM minihalo model, Petitjean et al. (1993b) found that the Lyα forest clouds had to contain most of the baryons at redshift 2-3, in order to match the observed column density distribution function. This is in agreement with Shapiro, Giroux and Babul (1994), who found that in a CDM model the fraction of baryons not yet collapsed into galaxies should be of the order of 50–90% .

Hydrodynamical simulations predict that 80–90 % of the baryons in the universe are in the Lyα clouds, with practically no change from $z = 4$ to $z = 2$.

In particular the majority of them are in clouds with column densities $\sim 10^{14}$ cm^{-2}(Miralda-Escudé et al. 1996). This result is not in contrast with the well known fact that most of the *neutral* gas is observed to be in very high column density systems (the damped Lyα systems associated with protogalaxies) having $N_{HI} > 10^{20}$ cm^{-2}. The reason is simply the high state of ionisation of most of the gas in the low column density systems and the increasing neutral fraction as a function of column density.

Independent of the cosmological model, the large transverse sizes of Lyα absorbers measured from QSO pairs give another, indirect indication that the baryon density in Lyα clouds must be large, or the absorbers must be extremely flattened (Rauch & Haehnelt 1995).

6.3 The Fluctuating Gunn-Peterson Effect and the Semi-analytical Approach

Since the discovery of the Lyα forest, there have been several attempts to detect a Gunn-Peterson (GP) effect due to the presence of a uniform intercloud medium, as distinct from individual clouds (Steidel & Sargent 1987; Jenkins & Ostriker 1991; Webb et al. 1992; Giallongo et al. 1993, 1994). However, there does not seem to be any theoretical justification for such a distinction for the case of a photoionised intergalactic medium. The galaxies and clusters we observe today must have collapsed from density fluctuations which were already substantial even at the highest redshifts of observed quasars. Moreover, in any theory of hierarchical gravitational collapse, fluctuations on subgalactic scales are even larger and collapse at an earlier epoch. Thus, a uniform intergalactic medium cannot exists and the Lyα (or GP) optical depth should vary even at the lowest column densities (Black 1981; McGill 1990; Bi, Börner & Chu 1992; Miralda-Escudé & Rees 1993; Reisenegger & Miralda-Escudé 1995).

In a CDM dominated structure formation scenario the accumulation of matter in overdense regions reduces the optical depth for Lyα absorption considerably below the average in most of the volume of the universe. Because the pressure of the photoionised gas prevents it from collapsing on scales smaller than the Jeans

scale, an intergalactic medium pervading all space must be present, and the inhomogeneous distribution of gas in voids must yield a fluctuating GP effect (see Eq. (4.16)). Given this, the traditional searches for the GP effect that try to measure the amount of matter between the absorption lines can easily overestimate the amount of ionising radiation necessary to keep the neutral hydrogen GP absorption below current detection limits.

As another consequence, the distinction between the low column density Lyα forest "lines", and the GP "trough", becomes somewhat artificial. Bi and collaborators (Bi, Börner & Chu1992; Bi 1993; Bi & Davidsen 1997) have shown that the optical depth fluctuations corresponding to the linear regime of gravitational collapse in the intergalactic medium can give a remarkably realistic representation of the Lyα forest (ignoring the higher column density lines, which are produced from non-linear structures, e.g., minihalo type objects). Their semi-analytical work is based on a log-normal density fluctuation field. For low densities where dissipation is not important the collapse of dark matter and baryons differs mainly by the presence of the gas pressure which effectively smooths the baryons distribution on scales below the Jeans length. Bi et al. treated the pressure as a modification to the power spectrum of the baryon density contrast $\delta_{\rm b}$, suppressing power on scales below the Jeans length:

$$\delta_{\rm b}(k) = \frac{\delta_{\rm DM}(k)}{1 + k^2(\lambda_{\rm J}/2\pi)^2} \tag{6.2}$$

where $\lambda_{\rm J}$ is the Jeans length, k the wavenumber, and $\delta_{\rm DM}$ the dark matter overdensity. This method can elucidate many of the basic features of low column density Lyα clouds. The schematic treatment of the equation of state and the lack of inclusion of shock heating limit the approach, however, to overdensities of $\delta < 5$, where gas physics beyond the Jeans criterion is not very important.

Gnedin and Hui (1996) and Hui, Gnedin and Zhang (1997) led a similar analysis for the low column density Lyα forest making use of the truncated Zeldovich approximation to compute the density and peculiar velocity fields. They adopted a density-peak *Ansatz*: each density peak in the real space is associated with an absorption line whose column density is assigned on the basis of the height and curvature of the peak. The column density distribution $f(N)$ is then a statistics of density peaks in real space. They found that the amplitude and slope of the

linear power spectrum on comoving scales of around 10–20 h^{-1} Mpc are the most important factors in determining the slope of the column density distribution. Decreasing the amplitude and/or steepening the slope of the power spectrum tends to steepen the distribution in the column density range $\sim 10^{12.5} - 10^{14.5}$ cm^{-2}. Changes in the equation of state $T = T(\rho)$ have an additional, but smaller impact. The overall normalization of $f(N)$ is given by $(\Omega_b h^2)^2/\Gamma$ where Γ is the photoionisation rate; a change in this quantity shifts $f(N)$ horizontally to larger N.

Furthermore, in § 4.5, we exhaustively spoke about the techniques, based on the equation for the optical depth due to a fluctuating GP effect (4.16), for recovering the initial power spectrum of density fluctuations directly from the fluctuations of the optical depth (Croft et al. 1998; Hui 1999; Nusser & Haehnelt 1999; McDonald & Miralda-Escudé 1999).

For a more accurate "semi-numerical approach", much cheaper than a high resolution hydrodynamic calculation, one can run a lower resolution N-body simulation, compute the density field from the particle distribution, impose the temperature-density relation, and extract spectra. This tecnhique uses a fully non-linear solution for the density and velocity fields, but it still assumes that gas traces dark matter and that all gas lies on the temperature-density relation, approximations that are good but not perfect. Similar techniques have been used by Gnedin and Hui (1998), who also incorporate an approximate treatment of gas pressure in the N-body calculation by modifying the gravitational potential, and by Petitjean, Mücket and Kates (1995), Mücket et al. (1996) and Riediger, Petitjean and Mücket (1998), who use a more elaborate method to compute gas temperatures.

Although not "hydrodynamically correct" the semi-analytical techniques present a major advantage over numerical simulations: They can be easily rerun testing the effect of changing the input parameters and/or they can be evolved for a long time span, e.g., down to redshift $z = 0$.

6.4 Modelling the Metal Absorbers at High Redshift

Bahcall & Spitzer (1969) and Burbidge et al. (1977) pointed out that to explain observations of metal absorption lines the cross section of absorbing gas would have to be much larger than the typical half-light radii of present-day spiral galaxies if these objects had the same comoving space density as such galaxies. First observational evidence in favor of large metal absorption cross-sections came from the work by Bergeron (1986) who discovered low redshift galaxies close to the line of sight to QSOs with known Mg II absorption systems, coincident in redshift with the absorbers. Surveys of optically thick (Lyman Limit) absorption systems and galaxies at intermediate redshift have reported results consistent with galaxies being surrounded by Mg II absorbing halos of radius $\sim$ 40 kpc (e.g. Bergeron 1995; Steidel 1995; Churchill, Steidel & Vogt 1996). Similarly, damped Lyα systems were interpreted by Wolfe and collaborators (e.g. Wolfe 1988; Lanzetta et al. 1991) as large protodisks, progenitors of present-day spiral galaxies with significantly larger radii at high redshift.

The large-halo/disk scenario can qualitatively explain the component and internal structure of heavy element absorption systems, especially the strong clustering measured for C IV systems (see Chapter 4). In this picture the individual absorption components could be clouds orbiting in a halo or co-rotating in a disk, produced and replenished e.g. by thermal instability (Bahcall 1975; Fall & Rees 1985; Mo 1994; Mo & Miralda-Escudé 1996).

