Mon. Not. R. Astron. Soc. **372,** 1333–1344 (2006) doi:10.1111/j.1365-2966.2006.10941.x

Tomography of the intergalactic medium with Lyα forests in close QSO pairs*

V. D'Odorico,[1]† M. Viel,[1,2] F. Saitta,[3] S. Cristiani,[1] S. Bianchi,[4] B. Boyle,[5,6] S. Lopez,[7] J. Maza[7] and P. Outram[8]

[1]*INAF – Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, Trieste I-34131, Italy*
[2]*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[3]*Dipartimento di Astronomia, Università degli Studi di Trieste, via G.B. Tiepolo 11, Trieste I-34131, Italy*
[4]*INAF – Istituto di Radioastronomia, Sezione di Firenze, Largo E. Fermi 5, Firenze I-50125, Italy*
[5]*Australia Telescope National Facility, PO Box 76, Epping, NSW 1710, Australia*
[6]*Anglo-Australian Observatory, PO Box 296, Epping, NSW 2121, Australia*
[7]*Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile*
[8]*Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE*

Accepted 2006 August 14. Received 2006 August 9; in original form 2006 March 17

ABSTRACT

We study the three-dimensional distribution of non-virialized matter at $z \sim 2$ using high-resolution spectra of quasi-stellar object (QSO) pairs and simulated spectra drawn from cosmological hydrodynamical simulations. We have collected the largest sample of QSO pairs ever observed with Ultraviolet and Visual Echelle Spectrograph (UVES) at the European Southern Observatory-Very Large Telescope (ESO-VLT), with angular separations between ∼1 and 14 arcmin. The observed correlation functions of the transmitted flux in the H I Lyman α forest along and transverse to the lines of sight are in good agreement implying that the distortions in redshift space due to peculiar velocities are small. The clustering signal is significant up to velocity separations of $\sim$200 km s^{-1}, or about $3\,h^{-1}$ comoving Mpc. The regions at lower overdensity ($\rho/\bar{\rho} \lesssim 6.5$) are still clustered but on smaller scales ($\Delta v \lesssim 100\,\mathrm{km\,s^{-1}}$). The observed and simulated correlation functions are compatible at the 3σ level. A better concordance is obtained when only the low overdensity regions are selected for the analysis or when the effective optical depth of the simulated spectra is increased artificially, suggesting a deficiency of strong lines in the simulated spectra. We found that also a lower value of the power-law index of the temperature–density relation for the Lyman α forest gas improves the agreement between observed and simulated results. If confirmed, this would be consistent with other observations favouring a late He II reionization epoch (at $z \sim 3$). We remark the detection of a significant clustering signal in the cross-correlation coefficient at a transverse velocity separation $\Delta v_\perp \sim 500\,\mathrm{km\,s^{-1}}$ whose origin needs further investigation.

Key words: intergalactic medium – quasars: absorption lines – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

The study of the evolution of cosmic structures is one of the major topic of present-day cosmology.

It is now more than a decade that thanks to semi-analytical and hydrodynamical simulation results (e.g. Cen et al. 1994; Zhang, Anninos & Norman 1995; Hernquist et al. 1996; Miralda-Escudé et al. 1996; Bi & Davidsen 1997; Davé et al. 1997; Zhang et al. 1997; Theuns, Leonard & Efstathiou 1998a; Machacek et al. 2000) the majority of the absorption features observed in Lyman α forests of high-redshift quasi-stellar object (QSO) spectra is identified with the fluctuations of the intermediate- and low-density intergalactic medium (IGM), arising naturally in the hierarchical process of structure formation. The physics of this highly ionized gas is simple, governed mainly by the Hubble expansion and the gravitational instability (Rauch et al. 2005). If photoionization equilibrium is assumed, a mean relation between temperature and overdensity is obtained (Hui & Gnedin 1997):

*Based on observations collected at the European Southern Observatory Very Large Telescope, Cerro Paranal, Chile – Programs 65.O-0299(A), 68.A-0216(A), 69.A-0204(A), 69.A-0586(A), 70.A-0031(A), 166.A-0106(A).

†E-mail: dodorico@oats.inaf.it

$$T = T_0\,(\rho/\bar{\rho})^{\gamma-1}, \quad (1)$$

where T_0 and γ depend on the ionization history of the Universe (for an early reionization, $T_0 = 10^4$ K and $\gamma = 1.6$ are generally adopted). The combination of photoionization equilibrium and equation (1) leads to a power-law relation between measured H I Lyman α optical depth and gas overdensity (see e.g. Weinberg et al. 1999, for a review), which, adopting the standard values for the parameters, can be written as

$$\tau_{\rm HI} \simeq 1.3 \times 10^{-3}(1+z)^{9/2}\,(\rho/\bar{\rho})^{3/2}. \quad (2)$$

A critical test of the nature of the Lyman α absorbers as proposed by simulations, comes from the determination of their sizes and spatial distribution. Correlations in the Lyman α forest were detected with a 4–5σ confidence by various authors at typical scales $\Delta v \lesssim 350\,{\rm km\,s^{-1}}$ observing at high-resolution individual lines of sight (Lu et al. 1996 at $z \sim 3.7$; Cristiani et al. 1997 at $z \sim 3$; Kim, Cristiani & D'Odorico 2001 at $z \sim 2$). This velocity range corresponds to scales $\lesssim 2.5\,h^{-1}$ Mpc (considering peculiar velocities negligible). In this observational approach, however, the three-dimensional information is convoluted with distortions in the redshift space, due to peculiar motions and thermal broadening.

Multiple lines of sight offer an invaluable alternative to address the spatial distribution of the absorbers, enabling a more direct interpretation of the observed correlations. The interesting range of separations lies about the Jeans scale of the photoionized IGM ($\sim$1 arcmin or $\sim 1.4\,h^{-1}$ comoving Mpc at $z \sim 2$). At this scale there should be a transition from a smooth gas distribution, which produces nearly identical absorption features in neighbouring lines of sight, to a correlated density distribution, where the correlation strength decreases with increasing separation of the lines of sight (e.g. Viel et al. 2002). Indeed, the spectra of multiple images of lensed quasars with separations of the order of a few arcsec (Smette et al. 1992, 1995; Impey et al. 1996; Rauch et al. 2001) show nearly identical Lyman α forests, implying that the absorbing objects have large sizes ($>50\,h^{-1}$ kpc). For pairs with larger separation the correlation between the absorption features becomes weaker. Studies based on the statistics of coincident and anticoincident absorption lines provided evidence for dimensions of a few hundred kpc (Bechtold et al. 1994; Dinshaw et al. 1994, 1995, 1997; Fang et al. 1996; Crotts & Fang 1998; D'Odorico et al. 1998; Petitjean et al. 1998; Lopez, Hagen & Reimers 2000; Young, Impey & Foltz 2001; Becker, Sargent & Rauch 2004). The obtained large sizes were conclusive to exclude models of the Lyman α forest as a population of pressure confined small clouds (e.g. Sargent et al. 1980) or clouds in dark matter minihaloes (e.g. Miralda-Escudé & Rees 1993).