Unambiguous evidence in favor of large absorption cross-sections and the identification of the absorbing objects with massive galaxies is restricted to low redshift observations. Massive disks are not the only viable dynamical model for the observed velocity structure in damped Lyα systems (Haehnelt, Steinmetz & Rauch 1998; Ledoux et al. 1998). Moreover, observations of galaxies at high redshift are strongly biased towards objects with high current star formation rates. Thus it is not clear whether the transverse separations on the sky of galaxy-absorber pairs coincident in redshift do reliably indicate the presence and sizes of any (hypothetical) halos/disks. It is difficult to ascertain that there is not an undetected (not necessarily faint) galaxy closer to the line of sight.

The possibility of "smaller" but more numerous objects as sources of the metal

absorption has been explored earlier. Tyson (1988) suggested identifying damped Lyα systems with gas-rich dwarf galaxies instead of large proto-disks. York et al. (1986) discussed clusters of such objects to explain the component structure of CIV systems. Individual galaxy halos cannot produce potential wells deep enough to explain the largest velocity splittings of C IV systems (up to 1000 km s^{-1}) as virialised motions. Pettini et al. (1983) and Morris et al. (1986) concluded that the velocity splitting and large cross-sections are equally difficult to understand if objects similar to present-day galaxy clusters placed at high redshift were causing the absorption.

Earlier attempts at understanding QSO absorption systems have mostly employed heuristic models. The ionisation state of the gas was calculated with simplified assumptions about the geometry, temperature and dynamical structure of the gas. In the new picture traced by realistic hydrodynamical simulations of the fate of gas in a universe subject to hierarchical structure formation (see § 6.2), a coherent filamentary large scale structure of dark matter and baryons is responsible for Lyα forest absorption lines. Denser condensations embedded in the filaments are unstable against rapid collapse and cooling of the gas and probably form stars at their centers.

In such a hierarchical picture metal absorption systems at high redshift are more likely to arise from groups of relatively small, merging protogalactic objects, rather than from ensembles of clouds in huge virialised halos (Haehnelt, Steinmetz & Rauch 1996; Rauch, Haehnelt & Steinmetz 1997). This hypothesis is asserted by observational and theoretical arguments:

1. Carbon is found to be associated to more than half of the high redshift Lyα lines with $N_{\rm HI} \sim$ a few $\times 10^{14}$cm^{-2} (Cowie et al. 1995; Tytler et al. 1995). In the numerical simulations such column densities typically correspond to gas densities smaller than those expected for fully collapsed objects at these redshifts (indicating baryonic overdensities of order ten), so at least some metal absorption systems appear to occur outside virialised regions.
2. Typical observed temperatures of C IV systems are somewhat larger than expected if the gas were heated only by photoionisation (Rauch et al. 1996). Such an enhancement is predicted by numerical simulations (Haehnelt, Steinmetz & Rauch 1996, see Fig. 6.2) and is most likely due to shock and (to a smaller ex-

tent) adiabatic heating during the gravitational compression of the gas. Absorber sizes along the line of sight inferred from ionisation balance calculations are of order ten kpc. This is uncomfortably large for a cloudlet-in-halo model (Haehnelt, Rauch, & Steinmetz 1996).

3. Simple scaling laws predict that in a hierarchical scenario the typical ratio of cooling time to dynamical time decreases as $(1+z)^{-3/2}$. At high redshift the gas will generally cool rapidly out to the virial radius. It will be difficult to maintain large, massive, hot halos for an extended period of time. Similarly for fixed circular velocity both the mass and the size of the dark matter component of typical objects at high redshift decrease as $(1+z)^{-3/2}$ (see Kauffmann 1996; Steinmetz 1997).

Rauch, Haehnelt and Steinmetz (1997, see also Haehnelt, Steinmetz & Rauch 1996) have carried out a SPH hydrodynamical simulation to investigate the spatial distribution and absorption properties of metal-enriched gas in regions of ongoing galaxy formation. They make the simple hypothesis that gas is homogeneously contaminated by metals (C, Si, N, O). Different ionic species are found to probe different density and temperature regimes. In the diffuse gas which gives rise to the lowest H I column densities (10^{12} to 10^{14} cm^{-2}) high ionisation species are prevalent and OVI λ103.1 nm is often the strongest metal absorption line. Towards higher H I column densities, which probes filaments embedded in the large-scale sheets, low column density C IV $\lambda\lambda$154.8, 155.0 nm lines from infalling gas with densities around 10^{-4} cm^{-3} (overdensities of about 10 to 100) dominate the metal absorption features. C IV remains the most easily visible metal ion in the as yet unvirialised regions around the protogalactic clumps which are later to merge into present-day galaxies. Still larger H I column densities occur for lines of sight approaching the central regions of the protogalactic clumps. Total densities here exceed 10^{-4} cm^{-3}, and species like C II and Si IV become increasingly prominent. At densities above 10^{-3} cm^{-3} we have reached the virialised region which is generally optically thick for radiation shortward of 1 Rydberg.

6.5 Comparison with the Observations

The simulations have been quite successful in matching the overall observed properties of the absorption systems, and the agreement ranges from the acceptable to the amazing.

The shape of the column density distribution and the Doppler parameter distribution are reasonably well reproduced by the simulations (Cen et al. 1994; Zhang et al. 1995,1997; Hernquist et al. 1996; Miralda-Escudé et al. 1996). Although the approximate range of Doppler parameters is hard to miss (with photoionisation being the great equaliser), subtle effects can raise or lower the mean line width by $\sim 30\%$ and change the shape of the Doppler parameter distribution. There may be some discrepancy for the Doppler parameters between different simulations (Zhang et al. 1997; Davé et al. 1997; Miralda-Escudé et al. 1996) but it is not yet clear whether this is due to different types of data analysis, different assumptions about the process of reionisation, or limited numerical resolution. A departure from Voigt profile shapes, especially the broad wings of weak lines signifying bulk motion broadening in sheets, is seen in the simulations (Cen et al. 1994) and appears to be present in real high resolution spectra (Rauch 1996).

Theuns et al. (1999) investigated the properties of the simulated Lyα forest in three popular cosmologies (critical, open and vacuum-dominated). All the models produce column density distributions which are well matched to the observations. On the other hand, they found that the observed Doppler parameter distribution is better reproduced by models with a higher baryon density $\Omega_{\rm b}h^2$ and a lower matter density. These models are hotter and have higher peculiar velocities shifting the b-parameter distribution towards larger b values.

The large transverse sizes of the absorbers seen against background QSO pairs and lensed QSOs are readily explained by the coherence length of the sheets and filaments (Miralda-Escudé et al. 1996; Charlton et al. 1997; Cen & Simcoe 1997). The weak clustering amplitude appears to be in agreement with the observations. The histogram of residual fluxes in the Lyα forest is reproduced very well by the models (Rauch et al. 1997).

The computed He II opacity requires a break in the standard UV background intensity (Haardt & Madau 1996) between the H I and He II photoelectric edges

by a factor of 100–150 to match the observations (Miralda-Escudé et al. 1996; Zhang et al. 1995, 1997). This is consistent with the break required to match the amount of Si and C absorption measured in the Lyα forest (Songaila & Cowie 1996; Savaglio et al. 1997; D'Odorico et al. 1998a). The observed column density ratios of the different ionic species at $z = 3$ are well reproduced if a carbon abudance [C/H] $= -2.5$, relative abundances as found in present-day galactic-halo stars, a UV background intensity at the Lyman limit $J_{-21} = 0.3$ and either a power-law spectrum ($J \propto \nu^{-1.5}$) or the Haardt & Madau (1996) spectrum are assumed (Rauch, Haehnelt & Steinmetz 1997).