The new theoretical scenario and the production of exceptional quality QSO spectra by high-resolution spectrographs at 10-m class telescopes, started the use of the Lyman α forest as a cosmological probe. From the study of absorption spectra along single lines of sight to distant QSOs and the implementation of various algorithms to convert the observed flux into density it was possible, for example, to determine the shape and amplitude of the power spectrum of the spatial distribution of dark matter (Croft et al. 1998, 1999, 2002; Nusser & Haehnelt 1999, 2000; Viel, Haehnelt & Springel 2004b) or, more straightforwardly, of the transmitted flux (McDonald et al. 2000; Kim et al. 2004).

The final goal of this paper is to investigate the distribution properties of matter in the IGM applying the modern interpretation of the Lyman α forest. In particular, we computed the correlation properties of the transmitted flux along and across the lines of sight for a new exceptional sample of high-resolution spectra of close QSO pairs and groups at $z_{\rm em} \sim 2$. Previous works following a similar approach (Rollinde et al. 2003; Coppolani et al. 2006) used sets of low-resolution spectra of QSO pairs with the main aim of deriving constraints on $\Omega_{0\Lambda}$ from the application of the Alcock–Paczyński test (Alcock & Paczyński 1979, see also Section 6). However, about $13(\Delta\theta/(1\,{\rm arcmin}))^2$ QSO pairs with separation $<\Delta\theta$ are needed to determine $\Omega_{0\Lambda}$ with a precision better than 10 per cent (McDonald 2003). So, we will not attempt to determine $\Omega_{0\Lambda}$ in this study, although our plan is to gather a larger sample of QSO pairs in a few years.

The present investigation is completed by the comparison with a set of mock spectra drawn from a hydrodynamical simulation, reproducing several different realizations of our sample of QSO associations. We analysed how variations of the mean flux (corresponding to variations of the intensity of the ionizing background) and of the temperature–density relation of the gas affect the correlation functions.

In this preliminary study, we do not take advantage of the presence, in our sample, of two groups of QSOs (one triplet and one sestet) to detect structures extending on large scales. Since new observing time have been allocated to our project at Ultraviolet and Visual Echelle Spectrograph-Very Large Telescope (UVES-VLT), a forthcoming paper will be devoted to the detailed analysis of coincidences and anticoincidences both of Lyman α and metal absorptions among multiple lines of sight, on improved signal-to-noise ratio (S/N) spectra.

The paper structure is as follows: in Section 2 we describe the observed data sample, the reduction procedure and the simulated spectra. Sections 3, 4 and 5 are devoted to the computation of the auto- and cross-correlation functions of both the observed and simulated spectra and to their comparison. We discuss some interesting aspects of our results and we explore the parameter space of the simulations in Section 6. The conclusions are drawn in Section 7.

Throughout this paper we adopt $\Omega_{0\rm m} = 0.3$, $\Omega_{0\Lambda} = 0.7$ and $h = H_0/72\,{\rm km\,s^{-1}\,Mpc^{-1}}$.

2 DATA SAMPLE

2.1 The observed spectra

The exploitation of the potential offered by multiple QSO lines of sight has been limited by the dearth of suitable groups of QSOs close and bright enough to permit high-resolution spectroscopy. Two major breakthroughs have dramatically improved this situation:

(i) the 2dF QSO Redshift Survey (Croom et al. 2004, www.2dfquasar.org), whose complete spectroscopic catalogue contains more than $\sim$23 000 QSOs in a single homogeneous data base, which is approximately 50 times more than the previous largest QSO survey to a similar depth ($B < 21$);

(ii) the UVES spectrograph (Dekker et al. 2000) at the Kueyen unit of the European Southern Observatory (ESO) VLT (Cerro Paranal, Chile) which has a remarkable efficiency especially in the extreme UV (close to the atmospheric cut-off). In this way, relatively low-redshift ($z \lesssim 2.5$) QSOs become accessible to high-resolution observations of the Lyman α forest down to faint apparent magnitudes and their surface density becomes high enough to provide several lines of sight for a tomography of the IGM.

We searched the 2dF QSO data base for the best groups with apparent magnitude $B \leqslant 20$ and $z > 1.8$. A triplet found in the Calán–Tololo QSO survey (Maza et al. 1993, 1995) was added to the sample. A great observational effort was carried out to collect

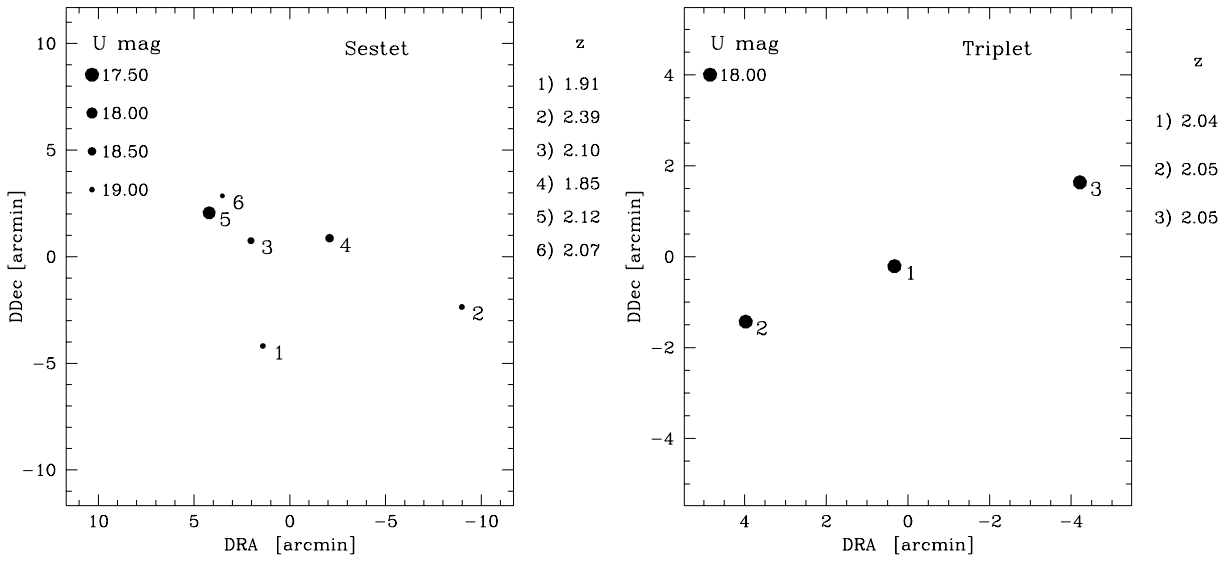

Figure 1. Relative positions, *U* magnitudes and redshifts of the QSOs composing the sestet (left-hand plot) and the triplet (right-hand plot) in our sample.

UVES spectra of the selected QSOs, which all have magnitudes fainter than $B \sim 19$, with the exception of the triplet for which $B \sim 18$. ESO allocated 51 h of observation to our project up to now, which allowed us to obtain acceptable S/N spectra of one pair (from now on called pair A, with angular separation of ∼1.3 arcmin), one sestet and the triplet of QSOs (see Fig. 1).

Reduction of the QSO spectra was conducted with the pipeline of the instrument (version 2.1, Ballester et al. 2000) provided by ESO in the context of the data reduction package MIDAS. In most cases, we could not apply the standard procedure due to the faintness of the spectra in the blue region and had to pre-filter the cosmic rays to make the optimal extraction work properly. Single extracted spectra were summed and rebinned and wavelengths were corrected to the vacuum-heliocentric reference frame. The final spectra have resolution $R \sim 40\,000$ in the Lyman α and C IV forest, while the S/N per pixel varies on average between S/N ∼ 3 and 10 in the Lyman α forest and between 4 and 15 in the C IV forest (see Table 1 for details).

Continuum determination, in particular in the Lyman α forest region, is a very delicate step in the process of spectra reduction. Tentative procedures realized up to now to objectively determine the continuum position through automatic algorithms do not give satisfactory results. We adopted a manual subjective method based on the selection of the regions free from clear absorption that are successively fitted with a spline polynomial of third degree.

The limitations introduced by the uncertainty in the true continuum level and shape should be less important in the computation of the cross-correlation function than in the case of single line of sight analysis. This because the undulations are uncorrelated between adjacent lines of sight, in particular, if the two QSOs do not have too similar redshifts (see Viel et al. 2002, for a discussion in the case of power spectra).

We added to the sample two more pairs from our archive: UM680/UM681 (pair U, separated by ∼1 arcmin) and Q2344+1228/Q2343+1232 (pair Q, separated by ∼5.57 arcmin). They were observed with UVES at the same resolution as the new data, and reduced following the same procedure. More details on their properties are given in the paper by D'Odorico, Petitjean & Cristiani (2002).

The list of the QSOs and the main characteristics of the spectra are reported in Table 1. The Lyman α forests of two of the closest pairs in our sample are shown in Fig. 2.

The total sample is formed by 21 pairs uniformly distributed between angular separations of ∼1 and 14 arcmin, corresponding to comoving spatial separations between ∼1.4 and $21.6\,h^{-1}$ Mpc. The median redhift of the Lyman α forest is $z \sim 1.8$. This is the largest sample of high-resolution spectra of QSO pairs ever collected, unique both for the number density – we have six QSOs in a region of ∼0.04 deg^2 – and the variety of line of sight separations investigated.

As a reference sample (see Sections 4 and 5.1), we considered eight UVES QSO spectra obtained in the framework of the ESO Large Program (LP) 'The Cosmic Evolution of the IGM' (Bergeron et al. 2004, LP sample). The QSOs were selected in order to match as much as possible the redshift range of our sample (see Table 2). The spectra have resolution $R \sim 45\,000$ and S/N ∼50 in the Lyman α forest region. They were normalized to the continuum and prepared for the analysis following the same procedure adopted for the spectra of the pair sample.

Table 1. Characteristics of the observed QSO spectra.

	Object	z	M_B	Lyman α range	S/N per pixel
Pair A	PA1	2.645	19.11	2.094–2.585	8–12
	PA2	2.610	19.84	2.094–2.550	3.5–6.5
Triplet	T1	2.041	18.20	1.633–1.991	3–10
	T2	2.05	18.30	1.592–1.999	3.5–8.5
	T3	2.053	18.10	1.665-2.002	2.5–6
Sestet	S1	1.907	19.66	1.665–1.859	2–5.5
	S2	2.387	19.53	1.858–2.331	3–7.5
	S3	2.102	19.31	1.633–2.051	4–10
	S4	1.849	19.59	1.575–1.802	2–8
	S5	2.121	18.85	1.633–2.069	3.5–8
	S6	2.068	20.19	1.592–2.017	3–8.5
Pair U	UM680	2.144	18.60	1.653–2.092	6.5–17
	UM681	2.122	19.10	1.634–2.070	7–17
Pair Q	Q2343+12	2.549	17.00	1.994–2.490	13–23
	Q2344+12	2.773	17.50	2.183–2.711	12–18

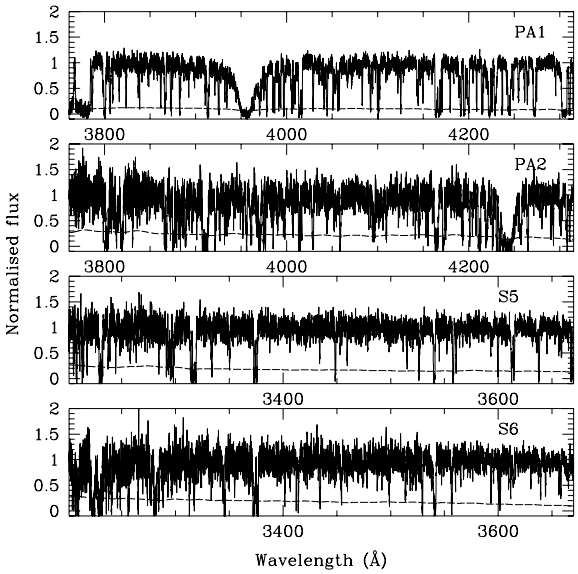

Figure 2. Lyman α forests of two of the closest pairs in our sample. Pair A (upper two panels) separated by 1.3 arcmin and the pair at 1.06 arcmin formed by the QSOs S5 and S6 of the sestet (lower two panels).

Table 2. List of the considered QSOs from the ESO Large Program (see text).

Object	Redshift	M_B	Lyman α range
HE 1341−1020	2.134	17.1	1.64431–2.08216
Q 0122−380	2.189	16.7	1.69072–2.13625
PKS 1448−232	2.220	16.96	1.71688–2.16674
HE 0001−2340	2.280	16.7	1.76750–2.22575
Q 0109−3518	2.406	16.44	1.87381–2.34966
HE 2217−2818	2.406	16.0	1.87381–2.34966
Q 0329−385	2.423	16.92	1.88816–2.36638
HE 1158−1843	2.453	16.93	1.91347–2.39589

2.2 The simulated spectra

In order to both assess the nature of the Lyman α forest inferred from simulations and to constrain the cosmological scenario of the same simulations, we compared the results obtained for our sample of observed QSO spectra with analogous results for a sample of mock Lyman α forests.