The evolution of the Lyα forest with time at high redshift is mainly driven by the Hubble expansion and the resulting increase in the mean ionisation of the gas, and to a lesser degree by the gas streaming along the filaments (Miralda-Escudé et al. 1996). Riediger et al. (1998, see also Mücket et al. 1996), from their simulation, find that the number of absorbers per redshift is given by a broken power law, with $\gamma \sim 2.6$ $(1.5 < z < 3)$ and $\gamma \sim 0.6$ $(0 < z < 1.5)$, $(\log N_{\rm HI} > 14)$, a remarkable agreement with the observed data. The break in the power law can be understood as a change with time in the dimensionality of the structures dominating the absorption. The sheetlike absorbers dominating the high redshift Lyα forest are expanding with time and are dropping below the detection threshold first because of their low column density, leaving the absorption from the less rapidly evolving gas distribution in the filaments and knots to dominate. There the column density also decreases, but since the original column was higher, the filaments remain visible for longer. Continuing infall also contributes to the increasing prominence of the more compact structures. Theuns, Leonard and Efstathiou (1998) find an excellent agreement between their simulated redshift evolution of the number of Lyα lines and observations. They explain the observed decrease in the rate of evolution at $z \lesssim 2$ by the steep decline in the photoionising background resulting from the rapid decline in quasar numbers at low redshift.

6.6 Conclusions

The new paradigm for the QSO absorption systems has considerable explanatory power, but that does not mean that it is correct. The interpretation of the absorption

systems, and the cosmological measurements planned or already performed to date depend on gravitational collapse as the dominant source of structure in the intergalactic medium. Even if the hierarchical models are basically correct, it is conceivable that local physical effects may upset some of the cosmological conclusions. A fluctuating radiation field may be a source of non-gravitational structure in the forest, as may be stellar feedback. Metal enrichment has been found to be common for absorption systems with H I column densities as low as $\log N_{\rm HI} \sim 14$ (Tytler et al.1995; Cowie et al.1995). If this is not due to a very early phase of metal enrichment we have to worry that some process other than gravitational collapse may have formed the metal enriched Lyα clouds. The origin of the ionising radiation, the epoch of reheating and the spectral slope are others sources of uncertainty. On the technical side: Do hydrodynamic codes already converge, or how much do the inferred cosmological parameters (Ω_b, the amount of small scale structure present) depend on the resolution and size of the simulations, and the numerical technique ?

Finally, the cosmological picture itself could be wrong, and the interpretation of the forest as absorption mostly due to the intergalactic medium (as opposed to distinct galaxies) may be doubtful. Galaxy halos or disks could be big/numerous enough to produce the low column density Lyα forest as well. The relationship between Lyα absorbers and galaxies at low redshift have received much attention recently (Morris et al. 1991; Salzer 1992; Morris et al. 1993; Salpeter & Hoffman 1995; Lanzetta et al. 1995; Le Brun, Bergeron & Boissé 1996; Bowen, Blades & Pettini 1996; Le Brun & Bergeron 1998). At the moment results are still controversial and no firm conclusion can be drawn from them.

Chapter 7

Conclusions and Perspectives

The formation and evolution of structures in the Universe is one of the central issues in modern cosmology. This study probes the abundance and form of dark matter, the mean baryonic density, the turnover scale in the perturbation power spectrum, and the formation processes of galaxies and clusters, that is, it provides a useful tool to discriminate between the various cosmological models.

On the observational side, surveys of galaxies have always been the common way to investigate the distribution properties of matter. Larger and larger surveys (see e.g., Da Costa 1998 for a review) composed a scenario of the local universe as a closely packed network of voids ~ 5000 km s^{-1} in diameter bounded by thin, large walls.

The advent of 8-10m class telescopes and the use of new photometric techniques (see e.g., Steidel, Pettini, & Hamilton 1995) have allowed to detect galaxies in emission up to redshift $z \sim 4$, the so-called Lyman break galaxies (Steidel et al. 1996; Giavalisco et al. 1998; Adelberger et al. 1998; Steidel et al. 1999). The most striking and robust result at present is that these star-forming objects must be much more clustered than the overall mass distribution on scales of $\sim 10\ h^{-1}$ Mpc, for any reasonable cosmological model. As a consequence, they are thought to be the progenitors of objects that end up in groups and clusters of galaxies by the present day (Governato et al. 1998). Most excitingly, the UV luminosity of these objects suggests, if interpreted correctly, a star formation rate that keeps constant up to redshift $z \simeq 4$ (Steidel et al. 1999).

A complementary probe of the high redshift universe is represented by QSO

absorption line systems. By definition, they trace the sites along the line of sight rich in gas, where most of the baryons resides at redshift $z \sim 2-3$. They present some appealing qualities. The complete range of redshift from zero to the highest QSO redshift, $z = 4.897$ (Schneider, Schmidt & Gunn 1991), is accessible. On the one hand, column densities as large as $\log N_{\rm HI} \sim 22$ are observed which reveal most certainly proto-galactic disks. On the other hand, due to the large absorption cross-section of the neutral hydrogen Lyα transition, the method is sensitive to column densities as low as $\log N_{\rm HI} \sim 12$ which arise in tenuous intergalactic gas.

In this thesis, I wanted to emphasize the potentialities and the results obtained from the investigation of QSO absorption spectra. Since the main aim was to explain my original contribution to the subject, I skipped or touched only marginally some very interesting topics I did not face in the course of my PhD work. In particular, I almost did not talk about Damped Lyman α systems and the element abundance estimate at high redshift (see e.g., Pettini 1999 for a recent review). Furthermore, I did not treat other applications of absorption lines like the remarkably militant debate about the deuterium to hydrogen ratio, the process of reionisation and the associated systems ($z_{\rm abs} \sim z_{\rm em}$).

My approach to the study of QSO absorption spectra has been mainly observational. I reduced and in some cases, directly obtained at the telescope, rough data, facing all the problems related to the optimal way to transform them in the final spectrum: the maximisation of signal-to-noise ratio, a reliable estimate of the continuum and an effective procedure to detect, identify and fit lines. These issues are briefly treated in § 3.7, where I reported the analysis of the high resolution spectrum of the high redshift QSO PKS 2126–158 (D'Odorico et al. 1998a).

The main issues I faced in the course of my PhD studies are summarised below.

The Doppler parameter and column density distributions

We investigated the statistical properties of the Lyα lines in the forest of PKS 2126–158 and of an enlarged sample of C IV doublets, obtained by merging the lines of this QSO with those of two other quasars observed at comparable resolution. Our results as regards the distribution of column densities and Doppler widths are in agreement with previous papers and with the Keck findings for the

Lyα forest lines (Hu et al. 1995). We could also carry out an investigation of the variations of line fitting parameters with resolution and signal-to-noise ratio by comparing our results with those of Giallongo et al. (1993). Eventually, we did not confirm the correlation between b and $N_{\rm HI}$ found by Pettini et al. (1990).

The spectral shape of the UV background flux

Two of the metal systems found in the spectrum PKS 2126–158 exhibit a well defined H I column density and were used to estimate the metallicity of the gas at $z \sim 3$. In order to make the column densities compatible with realistic assumptions about the cloud sizes and the silicon to carbon overabundance, it was necessary to assume a spectrum of the UV ionising background at $z \sim 3$ with a jump beyond the He II edge a factor of 10 larger than the standard predictions for the integrated quasar contribution (Haardt & Madau 1996). This could be originated by a contribution of primeval galaxies. The mean metallicity is found to be ~ -2.5 dex solar.

The clustering properties of absorbers

My reasearch interest has been mainly devoted to the study of the clustering properties of absorbers.

The predicted clustering properties of matter at high redshifts show detectable differences in different cosmological models. A comparison with suitable observational data could, thus, constrain the viability of such theoretical models. The information obtained from high redshift galaxies is strongly biased by two basic problems: the relation between luminosity and mass is still poorly known and we are selecting the bright end of the luminosity function for this class of objects. This thicken the relationship between the theoretical matter correlation function and the measured correlation function of objects. In the case of QSO absorption systems it is easier to relate observations to the underlying matter density field (see § 4.5), but on the other hand, the intrinsic lack of tridimenionality in QSO spectral data still requires the development of new statistical technique to compare them to the predicted results.

An historical review and all our results are reported in Chapter 4.

One of the principal characteristic of the Lyα forest population established by Sargent et al. (1980) was the Poissonian distribution of absorbers, that suggested their alternative nature with respect to galaxies. The truthfulness of this property was debated between confirmation and marginally significant detection of clustering on small scales. We compiled a statistical sample of 1600 Lyα lines in 15 quasars observed at high resolution ($\sim$ 10 km s^{-1}) in the framework of an ESO key programme with the addition of literature data. The two point correlation function computed along the line of sight in velocity space shows a significant signal on small velocity scales ($\Delta v \lesssim 300$ km s^{-1}) with amplitude and significance increasing with increasing column density. We found an apparent trend of increasing correlation with decreasing redshift and a continuity of properties between Lyα forest absorbers and metal absorbers suggesting physical association between the Lyα clouds with $\log N_{\rm HI} \gtrsim 14$ and the haloes of protogalactic systems (Cristiani et al. 1997).