We used simulations run with the parallel hydrodynamical (TreeSPH) code GADGET-2 (Springel, Yoshida & White 2001; Springel 2005). The simulations were performed with periodic boundary conditions with an equal number of dark matter and gas particles and used the conservative 'entropy formulation' of smoothed particle hydrodynamics (SPH) proposed by Springel & Hernquist (2002). Radiative cooling and heating processes were followed for a primordial mix of hydrogen and helium. We assumed a mean UV background produced by quasars and galaxies as given by Haardt & Madau (1996) with helium heating rates multiplied by a factor 3.3 in order to fit observational constraints on the temperature evolution of the IGM. More details can be found in Viel et al. (2004b).

The cosmological model corresponds to a 'fiducial' Λ cold dark matter (CDM) Universe with parameters $\Omega_{0\mathrm{m}} = 0.26$, $\Omega_{0\Lambda} = 0.74$, $\Omega_{0\mathrm{b}} = 0.0463$ and $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (the B2 series of Viel et al. 2004b). We have used 2×400^3 dark matter and gas particles in a $120\,h^{-1}$ comoving Mpc box. The gravitational softening was set to $5\,h^{-1}$ kpc in comoving units for all particles. We note that the parameters chosen here, including the thermal history of the IGM, are in perfect agreement with observational constraints including recent results on the cosmic microwave background (CMB) and other results obtained by the Lyman α forest community (e.g. Spergel et al. 2003; Viel et al. 2004b; Seljak et al. 2005).

The $z = 1.8$ output of the simulated box was pierced to obtain 50 triplets of lines of sight carefully reproducing the observed triplet mutual separations. The same was done for 50 sestets of lines of sight reproducing the observed sestet and 50 pairs of lines of sight at the same angular separation as pair U. 50 different realizations of pair A spectra and 50 of pair Q were obtained from the output box at redshift $z = 2.4$.

The velocity extent of the spectra is of $10\,882.3\,\mathrm{km\,s^{-1}}$ for the sestet, the triplet and pair U, and of $11\,683.9\,\mathrm{km\,s^{-1}}$ for pairs A and Q. The pixel velocity size is fixed for each spectrum and slightly increases with the redshift of the box, being $\mathrm{d}v \simeq 2.66, 2.85\,\mathrm{km\,s^{-1}}$ for $z = 1.8, 2.4$, respectively. These values are comparable with the average pixel size of the observed spectra, $\mathrm{d}v_{\mathrm{obs}} \simeq 2.64\,\mathrm{km\,s^{-1}}$.

The mean H I optical depth in the simulation box is determined by the UV background at the considered redshift. The corresponding observed quantity is the so-called effective optical depth, τ_{eff}, which is obtained from the average flux measured in QSO absorption spectra: $\bar{f} = \exp(-\tau_{\mathrm{eff}})$. The τ_{eff} of the simulated spectra can be rescaled to a given value simply by multiplying the optical depth of every pixel by a constant factor (e.g. Theuns et al. 1998b; Bolton et al. 2005). So for the spectra reproducing our sample of QSOs we have adopted the average values measured in the observed spectra: $\tau_{\mathrm{eff}} = 0.10$ for the triplet and pair U (average Lyman α redshift $z_{\mathrm{Lya}} \simeq 1.83$), 0.12 for the sestet ($z_{\mathrm{Lya}} \simeq 1.83$) and 0.21 for pairs A and Q ($z_{\mathrm{Lya}} \simeq 2.34$). The chosen values are within 1 per cent of the original values of the simulated spectra. In the computation of the correlation functions, we have used the average fluxes obtained as the mean of the flux values along the line of sight as for the observed spectra. We checked that the difference in the resulting correlation functions taking the true values for the average flux is less than 5 per cent.

Finally, we added to the simulated spectra a Gaussian noise in order to reproduce the observed average S/N (per pixel): S/N = 5 for the triplet, the sestet and pair A; S/N = 9 for pair U and S/N = 15 for pair Q.

3 CORRELATION PROPERTIES OF THE IGM: GENERAL PROCEDURE

On the basis of the interpretation of the Lyman α forest as due to a continuous density field with a univocal correspondence between density and transmitted flux, we computed the correlation properties of the transmitted flux in QSO lines of sight and regarded them as indicators of the correlation properties of matter in the IGM.

We selected in each normalized spectrum the region between the Lyman β emission (or the shortest observed wavelength, when the Lyman β was not included in the spectrum) and $5000\,\mathrm{km\,s^{-1}}$ from the Lyman α emission (to avoid proximity effect due to the QSO). Absorption lines due to ions of elements heavier than hydrogen 'contaminate' the Lyman α forest and can give spurious contribution to the clustering signal (see Kim et al. 2004, for a discussion in the case of single lines of sight). We flagged and removed the spectral regions where metal lines and Lyman α absorptions of damped and subdamped systems occurred inside the Lyman α forest.

Table 3. Meaning of the symbols used in the following figures.

	LP sample	Pair sample	Simulations
Auto-correlation	Cross	Solid triangle	Empty triangle
Cross-correlation		Solid square	Empty square
Cross-correlation coeff.		Solid dot	Empty dot

Given the normalized transmitted flux, f, as a function of the velocity $v_\parallel$ along the line of sight and the angular position θ on the sky, we define $\delta_f = (f - \bar{f})$, where the average flux, $\bar{f}$, is computed for every spectrum as the mean of the transmitted flux over all the considered pixels in that spectrum. We neglected the redshift evolution of the average transmitted flux in the Lyman α forest of the individual spectra, which translates into the redshift evolution of the mean H I opacity of the Universe (Kim, Cristiani & D'Odorico 2002; Schaye et al. 2003; Viel et al. 2004b), because we verified that its effect on the correlation function is negligible.

Table 3 provides details about the symbol convention used to plot our results.

4 THE FLUX AUTO-CORRELATION FUNCTION

The unnormalized auto-correlation function of the flux along the line of sight is defined as

$$\xi_\parallel^{\rm f}(\Delta v_\parallel) = \langle \delta_f(v_\parallel, \theta)\delta_f(v_\parallel + \Delta v_\parallel, \theta)\rangle, \quad (3)$$

following previous studies on the same subject (e.g. McDonald et al. 2000; Rollinde et al. 2003; Becker et al. 2004). The auto-correlation function for our sample of QSO spectra was obtained by averaging over all the pixels of all the QSOs. The results were binned in $50\,{\rm km\,s^{-1}}$ velocity bins. The 1σ error on this measure is estimated extracting 50 samples of 15 objects from our sample with a bootstrap method, computing the auto-correlation function for each sample and determining the standard deviation of the distribution.