We estimated also the clustering properties of C IV metal lines by considering a sample of 71 doublets at redshifts $2.3 \lesssim z \lesssim 4$ (D'Odorico et al. 1998a). A significant signal is found for scales smaller than $200 - 300$ km s^{-1} in agreement with previous results at lower resolution.

The single sightline approach to the estimate of the correlation among absorbers presents some drawbacks that cannot be neglected: purely internal velocity motions distort the redshift space possibly adding spurious signal to the measured clustering, and lines at quite different redshifts are mixed together. A way to minimise these effects is to study the cross-correlation between absorbers along adjacent lines of sight. Only few attempts have been made to compute the cross-correlation function up to now (e.g., Williger et al. 1996; Crotts, Burles & Tytler 1997). The reason is that the number of known suitable adjacent lines of sight is very small. With the advent of 10 m-class telescopes this number will increase dramatically.

We are engaged at the moment in a project to obtain additional data at the New Technology Telescope NTT on a group of 8 QSOs with separations spanning between a few arcminutes and $\sim$ 60 arcmin and two QSO pairs - 0056-35A,B and 2344+125A,B - separated by 3.7 arcmin and 5 arcmin respectively. Time has already been allocated in September and October 1999. The set of data will be ideal

to investigate the spatial correlation of C IV absorption systems in the redshift range $1.7 \lesssim z_{\rm abs} \lesssim 3.4$, on proper scales of $\sim 1\ h^{-1}$ Mpc or above. These scales are characteristic of the transition between the outskirts of individual haloes and the larger structures, filaments or sheets. Our spectra will have a resolution ~ 4 times better than in Williger et al. (1996), we expect a number of absorbers 4-5 times larger and thus an increase in the statistical significance of our result. Furthermore, we investigate a wider range of spatial scales and, in particular, smaller transverse scales thanks to the two QSO pairs.

The size and geometry of Lyα absorbers

We have exploited the 3-dimensional information enclosed in lines of sight to QSO pairs to estimate the transverse dimension of Lyα absorbers (D'Odorico et al. 1998b). Chapter 5 describes the algorithm we adopted and our results, together with a review of the similar analysis present in the literature.

We added to our new data on the QSO pair Q0307-195A,B all the spectra of QSO pairs at comparable resolution available in the literature, for a total of 8 couples. We adopted a Bayesian approach based on the statistic of the hits (i.e., when an absorption line appears in both QSO spectra) and misses (i.e., when a line is seen in any of the QSO spectra, but no line is seen in the other) (Mc Gill 1990; Fang et al. 1996). Some improvements have been applied to the standard methods in order to deal correctly with the extreme cases of very low numbers (1 or 0) of hits or misses and avoid any spurious dependence on the separation between lines of sight. In the hypothesis of single radius, unclustered, spherical or disc-like clouds we obtained very large proper radii of the order of $300 - 500\ h^{-1}$ kpc. No correlation between the typical inferred size and the redshift was detected while a correlation with the proper separation of the pair is possible, suggesting that our hypothesis are too simple to describe the real scenario. If we add the information on the equivalent width of the lines and consider clouds with a power law column density profile, a disc geometry with radii $R \simeq 100 - 200\ h^{-1}$ kpc reasonably reproduce the data.

Our results together with analysis carried out with the use of hydrodynamical simulations indicate that the adopted models are too simple to reliably compute

the size of absorbers.

Improvements in this sense can be effected on the one hand by the use of Monte Carlo simulations following the approach of Dinshaw et al. (1997, and references therein), by adopting more realistic models for the clouds. On the other hand, big hydrodynamical simulations allow to investigate the relationship absorption/absorber (almost) without the need of a defined model for the clouds. Eventually, a necessary condition is to have more and more observational data on QSO pairs and close groups.

The Hubble Deep Field South quasar

Another topic I did not touched in the course of this thesis is the relationship between galaxies and absorbers. The connection between apparently normal galaxies and strong metal absorptions has been explored directly only recently. This is done by identification in the same field as the quasar, of galaxies that coincide in redshift with absorptions seen in the spectrum (e.g. Bergeron & Boissé 1991; Steidel, Dickinson & Persson 1994). At intermediate redshift ($z \sim 0.7$), Mg II systems with equivalent width $w_{\rm r} > 0.03$ nm are primarily associated with large $R \sim 35\ h^{-1}$ kpc haloes of a population of luminous ($\sim L^*_{\rm B}$) galaxies which appear not to evolve significantly over the observed redshift range. By analogy, the C IV systems, which are observable from the ground at $1.2 \lesssim z \lesssim 5$, are thought to arise in the halo of normal galaxies.

Most of the numerous Lyα lines are thought to be of intergalactic origin at high redshift, while at low redshift the relationship with galaxies has not yet been cleared (Lanzetta et al. 1995; Le Brun, Bergeron & Boissé 1996; Le Brun et al. 1997). Very deep imaging is necessary to test whether absorption lines and in particular Lyα lines either are produced by a population of gas-rich dwarf galaxies or are of pure intergalactic origin.

At the moment, the best field in which this question can be addressed is the STIS HDF-S field, which is centered on a bright, high-redshift quasar (J2233-606, $V \sim 17.5, z_{\rm em} = 2.24$; see e.g. Savaglio et al. 1999). Indeed a very deep image of the field (without any filter) has been obtained with STIS and deep imaging in U,B,V,R,I,J,H and K is planned in the course of the ESO Imaging Survey. More-

over, Tresse et al. (1999) report identification of a new $z_{\rm em} = 1.335$ quasar with I=20.8 at a projected angular separation of $44.5''$ (or 200 h^{-1} kpc) from J2233-606. This makes the QSO pair an ideal target for QSO environment studies. The QSO J2233-606 will be one of the first targets for the UV and Visual Echelle Spectrograph (UVES) first light observations in Fall 1999 on the Very Large Telescope (VLT). This will be the best spectrum in the wavelength range $\lambda\lambda 330 - 500$ nm. In addition, we have obtained an EMMI echelle spectra of the QSO at a resolution $R \sim 7700$ in the wavelength range $\lambda\lambda 390 - 770$ nm in September 1998.

At the moment, I am working on a complex metal system in the spectrum of J2233-606 that presents lines of many different ions and of different ionisation stages of the same elements. This particularly favourable juncture allows to put tightening constraints on the shape of the UV background and of the ionising parameter (e.g. Bergeron & Stasinska 1986).

We can expect new observational facilities to enlarge the scope of absorption line studies considerably. The projected Cosmic Origins Spectrograph (COS) will increase the spectroscopic efficiency of the HST, benefitting almost all of the observational areas mentioned here. The Sloan Digital Sky Survey (SDSS) should produce large numbers of QSOs useful for spectroscopic follow-up, to study (for example) the large scale structure from intergalactic absorption in three dimensions. Even low resolution QSO spectra will be useful for cosmological purposes. And the use of large optical telescopes is inevitable for studies of gravitationally lensed QSOs, the absorber-galaxy connection at any redshift, and of course for most projects involving narrow, metal absorption lines, a subject of increasing relevance for our understanding of the process of galaxy formation.

Bibliography

Adelberger K.L., Steidel C.C., Giavalisco M., Dickinson M., Pettini M., Kellog M., 1998, ApJ 505, 18

Alcock C., Paczyński B., 1979, Nature 281, 358

Arons J., McCray R., 1969, ApJ 158, L91

Arons J., 1972, ApJ 172, 553

Atwood B., Baldwin J.A., Carswell R.F., 1985, ApJ 292, 58

Babul A., 1991, MNRAS 248, 177

Bahcall J.N., Salpeter E.E., 1965, ApJ 142, 1677

Bahcall J.N., 1968, ApJ 153, 679

Bahcall J.N., Peebles P.J.E., 1969, ApJ 156, L7

Bahcall J.N., Spitzer L., 1969, ApJ 156, L63

Bahcall J.N., 1975, ApJ 200, L1

Bahcall J.N., Bergeron J., Boksenberg A., Hartig G.F., Jannuzi B.T., et al., 1993, ApJS 87, 1

Bahcall J.N., Bergeron J., Boksenberg A., Hartig G.F., Jannuzi B.T., 1996, ApJ 457, 19

Bajtlik S., Duncan R.C., Ostriker J.P., 1988, ApJ 327, 570

Bajtlik S., 1995, in: *QSO absorption lines*, Proc. ESO Workshop, Meylan G. ed., Springer, Berlin, p. 337

Baldwin J.A., Burbidge E.M., Burbidge G.R., Hazard C., Robinson L.B., et al., 1974, ApJ 193, 513

Barcons X., Webb J.K., 1990, MNRAS 244, 30p

Black J., 1981, MNRAS 197, 553

Bechtold J., 1987, in: Proc. Third IAP Workshop, *High Redshift and Primeval Galaxies*, J. Bergeron et al. eds., Editions Frontiéres, Gif-sur-Yvette, p. 397

Bechtold J., Green R.F., York D.G., 1987, ApJ 312, 50

Bechtold J., 1994, ApJS 91, 1

Bechtold J., Crotts A.P.S., Duncan R.C., Fang Y., 1994, ApJ 437, L83

Bechtold J., 1995, in: *QSO Absorption Lines*, Proc. ESO Workshop, ed. Meylan G., Berlin, Springer, p.299

Bergeron J., Boissé P., 1984, A&A 133, 374

Bergeron J., 1986, A&A 155, L8

Bergeron J., Stasińska G., 1986, A&A 169, 1

Bergeron J., Boissé P., 1991, A&A 243, 344

Bergeron J., 1995, in: *QSO Absorption Lines*, Proc. ESO Workshop, ed. G. Meylan, Springer, Berlin, p. 127

Bi H., Börner G., Chu Y., 1989, A&A 218, 19

Bi H., Börner G., Chu Y., 1991, A&A 247, 276

Bi H., Börner G., Chu Y., 1992, A&A 266, 1

Bi H., 1993, ApJ 405, 479

Bi H., Davidsen A.F., 1997, ApJ 479, 523

Boissé P., Le Brun V., Bergeron J., Deharveng J.M., 1998, A&A 333, 841

Boksenberg A., 1972, in *Auxiliary Instrumentation for Large Telescopes*, Proc. ESO/CERN Conf., Geneva, p 295

Boksenberg A., Sargent W.L.W., 1975, ApJ 198, 31

Bond J.R., Szalay A.S., Silk J., 1988, ApJ 324, 627

Bond J.R., Wadsley J.W., 1997, in: *Computational Astrophysics*, Proc. 12th Kingston Conference, eds D. Clark, M. West, PASP conf. series, p323

Bowen D.V., Blades J., Pettini M., 1996, ApJ 464, 141

Burbidge E.M., Lynds C.R., Burbidge G.R., 1966, ApJ 144, 447

Burbidge G., O'Dell S.L., Roberts D.H., Smith H.E., 1977, ApJ 218, 33

Carlberg R.G., et al. 1998, in *Large Scale Structure in the Universe*, Royal Society Discussion Meeting, March 1998, astro-ph/9805131

Carswell R.F., Whelan J.A.J., Smith M.G., Boksenberg A., Tytler D., 1982, MNRAS 198, 91

Carswell R.F., Morton D.C., Smith M.G., Stockton A.N., Turnshek D.A., Weymann R.J., 1984, ApJ 278, 486

Carswell R.F., Rees M.J., 1987, MNRAS 224, 13

Carswell R.F., Webb J.K., Baldwin J.A., Atwood B., 1987, ApJ 319, 709

Carswell R.F., Lanzetta K.M., Parnell H.C., Webb J.K., 1991, ApJ 371, 36

Cen R., Miralda-Escudé J., Ostriker J.P., Rauch M., 1994, ApJ 437, L9

Cen R., Simcoe R.A. 1997, ApJ 483, 8

Charlton J.C., Salpeter E.E., Hogan C.J., 1993, ApJ 402, 493

Charlton J.C., Salpeter E.E., Linder S.M., 1994, ApJ 430, 29

Charlton J.C., Churchill C.W., Linder S.M., 1995, ApJ 452, 81L

Charlton J.C., Anninos P., Zhang Y., Norman M., 1997, ApJ 485, 26

Chernomordik V.V., Ozernoy L.M., 1993, ApJ 404, L5

Chernomordik V.V., 1995, ApJ 440, 431

Churchill C.W., Steidel C.C., Vogt S.S., 1996, ApJ 471, 164

Cooke A.J., Espey B., Carswell R.F., 1997, MNRAS 284, 552

Courbin F., Lidman C., Magain P., 1998, A&A 330, 57

Cowie L.L., Songaila A., Kim T.-S., Hu E.M., 1995, AJ 109, 1522

Cristiani S., D'Odorico S., Fontana A., Giallongo E., Savaglio S., 1995, MNRAS 273, 1016

Cristiani S., D'Odorico S., D'Odorico V., Fontana A., Giallongo E., Savaglio S., 1997, MNRAS, 285, 209

Croft R.A., Weinberg D.H., Katz N., Hernquist L., 1998, ApJ 495, 44

Croft R.A., Weinberg D.H., Pettini M., Katz N., Hernquist L., 1999, ApJ 520, 1

Crotts A.P.S., 1985, ApJ 298, 732

Crotts A.P.S., 1987, MNRAS 228, 41

Crotts A.P.S., 1989, ApJ 336, 550

Crotts A.P.S., Bechtold J., Fang Y., Duncan R.C., 1994, ApJ 437, L79

Crotts A.P.S., Burles S., Tytler D., 1997, ApJ 489, L7

Crotts A.P.S., Fang Y., 1998, ApJ 502, 16

Davè R., Hernquist L., Weinberg D.H., Katz N., 1997, ApJ 477, 21

Davidsen A.F., Kriss G.A., Zheng W., 1996, Nature 380, 47

Davis M., Peebles P.J.E., 1983, ApJ 267, 465

de Lapparent V., Geller M.J., Huchra J.P., 1988, ApJ 332, 44

Dinshaw N., Impey C.D., Foltz C.B., et al., 1994, ApJ 437, L87

Dinshaw N., Foltz C.B., Impey C.D., et al., 1995, Nature 373, 223

Dinshaw N., Impey C.D., 1996, ApJ 458, 73

Dinshaw N., Weymann R.J., Impey C.D., et al., 1997, ApJ 491, 45

Dinshaw N., Foltz C.B., Impey C.D., et al., 1998, ApJ 494, 567

Dobrzycki A., Bechtold J., 1991, in: *The Space Distribution of Quasars*, Proc. Workshop, Crampton ed., ASP Conference series 21, p. 272

Dobrzycki A., Bechtold J., 1991, ApJ 377, L69

D'Odorico S., 1990, ESO The Messenger 61, 51

D'Odorico S., 1997, in: *The Early Universe with the VLT*, Proc. ESO Workshop, ed. J. Bergeron, Berlin: Springer, p.54

D'Odorico V., Cristiani S., D'Odorico S., Fontana A., Giallongo E., 1998, A&AS 127, 217

D'Odorico V., Cristiani S., D'Odorico S., Fontana A., Giallongo E., Shaver P., 1998, A&A 339, 678

Donahue M., Shull J.M., 1991, ApJ 383, 511

Doroshkevich A.G., Shandarin S.F., 1977, MNRAS 179, 95

Doroshkevich A.G., Mücket J.P., 1985, Soviet AL 11, 137

Duncan R.C., Ostriker J.P., Bajtlik S., 1989, ApJ 345, 39

Elowitz, R.M., Green, R.F., Impey, C.D., 1995, ApJ 440, 458

Elvis M., Wilkes B.J., McDowell J.C., Green R.F., Bechtold J., et al., 1994, ApJS 95, 1