The obtained correlation function is in very good agreement with the analogous determination by Rollinde et al. (2003) and in qualitative agreement with the Lyman α lines two point correlation function at similar redshifts (Kim et al. 2001).

The same procedure was adopted to compute $\xi_\parallel^{\rm f}$ for the QSO spectra of the control sample taken from the UVES LP. These spectra have a much larger S/N $\sim$ 50 in the Lyman α forest region) than our spectra, they have been selected to be free from damped Lyman α absorptions and they do not belong to known close QSO groups.

The auto-correlation function for the simulated spectra was computed as the arithmetic mean of the correlation functions obtained for 50 realizations of the observed sample and the error is the corresponding standard deviation. It is important to recognize that the computed error bars both for the observed and simulated $\xi_\parallel^{\rm f}$ are strongly correlated. This is due to the fact that every pixel contributes to the correlation function in several velocity bins.

In Fig. 3, we show the comparison of the three auto-correlation functions in velocity space, error bars are included just to give a qualitative idea of their size.

The agreement at the 3σ level is good. However, we would like to point out two kinds of discrepancies at the 1σ level which, in our opinion, are the signatures of two interesting effects.

(1) Both observed correlation functions have an amplitude in the first bin (centred at $50\,{\rm km\,s^{-1}}$) which is larger than the simulated one. We ascribe this to a scantiness of strong lines ($\log N({\rm H\,I}) \gtrsim 15.5$) in the simulated spectra (see e.g. Theuns et al. 2002) due to the limited extent of simulations in comoving space, which implies that large overdensities (giving rise to strong Lyman α lines) are not appropriately probed. We found support to this hypothesis from the analysis in Sections 4.1 and 6.1.1.

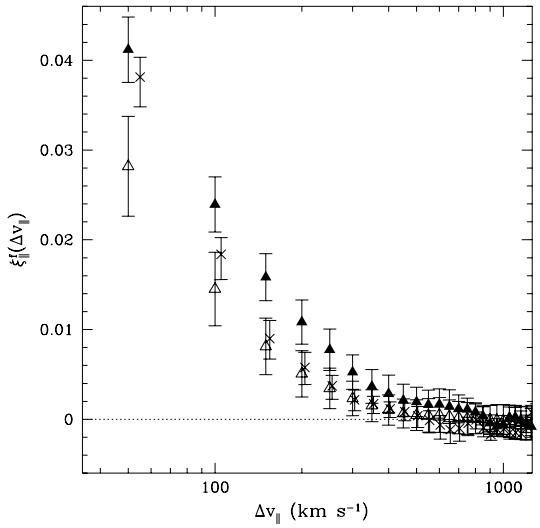

Figure 3. Comparison of the auto-correlation function for our sample of observed QSO spectra (solid triangles), for the LP sample of observed QSO spectra (crosses, shifted in velocity for clarity) and for the simulated sample of spectra (empty triangles) as a function of velocity separation along the line of sight. Error bars on $\xi_\parallel^{\rm f}$ are highly correlated and they are plotted only to give an idea of their extent.

(2) The correlation function computed with our sample of QSOs tends to show a stronger clustering than the simulated one up to large separations, while the LP correlation function is compatible with the mock correlation function already at the second bin. The main difference between the two samples is the fact that our QSOs were selected to belong to close pairs and associations, thus they could trace overclustered regions. Although we excluded the spectral portion within $5000\,{\rm km\,s^{-1}}$ of the QSO emission, an enhanced clustering signal could still be present. We discuss this hypothesis in Section 4.2.

We notice that cosmic variance could also be at the origin of these small differences. Indeed auto-correlation functions of single observed QSO spectra show large variations among them and we are working with relatively small samples. On the other hand, all the simulated spectra have been drawn from the same box of $120\,h^{-1}$ comoving Mpc, which although notably large for a simulation, possibly underestimates the cosmic variance (see McDonald et al. 2000, for possible ways of correcting the simulated correlation function for errors induced by the finite box size).

4.1 Contribution of the stronger absorptions

Lyman α lines of increasing H I column density correspond to structures of increasing overdensity and smaller characteristic size. Going from $\log N({\rm H\,I}) \sim 12$ to $\log N({\rm H\,I}) \gtrsim 15$ we probe from underdense, modestly overdense regions, to highly overdense filaments and their intersections. We expect that the major contributors to the clustering signal are the stronger absorption lines, which trace the higher density peaks. The aim of this section is two-fold. On the one hand, by excluding the stronger lines we want to verify the clustering properties of the gas in the 'true' IGM, far from the densest

regions. On the other hand, we want to test the relative behaviour of the observed and simulated auto-correlation functions without those lines.

Note that a careful study of the impact of strong absorption systems on the flux power spectrum was performed by Viel et al. (2004a) and McDonald et al. (2005).

To select the strong absorption lines without carrying on the time-consuming Voigt fitting procedure of the observed and simulated spectra, we can relate the transmitted flux values to a corresponding column density assuming a typical Doppler parameter. We eliminated the pixels with $f < f_{\rm thres} = 0.1$ or $\tau_{\rm H\,I} \gtrsim 2.3$. This approximately corresponds to excluding the absorption lines with $\log N({\rm H\,I}) \gtrsim 13.96$, if we adopt the formula for the optical depth at the centre of the line:

$$\tau_{0,{\rm Ly}\alpha} = 7.58295 \times 10^{-13} \frac{N\,({\rm cm}^{-2})}{b\,({\rm km\,s}^{-1})}, \quad (4)$$

and a Doppler parameter $b = 30\,{\rm km\,s}^{-1}$ characteristic of the Lyman α forest absorptions at redshift $z \sim 2$ (Kim et al. 2001).

This technique is efficient for data with ${\rm S/N} \gtrsim 1/f_{\rm thres} = 10$, so we applied the selection only to the LP QSO sample and to the sample of simulated spectra for which, in this case, we adopted a Gaussian noise giving ${\rm S/N} = 50$.

Fig. 4 shows the resulting correlation functions. The amplitude of the observed $\xi^{\rm f}_{\parallel}$ has reduced to about one third of the initial value but still shows a significant clustering signal. This is an indication that matter is still clustered also far from the most overdense regions. The threshold on $\tau_{\rm H\,I}$, transformed with equation (2), corresponds to $\rho/\bar{\rho} \lesssim 6.5$.

What is worth noting is that the amplitude of the simulated correlation function has decreased of about one half and is now fully compatible with the observed correlation function. This suggests that strong lines are deficient in the simulated spectra so that their contribution to the correlation function is smaller than for the observed spectra. When they are excluded from the computation the two correlation functions become consistent.