Emerson D., 1996, *Interpreting Astronomical Spectra*, John Wiley & Sons, Chichesterr, England

Espey B.R., 1993, ApJ 411, L59

Fall S.M., Rees M.J., 1985, ApJ 298, 18

Fall S.M., Pei Y.C., 1993, ApJ 402, 479

Fan X., Tytler D., 1994, ApJS 94, 17

Fang L.Z., 1991, A&A 244, 1

Fang Y., Duncan R.C., Crotts A.P.S., Bechtold J., 1996, ApJ 462, 77

Ferland G.J., 1996, *Hazy, a Brief Introduction to Cloudy*, University of Kentucky Department of Physics and Astronomy Internal Report

Fernandez-Soto A., Lanzetta K.M., Barcons X., Carswell R.F., Webb J.K., et al., 1996, ApJ 460, L85

Fisher K. B., Davis M., Strauss M.A., Yahil A., Huchra J., 1994, MNRAS 266, 50

Foltz, C.B., Weymann, R.J., Roser, H.J., et al., 1984, ApJ 281, L1

Foltz, C.B., Weymann, R.J., Peterson B.M., Sun L., Malkan M.A., Chaffe F.H., Jr, 1986, ApJ 307, 504

Fontana A., Ballester P., 1995, ESO The Messenger 80, 37

Francis P.J., Hewett P.C., 1993, AJ 105, 1633

Frank J., King A., Raine D., 1992, *Accretion Power in Astrophysics*, Cambridge Univ. Press

Fransson C., Epstein R., 1982, A& A 198, 1127

Goldreich P., Sargent W.L.W., 1976, CommAp 6, 133

Giallongo E., Cristiani S., Fontana A., Trèvese D., 1993, ApJ 416, 137

Giallongo E., D'Odorico S., Fontana A., McMahon R., Savaglio S., et al., 1994, ApJ 425, L1

Giallongo, E., Cristiani, S., D'Odorico, S., Fontana A., Savaglio S., 1996, ApJ 466, 46.

Giavalisco M., Steidel C.C., Adelberger K.L., Dickinson M.E., Pettini M., Kellog M., 1998, ApJ 503, 543

Giroux M.L., Shull J. M., 1997, AJ 113, 1505

Gnedin N.Y., Hui L., 1996, ApJ 472, L73

Gnedin N.Y., Hui L., 1998, MNRAS 296, 44

Governato F., Baugh C.M., Frenk C.S., Cole S., Lacey C.G., et al., 1998, Nature 392, 359

Guillemin P., Bergeron J., 1997, A&A 328, 499

Gunn J.E., Peterson B.A., 1965, ApJ 142, 1633

Haardt F., Madau P., 1996, ApJ 461, 20

Haehnelt M.G., Steinmetz M., Rauch M., 1996, ApJ 465, 95

Haehnelt M.G., Rauch M., Steinmetz M., 1996, MNRAS 283, 1055

Haehnelt M.G., Steinmetz M., 1998, MNRAS 298, 21p

Haehnelt M.G., Steinmetz M., Rauch M., 1998, ApJ 495, 647

Heisler J., Hogan C.J., White S.D.M., 1989, ApJ 347, 52

Hernquist L., Katz N., Weinberg D., Miralda-Escudé J., 1996, ApJ 457, L51

Hogan C.J., Anderson S.F., Rugers M.H., 1997, AJ 113, 1495

Hoyle F., Fowler W.A., 1963, Nature 197, 533

Hu E.M., Kim T.-S., Cowie L.L., Songaila A. Rauch M., 1995, AJ 110, 1526

Hui L., Gnedin N.Y., Zhang Y., 1997, ApJ 486, 599

Hui L., Gnedin N.Y., 1997, MNRAS 292, 27

Hui L., Stebbins A., Burles S., 1999, ApJ 511, L5

Hui L., 1999, ApJ 516, 519

Hunstead R.W., Murdoch H.S., Pettini M., Blades J.C., 1988, ApJ 329, 527

Ikeuchi S., 1986, Ap&SS 118, 509

Ikeuchi S., Ostriker J.P., 1986, ApJ 301, 522

Impey C.D., Neugebauer G., 1988, AJ 95, 307

Impey C.D., Petry C.E., Malkan M.A., Webb W., 1996, ApJ 463, 473

Jakobsen P., Perryman M.A.C., Di Serego Alighieri S., et al., 1986, ApJ 303, L27

Jakobsen P., Boksenberg A., Deharveng J.M., Greenfield P., Jedrzejewski R., et al., 1994, Nature 370, 35

Jannuzi B.T., Bahcall J.N., Bergeron J., Boksenberg A., Hartig G.F., et al., 1998, ApJS 118, 1

Jauncey D.L., Wright A.E., Peterson B.A., Condon J.J., 1978, ApJ 223, L1

Jenkins E.B., Ostriker J.P., 1991, ApJ 376, 33

Kaiser N., 1987, MNRAS 227, 1

Kaiser N., Peacock J.A., 1991, ApJ 379, 482

Katz N., Weinberg D.H., Herquist L., Miralda-Escudé J., 1996, ApJ 457, L57

Kembhavi A.K., Narlikar J.V., 1999, *Quasars and Active Galactic Nuclei - An introduction*, Cambridge University Press, Cambridge

Kim T.-S., Hu E.M., Cowie L.L., Songaila A., 1997, AJ 114, 1

Kinman T.D., 1966, ApJ 144, 1232

Kirkman D., Tytler D., 1997, ApJ 484, 672

Kulkarni V.P., Fall S.M., 1993, ApJ 413, 63

Kulkarni V.P., Huang K., Green R.F., Bechtold J., Welty D.E., York D.G., 1996, MNRAS 279, 197

Lanzetta K.M., et al., 1991, ApJS 77, 1

Lanzetta K.M., Bowen D.B., Tytler D., Webb J.K., 1995, ApJ 442, 538

Lanzetta K.M., Wolfe A.M., Turnshek D.A., 1995, ApJ 440, 435

Lanzetta K.M., Wolfe A.M., Altan H., Barcons X., Chen H-W., et al., 1997, AJ 114, 133

Le Brun V., Bergeron J., Boissé P., 1996, A&A 306, 691

Le Brun V., Bergeron J., Boissé P., Deharveng J.M., 1997, A&A 321, 733

Le Brun V., Bergeron J., 1998, A&A 332, 814

Ledoux C., Petitjean P., Bergeron J., Wampler E.J., Srianand R., 1998, A&A 337, 51

Le Fèvre O., Hudon D., Lilly S.J., Crampton D., Hammer F., Tresse L. 1996, ApJ, 461, 534

Liu X.D., Jones B.J.T., 1990, MNRAS 242, 678

Lu L., 1991, ApJ 379, 99

Lu L., Wolfe A.M., Turnshek D.A., 1991, ApJ 367, 19

Lu L., Sargent W.L.W., Womble D.S., Takada-Hidai M., 1996, ApJ 472, 509

Lucy L.B., 1974, AJ 79, 745

Lynds C.R., 1971, ApJ 174, L73

Lynds C.R., Stockton A.N., 1966, ApJ 144, 446

MacAlpine G.M., Feldman F.R., 1982, ApJ, 261, 412

McDonald P., Miralda-Escudé J., 1999, ApJ 518, 24

McGill C., 1990, MNRAS 242, 544

Meiksin A., Madau P., 1993, ApJ 412, 34

Meiksin A., 1994, ApJ 431, 109

Meiksin A., Bouchet F.R., 1995, ApJ 448, L85

Meiksin A., 1997, in: *Young Galaxies and QSO Absorption Systems*, eds. S.M. Viegas, R. Gruenwald, R.R. de Carvalho, ASP Conf. Series, San Francisco, 1

Melott A., 1980, ApJ 268, 630

Meyer D.M., York D.G., 1987, ApJ 315, 5

Milgrom M., 1988, A&A 202, 9

Miralda-Escudé J., Ostriker J.P., 1990, ApJ 350, 1

Miralda-Escudé J., Rees M.J., 1993, MNRAS 260, 617

Miralda-Escudé J., Cen R., Ostriker J.P., Rauch M., 1996, ApJ 471, 582

Mo H.J., Xia X.Y., Deng Z.G., Börner G., Fang L.-Z., 1992, A&A 256, L23

Mo H.J., 1994, MNRAS 269, 49

Mo H.J., Miralda-Escudé J., 1996, ApJ 469, 589

Morris S.L., Weymann R.J., Foltz C.B., Turnshek D.A., Shectman S., et al., 1986, ApJ 310, 40

Morris S.L., Weymann R.J., Savage B.D., Gilliland R.L., 1991, ApJ 377, L21

Morris S.L., Weymann R.J., Dressler A., McCarthy P.J., Smith B.A., Terrile R.J., Giovanelli R., Irwin M., 1993, ApJ 419, 524

Morris S.L., van den Bergh S. 1994, ApJ 427, 696

Morton D.C., Smith W.H., 1973, ApJS 26, 333

Morton D.C., 1991, ApJS 77, 119

Mücket J.P., Petitjean P., Kates R.E., Riediger R., 1996, A&A 308, 17

Murakami I., Ikeuchi S., 1990, PASJ 42, L11

Murakami I., Ikeuchi S., 1993, ApJ 409, 42

Murdoch H.S., Hunstead R.W., Pettini M., Blades J.C., 1986, ApJ 309, 19

Norris J., Hartwick F.D.A., Peterson B.A., 1983, MNRAS 273, 450

Nusser A., Hahehnelt M., 1999, MNRAS 303, 179

Ostriker J.P., Cowie L.L., 1981, ApJ 243, 127

Ostriker J.P., Ikeuchi S., 1983, ApJ 268, L63

Ostriker J.P., Bajtlik S., Duncan R.C., 1988, ApJ 327, 350

Pando J., Fang L.-Z., 1996, ApJ 459, 1

Peacock J., 1991, Nature 349, 190

Peebles P.J.E., 1980, *The Large Scale Structure of the Universe*, Princeton Univ. Press, Princeton

Peterson B.A., 1978, in *The Large Scale Structure of the Universe*, Longair M.S., Einasto J. eds., D. Reidel, Dordrecht, p. 389

Peterson B.M., 1993, PASP 105, 1084

Petitjean P., Bergeron J., 1990, A&A 231, 309

Petitjean P., Bergeron J., Carswell R.F., Puget J.L., 1993, MNRAS 260, 67

Petitjean P., Webb J.K., Rauch M., Carswell R.F., Lanzetta K.M., 1993, MNRAS 262, 499

Petitjean P., Bergeron J., 1994, A&A 283, 759

Petitjean P., Rauch M., Carswell R.F., 1994, A& A 291, 29

Petitjean P., Mücket J.P., Kates R.E., 1995, A&A 295, L12

Petitjean P., 1998, in *Formation and Evolution of galaxies*, Proc. of the Les Houches School, eds. O. Le Fevre and S. Charlot, Springer-Verlag, astro-ph/9810418

Petitjean P., Surdej, S., Smette, A., et al., 1998, A&A 334, L45

Pettini M., Hunstead R.W., Murdoch H.S., Blades J.C., 1983, ApJ 273, 436

Pettini M., Hunstead R.W., Smith L.J., Mar D.P., 1990, MNRAS 246, 545

Pettini M., Smith L.J., King D.L., Hunstead R.W., 1997, ApJ 486, 665

Pettini M., Ellison S.L., Steidel C.C., Bowen D.V., 1999, ApJ 510, 576

Pierre M., Shaver P.A., Iovino A., 1988, A&A 197, 3

Press, S.J., 1989, *Bayesian Statistics, Principles, Models & Applications*, Wiley, New York

Press W.H., Rybicki G.B., 1993, ApJ 418, 585

Press W.H., Rybicki G.B., Schneider D.P. 1993, ApJ, 414, 64

Prochaska J.X., Wolfe A.M., 1997, ApJ 487, 73

Prochaska J.X., Wolfe A.M., 1998, ApJ 507, 113

Prochaska J.X., Wolfe A.M., 1999, ApJS 121, 369

Quashnock J.M., Vanden Berk D.E., York D.G. 1996, ApJ, 472, L69

Quashnock J.M., Vanden Berk D.E. 1998, ApJ, 500, 28

Rao S.M., Turnshek D.A., 1998, ApJ 500, L115

Rauch M., Carswell R.F., Chaffee F.H., Foltz C.B., Webb J.K., et al., 1992, ApJ 390, 387

Rauch M., Carswell R.F., Webb J.K., Weymann R.J., 1993, MNRAS 260, 589

Rauch M., Haehnelt M.G., 1995 MNRAS, 275, 76

Rauch M., Sargent W.L.W., Womble D.S., Barlow T.A., 1996, ApJ 467, 5L

Rauch M., 1996, in: *Cold Gas at High Redshift*, Proc. of the Kluwer Conf., M.N. Bremer et al. eds., Dordrecht, Kluwer, p. 137

Rauch M., Haehnelt M.G., Steinmetz M., 1997, ApJ 481, 601

Rauch M., Miralda-Escudé J., Sargent W.L.W., Barlow T.A., Weinberg D.H., et al. 1997, ApJ 489, 7

Rauch M., 1998, in *Structure and Evolution of the IGM*, Proc. of the XIII IAP Meeting, eds. S. Charlot, P. Petitjean, Editions Frontiéres, Gif-sur-Yvette,

Rauch M., 1998, ARA&A 36, 267

Rees M.J., 1984, ARA&A 22, 471

Rees M.J., 1986, MNRAS 218, 25

Rees M.J., 1988, in: Proceedings of the QSO Absorption Line Meeting, Baltimore 1987, *QSO Absorption Lines, Probing the Universe*, eds. J.C. Blades, D.A. Turnshek, C.A. Norman, Cambridge, Cambridge University Press

Rees M.J., Setti G., 1970, A&A 8, 410

Reimers D., Köhler S., Wisotzki L., Groote D., Rodríguez-Pascal P., Wamsteker W., 1997, A& A 327, 890

Reisenegger A., Miralda-Escudé J., 1995, ApJ 449, 476

Remy M., Claeskens J.-F., Surdej J., Hjorth J., Refsdal S., et al., 1998, New Astronomy 3, 379

Richards G.T., York D.G., Yanni B., Kollgaard R.I., Laurent-Mühleisen S.A., et al., 1999, ApJ 513, 576

Riediger R., Petitjean P., Mücket J.P., 1998, A&A 329, 30.

Rodríguez-Pascual P.M., De La Fuente A., Sanz J.L., Recondo M.C., Clavel J., Santos-Lleó M., Wamsteker W., 1995, ApJ 448, 575

Salpeter E.E., Hoffman G.L., 1995, ApJ 441, 51

Salzer J.J., 1992, AJ 103, 385

Sanders D.B., Phinney E.S., Neugebauer G., Soifer B.T., Matthews K., 1989, ApJ 347, 29

Sargent W.L.W., Young P.J., Boksenberg A., Tytler D., 1980, ApJS 42, 41

Sargent W.L.W., Young P.J., Schneider D.P., 1982, ApJ 256, 374

Sargent W.L.W., Boksenberg A., 1983, in *Quasar and Gravitational Lenses*, 24th Liège Astrophysical Colloquium, p. 518

Sargent W.L.W., Boksenberg A., Steidel C.C., 1988, ApJS 68, 539

Sargent W.L.W., Steidel C.C., Boksenberg A., 1988, ApJ 334, 22

Sargent W.L.W., Steidel C.C., Boksenberg A., 1989, ApJ 69, 703

Sargent W.L.W., Steidel C.C., Boksenberg A., 1990, ApJ 351, 364

Savaglio S., Cristiani S., D'Odorico S., Fontana A., Giallongo E., Molaro P., 1997, A&A 318, 347