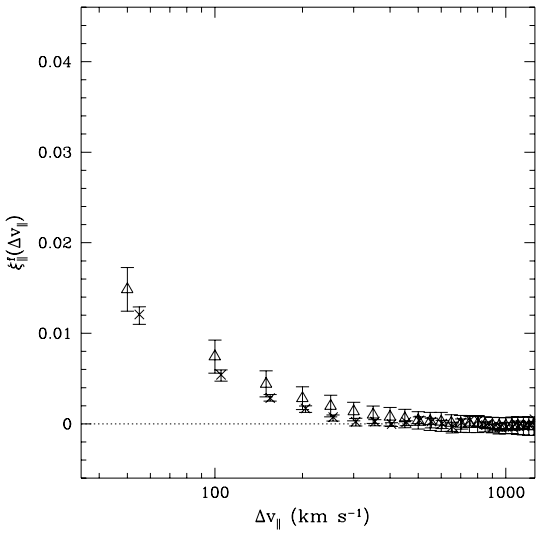

Figure 4. Same as Fig. 3 but selecting only the pixels with flux value $f > 0.1$ both in the observed and simulated spectra. This is equivalent to exclude the lines with column density $\log N({\rm H\,I}) \gtrsim 13.96$.

4.2 Selection effects in the observed spectra

There is the possibility that our observed sample is affected by a selection bias since the QSOs were required to be at close angular separations and, in most cases, they have similar emission redshifts. As a consequence, the observed lines of sight could be biased toward more clustered regions than average.

We tried to estimate this effect on the auto-correlation function by excluding larger and larger portions of the observed spectra from the QSO emissions (up to $20\,000\,{\rm km\,s}^{-1}$). No significant decrease of the amplitude of the correlation function has been detected.

This test could prove inadequate as far as the sestet is concerned since the redshifts of some member QSOs are separated by up to $\sim$50 000 km s^{-1} (see Fig. 1). As a consequence, our measure could be affected by the transverse proximity effect (e.g. Jakobsen et al. 2003; Worseck & Wisotzki 2006) or by possible clustering associated with the presence of a foreground QSO. Indeed, in the case of pair Q, we found a damped Lyman α system in the higher redshift QSO spectrum at the redshift of the other QSO (D'Odorico et al. 2002). From a preliminary analysis of the sestet, strong Lyman α absorptions ($\log N({\rm H\,I}) \sim 15$) are observed in the spectra of S3, S5 and S6 at $\sim$600–750 km s^{-1} from the emission redshift of the lower redshift QSO, S1, and a strong Lyman α with associated C IV absorption is present in the spectrum of S2, the furthest away QSO, at $\sim$1000 km s^{-1} from the redshift of S1.

Considering the other foreground object, S4, Lyman α absorptions within 200 km s^{-1} of its redshift are observed along the line of sight to S2 ($\log N({\rm H\,I}) \sim 14$) and S3 ($\log N({\rm H\,I}) \sim 15$) without associated metals detected. As already mentioned in the Introduction, we will carry out a detailed study of the coincidences among multiple lines of sight in a future paper, with data of improved S/N.

In the following, we will use the auto-correlation function determined from the LP sample to avoid systematic uncertainties introduced by the above mentioned possible selection effect.

5 CLUSTERING OF THE FLUX ACROSS THE LINE OF SIGHT

In this section we exploit the potentialities of our sample of QSO pairs by determining the clustering properties of the IGM across the lines of sight. The great advantage with respect to the correlation function along the line of sight, in particular for a sample like ours showing a large variety of pair separations, is that we have the guarantee of sampling true spatial separations between the pixels, the effect of peculiar velocities being negligible or absent.

5.1 The cross-correlation function in redshift space

As a first approach, we computed the cross-correlation function extending in a natural way the procedure adopted for the auto-correlation function.

Every pixel along the line of sight is considered as an element of the density field at the QSO angular position in the sky and at a distance from the observer (comoving along the line of sight) corresponding to the redshift of the pixel:

$$r_\parallel(z) = \frac{c}{H_0} \int_0^z \frac{{\rm d}z'}{E(z')}, \quad (5)$$

where $E(z)$ is defined as

$$E(z) = \sqrt{\Omega_{0{\rm m}}\,(1+z)^3 + \Omega_{0\Lambda}}. \quad (6)$$

In equation (5), there is the implicit hypothesis that peculiar velocities contribute negligibly to the measured redshift in the Lyman α forest. This statement is supported by recent measurements using QSO pairs (Rauch et al. 2005), and we will see a posteriori that the comparison of the auto- and cross-correlation functions makes this procedure allowable.

The cross-correlation function of the transmitted flux between two lines of sight at angular separation $\Delta\theta$ is defined as

$$\xi^{\rm f}_{\times}(\Delta r) = \langle \delta_f(\theta, r_{\parallel,1}) \delta_f(\theta + \Delta\theta, r_{\parallel,2}) \rangle, \quad (7)$$

where $\Delta r = \sqrt{(r^2_{\parallel,1} + r^2_{\parallel,2} - 2\, r_{\parallel,1}\, r_{\parallel,2}\, \cos\Delta\theta)}$ is the spatial separation between pixel 1 at $r_{\parallel,1}$ along one line of sight and pixel 2 at $r_{\parallel,2}$ along the paired line of sight.

The error on the correlation function cannot be estimated with a bootstrap technique, as done before, due to the limited number of pairs contributing to each separation (in particular at the smaller separations). In order to give an evaluation of the significance of the observed signal, the same cross-correlation function was computed replacing one QSO in each pair with a control QSO, then repeating the operation replacing the other QSO of every pair with another control QSO. The control sample is formed by the LP QSO spectra. We chose LP QSOs at redshifts close to those of the original objects and shifted the spectra in order to match it exactly. We derived an indicative error as the rms deviation from zero of the two control functions in two regions: $\Delta r < 5\, h^{-1}$ Mpc, for which the uncertainty is larger, and $\Delta r > 5\, h^{-1}$ Mpc.

We computed the cross-correlation function also for the sample of mock spectra. The simulated spectra are characterized by the redshift of the output box and a velocity extent (see Section 2.2). In order to assign a redshift value to every pixel, we gave the central pixel of every spectrum the redshift of the corresponding output box, then we numbered the pixels one by one transforming the velocity size of the pixel into a redshift size. Once the redshifts were determined pixel by pixel, we followed the same procedure adopted for the observed spectra for the 50 simulated samples and computed the average cross-correlation function and its 1σ standard deviation.

The result of our computation is shown in Fig. 5 compared with the observed cross-correlation function of the pairs and of the control sample. There is a very good agreement between the two functions. Only in the first bin there is an indication of a lower value for the simulated cross-correlation with respect to the observed one. This is consistent with the result obtained for the auto-correlation function if we take into account the fact that the first bin of $\xi^{\rm f}_{\times}$ corresponds to the second bin of $\xi^{\rm f}_{\parallel}$.

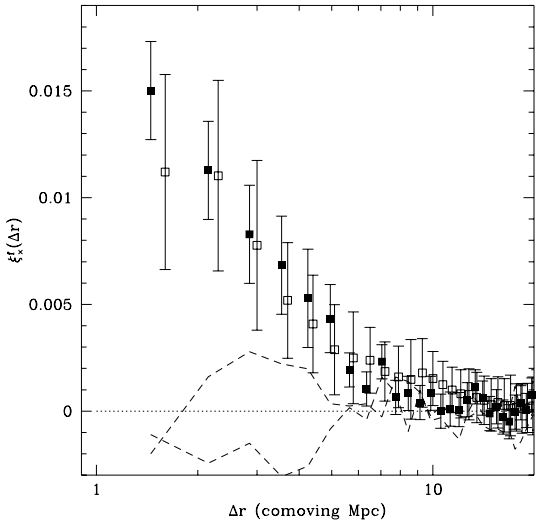

Figure 5. Cross-correlation function of the transmitted flux for our sample of pairs and groups of QSOs (solid squares) compared with the analogous correlation function computed with the sample of mock spectra (empty squares). The error bars on the simulated data are 1σ standard deviations for 50 simulated samples while the error bars on the observed data have been obtained from the rms deviation of the two control samples plotted as dashed lines (see text). Simulated data are slightly shifted in Δr for clarity.

5.2 The cross-correlation coefficient

A measure of the transverse clustering properties of the IGM which is less affected by peculiar velocities is the flux cross-correlation coefficient,

$$\chi^{\rm f}_{\times}(\Delta\theta) = \langle \delta_f(\theta, v_{\parallel}) \delta_f(\theta + \Delta\theta, v_{\parallel}) \rangle, \quad (8)$$

where every pixel along one line of sight is correlated with the one face-to-face in redshift space along the paired line of sight and the result is averaged over all the pixels in the common redshift interval.

Every pair of QSOs at angular separation $\Delta\theta$ gives one value of $\chi^{\rm f}_{\times}(\Delta\theta)$, and a sample with several pairs at different separations, as is our sample, gives an estimate of the correlation function. At a given redshift, the angular separation $\Delta\theta$ corresponds to a velocity separation $\Delta v_{\perp} = cF(z)\,\Delta\theta$, where c denotes the speed of light, and $F(z)$ is a dimensionless function of redshift that includes all the dependence on the global cosmological metric. In the cosmological model that we have adopted:

$$F(z) = \frac{E(z) \int_0^z [{\rm d}z/E(z)]}{(1+z)}, \quad (9)$$

and $E(z)$ is defined in equation (6).

We computed the range of velocity separations covered by each of our pair of spectra then we grouped the pairs in velocity bins of variable width and computed the average cross-correlation coefficient for every group. Given the small number of pairs in every group (a maximum of three QSO pairs) the uncertainties reported on these determinations are computed by applying a simple error propagation to equation (8), thus they account only for the pixel statistics and the noise associated with the transmitted flux but they are not representative of the true error due to the cosmic variance.

In the case of the simulated spectra, we had 50 realizations of each of our QSO pairs so we could obtain in every velocity interval defined for the observed pairs an average cross-correlation coefficient with its error, that in this case is the standard deviation of the distribution of values.

Results are shown in Fig. 6. The two points at the smallest separations are given by pair U + S5S6 and pair A, respectively. The two samples of values are in very good agreement except for the bin at $\Delta v_{\perp} \sim 500\,{\rm km\,s^{-1}}$ ($\simeq$5 arcmin at $z \sim 1.8$) for which the observed value is $\sim 7\sigma$ above the simulated value. The observed $\chi^{\rm f}_{\times}(\Delta\theta)$ in this velocity bin is obtained by averaging two pairs: S1S3 and T1T3 for which $\chi^{\rm f}_{\times} \simeq 0.040 \pm 0.003$ and 0.0177 ± 0.0015, respectively.

The two pairs of spectra were searched for peculiar features that could boost the signal by selecting those pixels for which $\delta_f(\theta, v_{\parallel})\delta_f(\theta + \Delta\theta, v_{\parallel}) > 0.4$. In Fig. 7, we show the results of this research. In the case of S1 (up left-hand plot) about 44 per cent of selected pixels have values $f > 2$ or $f < 0$, that is they are strongly affected by noise (in the case of S3 only $\sim$14 per cent of pixels are in those range of values). As a consequence, the large value of the cross-correlation coefficient for pair S1S3 could be due to the

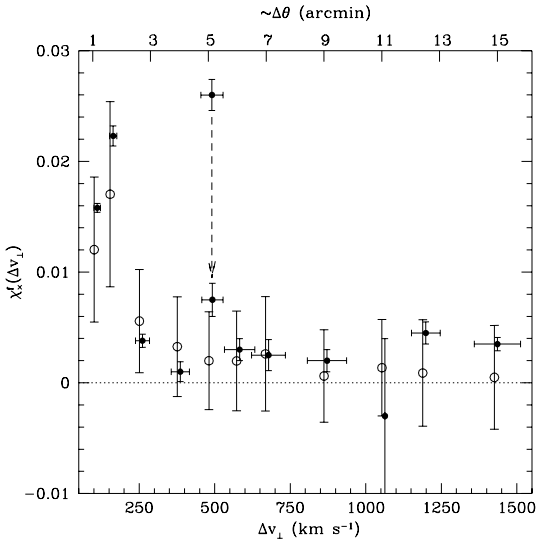

Figure 6. Comparison of the cross-correlation coefficients for our sample of observed spectra (solid dots) and of simulated ones (empty dots) as a function of the velocity separation, $\Delta v_\perp = c\,F(z)\,\Delta\theta$, corresponding to the angular separation, $\Delta\theta$, of the QSO pairs. The angular separation computed at $z = 1.8$ is reported in the top axes. Observed values are shifted by $10\,\mathrm{km\,s^{-1}}$ for clarity. The lower observed value at $\Delta v_\perp \simeq 500\,\mathrm{km\,s^{-1}}$ was obtained excluding pair S1S3 and removing a strong coincident Lyman α line in the T1T3 pair (see text). Error bars on the observed values along the x-axis represent the velocity range covered by the considered pairs. Errors on the $\chi^{\mathrm{f}}_{\times}$ represent the uncertainty on that measure for the observed values and the cosmic variance for the simulated values (see text).