Savaglio S., Ferguson H.C., Brown T.M., Espey B.R., Sahu K.C., 1999, ApJ 515, L5

Scheuer P.A.G., 1965, Nature 207, 963

Schneider D., Schmidt M., Gunn J.E., 1991, AJ 102, 837

Schwarz J., Ostriker J.P., Yahil A., 1975, ApJ 202, 1

Shapiro P.R., Giroux M.L., Babul A., 1994, ApJ 427, 25

Shaver P.A., Robertson J.G. 1983, ApJ 268, L57

Shklovski I.S., 1965, SovA 8, 638

Shields G.A., 1978, Nature 272, 706

Shull J.M., Stocke J.T., Penton S., 1996, AJ 111, 72

Slezak E., Bijaoui A., Mars G., 1990, A& AS 227, 301

Smette A., Surdej J., Shaver P.A., et al., 1992, ApJ 389, 39

Smette A., Robertson J.G., Shaver P.A., et al., 1995, A&AS 113, 199

Songaila A., Cowie L.L., 1996, AJ 112, 335

Songaila A., 1998, AJ 115, 2184

Spitzer L., 1956, ApJ 124, 20

Spitzer L. Jr., 1978, *Physical Processes in the Interstellar Medium*, John Wiley and Sons, New York, 52

Steidel C.C., 1990, ApJS 72, 1

Steidel C.C., Sargent W.L.W., 1992, ApJS 80, 1

Steidel C.C., Hamilton D. 1993, AJ 105, 2017

Steidel C.C., Dickinson M., Persson S.E., 1994, ApJ 437, L75

Steidel C.C., Pettini M., Dickinson M., Persson S.E., 1994, AJ 108, 2046

Steidel C.C., 1995, in: *QSO Absorption Lines*, Proc. ESO Workshop, ed. Meylan G., Berlin, Springer Verlag, p139

Steidel C.C., Pettini M., Hamilton D., 1995, AJ 110, 2519

Steidel C.C., Giavalisco M., Dickinson M., Adelberger K., 1996, AJ 112, 352

Steidel C.C., Giavalisco M., Pettini M., Dickinson M., Adelberger K., 1996, ApJ 462, L17

Steidel C.C., Dickinson M., Giavalisco M., Pettini M., Kellog M. 1998, ApJ 492, 428

Steidel C.C., Adelberger K.L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ 519, 1

Steinmetz M., 1997, in: *The Early Universe with the VLT*, Proc. ESO Workshop, ed. J. Bergeron, Berlin, Springer, p.156

Stockton A.N., Lynds C.R., 1966, ApJ 144, 451

Stone R.P.S., Baldwin J.A., 1983, MNRAS 204, 347

Storrie-Lombardi L., McMahon R.G., Irwin M.J., 1996, MNRAS 283, L79

Theuns T., Leonard A.P.B., Efstathiou G., 1998, MNRAS 297, L49

Theuns T., Leonard A.P.B., Schaye J., Efstathiou G., 1999, MNRAS 303, L58

Tresse L., Dennefeld M., Petitjean P., Cristiani S., White S., 1999, A&A 346, L21

Turnshek D.A., Kopko M. Jr., Monier E., Noll D., Espey B.R., et al., 1996, ApJ 463, 110

Tyson N.D., 1988, ApJ 329, L57

Tytler D., 1987, ApJ 321, 49

Tytler D., 1987, ApJ 321, 69

Tytler D., Sandoval J., Fan X.-M., 1993, ApJ 405, 57

Tytler D., Fan X.-M., Burles S., Cottrell L., Davis C., et al., 1995, in *Quasar Absorption Lines*, Proc. of the ESO Workshop, ed. G. Meylan, Berlin, Springer, p. 289

Ulmer A., 1996, ApJ 473, 110

Umemura M., Ikeuchi S., 1985, ApJ 299, 583

Urry C.M., Padovani P., 1995, PASP 107, 803

Verner D.A, Verner E.M., Ferland G.J., 1996, BAAS 188, 5418

Vishniac E.T., Ostriker J.P., Bertschinger E., 1985, ApJ 291, 399

Vishniac E.T., Bust G.S., 1987, ApJ 319, 14

Vladilo G., 1998, ApJ 493, 583

Vogt S.S., Allen S.L., Bigelow B.C., Bresee L., Brown B. et al., 1994, S.P.I.E. 2198, 362

Wagoner R., 1967, ApJ 149, 465

Walker T.P., Steigman G., Schramm D.N., et al., 1991, ApJ 376, 51

Wampler E.J., Chugai N.N., Petitjean P. 1995, ApJ 443, 586

Wang B., 1993, ApJ 415, 174

Wang B., 1995, ApJ 444, L17

Webb J. K., 1987, in: *Observational Cosmology*, Proc. of the IAU Symposium 124, Hewett A., Burbidge G., Fang L.Z. eds., Reidel, Dordrecht, p. 803

Webb J.K., Larsen I., 1988, in *High Redshift and Primeval Galaxies*, eds. Bergeron J., Kunth D., Rocca-Volmerange B., Editions Frontière, Gif-sur-Yvette

Webb J. K., Barcons X., 1991, MNRAS 250, 270

Webb J.K., Barcons X., Carswell R.F., Parnell H.C., 1992, MNRAS 255, 319

Weedman D.W., 1986, *Quasar Astronomy*, Cambridge University Press, Cambridge, England

Weinberg D.H., Ostriker J.P., Dekel A., 1989, ApJ 336, 9

Weinberg D.H., 1992, MNRAS 254, 315

Weinberg D.H., Katz N., Hernquist L., 1998, in: *Origins*, eds. C.E. Woodward, J.M. Shull, and H.A. Thronson, Jr., ASP Conference Series, Vol. 148, p.21

Weisheit J.C., 1978, ApJ 219, 829

Weymann R.J., Williams R.E., Peterson B.M., Turnshek D.A., 1979, ApJ 218, 619

Weymann R.J., Carswell R.F., Smith M.G., 1981, ARA&A 19, 41

Weymann R.J., Foltz C.B., 1983, ApJ 272, L1

Weymann R.J., Morris S.L., Foltz C.B., Hewett P.C. 1991, ApJ 373, 23

Weymann R.J., Jannuzi B.T., Lu L., Bahcall J.N., Bergeron J., et al., 1998, ApJ 506, 1

Wilkes B.J., Elvis M., 1987, ApJ 323, 243

Williger G.M., Carswell R.F., Webb J.K., Boksenberg A., Smith M.G., 1989, MNRAS 237, 635

Williger G.M., Baldwin J.A., Carswell R.F., Cooke A.J., Hazard C., et al., 1994, ApJ 428, 574

Williger G.M., Hazard C., Baldwin J.A., et al., 1996, ApJS 104, 145

Wolfe A.M., 1986 in: Proc. NRAO Conf. on Gaseous Halos of Galaxies. NRAO SP, Bregman J., Lockman J. (eds.) p. 259

Wolfe A.M., Turnshek D.A., Smith H.E., Cohen R.D., 1986, ApJ 61, 249

Wolfe A.M., 1988, in *QSO Absorption Line: Probing the Universe*, Proc. of the QSO Absorption Line Meeting, Baltimore, 1987, Cambridge University Press

Wolfe A.M., Prochaska J.X., 1998, ApJ 494, L15

Womble D.S., Sargent W.L.W., Lyons R.S., 1996, in: *Cold Gas at High Redshift*, Proc. of the Kluwer Conf., eds. Bremer M.N. et al., Kluwer, p. 137

York D.G., Dopita M., Green R., Bechtold J., 1986, ApJ 311, 610

York D.G., Yanny B., Crotts A., Carilli C., Garrison E., Matheson L., 1991, MNRAS 250, 24

Young, P.J., Sargent, W.L.W., Boksenberg, A., Carswell R.F., Whelan J.A.J., 1979, ApJ, 229, 891

Young P.J., Sargent W.L.W., Boksenberg A., 1982a, ApJS 48, 455

Young P.J., Sargent W.L.W., Boksenberg A., 1982b, ApJ 252, 10

Zeldovich Y.B., 1970, A&A 5, 84

Zhang Y., Anninos P., Norman M.L., 1995, ApJ 453, L57

Zhang Y., Anninos P., Norman M.L., 1996, ApJ 459, 12

Zhang Y., Anninos P., Norman M.L., Meiksin A., 1997, ApJ 485, 496

Zhang Y., Meiksin A., Anninos P., Norman M.L., 1998, ApJ 495, 63

Zheng W., Davidsen A.F., Kriss G.A., 1998, AJ 115, 391.