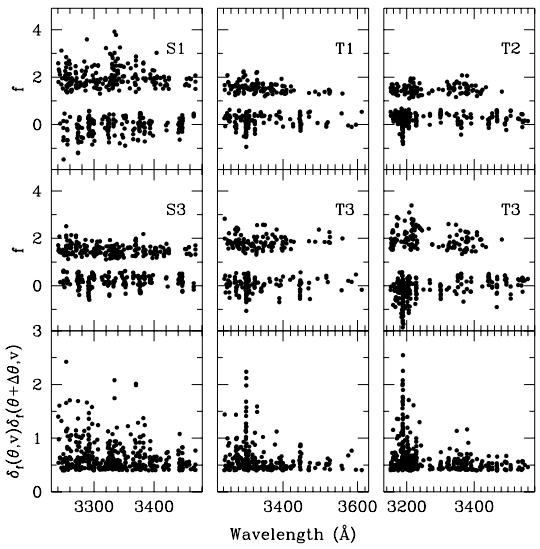

Figure 7. Selection of the pixels giving values $\delta_f(\theta, v_\parallel)\,\delta_f(\theta + \Delta\theta, v_\parallel) > 0.4$ for the pairs S1S3 (left), T1T3 (centre) and T2T3 (right). The two upper plots show the values of the flux for the selected pixels in the indicated QSO while the lower plot report the value of the above product.

low S/N in the spectrum of S1 and cannot be determined reliably. On the other hand, in the case of pair T1T3, Fig. 7 (centre plots) shows a feature giving a strong signal at $\lambda \simeq 3300$ Å. We identified in the spectra of T1 and T3 two coincident strong H I Lyman α absorptions ($\log N(\mathrm{H\,I}) \sim 15$) at this wavelength, the former with a clear associated C IV doublet while for the second there is the possible detection of the C IV λ 1548 line only. No corresponding H I Lyman α line is observed at this redshift along the third line of sight of the triplet, T2. If we mask the redshift interval covered by the coincident lines and recompute $\chi^{\mathrm{f}}_{\times}$ for T1T3 we get 0.0075 ± 0.0015, which is in better agreement with the simulated value as shown in Fig. 6.

In the right-hand plots of Fig. 7, we show the case of pair T2T3 (at 8.9 arcmin angular separation) where a coincident absorption system is present with characteristics very similar to the one in pair T1T3, and the cross-correlation coefficient has a value consistent with zero. This suggests that the presence of a coincident absorption system could be a necessary but not sufficient condition to explain a large value of $\chi^{\mathrm{f}}_{\times}$. It is interesting to note that also the recently published two point correlation function of the C IV absorptions in the LP QSO spectra (Scannapieco et al. 2006) shows an excess of clustering signal at $\sim$500 km s^{-1} whose origin has not been satisfactorily explained.

New observations are going to be carried out with UVES to increase the S/N in the two QSO pair spectra S1S3 and T1T3, in order to confirm the large value of the cross-correlation coefficient and possibly detect other coincident C IV systems. On the other hand, we will try and collect other QSO pairs at the same velocity separation to increase the statistics and we will look into physical mechanisms that could explain our results.

In Fig. 8, we compare the cross-correlation coefficients with the cross-correlation function. The angular separation $\Delta\theta$ between two QSO lines of sight was transformed into a comoving spatial separation, Δr, with the formula

$$\Delta r = \frac{c\Delta\theta}{H_0}\int_0^z \frac{\mathrm{d}z'}{E(z')}, \tag{10}$$

where $E(z)$ was defined in equation (6). Considering the error bars computed from the simulated pairs, there is a good agreement within 1σ. The large variations from one data point to the other in the cross-correlation coefficients should be due mainly to the small number of QSO pairs (between one and three) contributing to each point. On the other hand, the smoothness of $\xi^{\mathrm{f}}_{\times}$ is artificially increased by the fact that the values in the different bins are not independent.

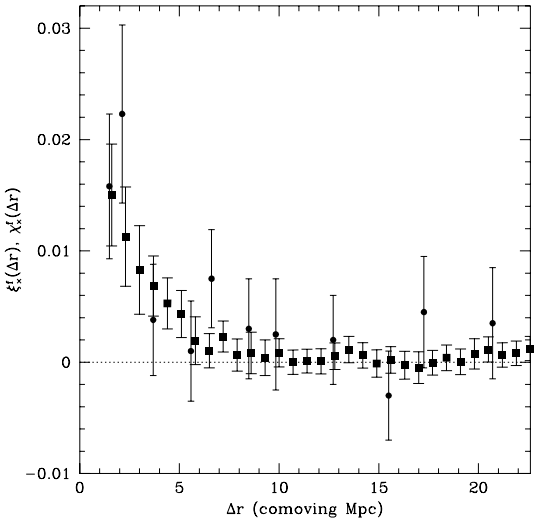

Figure 8. Cross-correlation function of the transmitted flux, $\xi^{\mathrm{f}}_{\times}$, for our sample of pairs and groups of QSOs (squares, slightly shifted in Δr) compared with the observed cross-correlation coefficients, $\chi^{\mathrm{f}}_{\times}$ (dots), as a function of spatial separation (see text). Error bars both on $\chi^{\mathrm{f}}_{\times}$ and $\xi^{\mathrm{f}}_{\times}$ have been determined from simulations.

6 DISCUSSION

From the cosmological point of view, one of the most interesting and challenging applications of the kind of calculations carried out in this paper is the constraint on the geometry of the Universe (in particular, the estimate of $\Omega_{0\Lambda}h^{-2}$) by linking angular separations and redshift differences in the hypothesis that the observed correlation properties are isotropic (Alcock & Paczyński 1979; Hui, Stebbins & Burles 1999; McDonald & Miralda-Escudé 1999).

A large number of QSO pairs at different angular separations (see McDonald 2003) is needed to obtain a measure of the cosmological parameters as accurate as those recently produced by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) team (Spergel et al. 2003). This is why we do not attempt to derive a measure of $\Omega_{0\Lambda}$ with the present sample.

Nevertheless, a comparison of the observed auto- and cross-correlation functions is interesting to qualitatively evaluate the distortion due to peculiar velocities and the correctness of the adopted cosmological parameters. The two functions are shown in Fig. 9. Note that, due to the minimum angular separation between our QSO pairs, there is no transverse clustering signal corresponding to the first bin of the auto-correlation (centred at $\Delta v = 50\,\mathrm{km\,s^{-1}}$ or $\Delta r \simeq 740$ kpc). The correspondence between $\xi^{\mathrm{f}}_{\parallel}$ and $\xi^{\mathrm{f}}_{\times}$ is very good suggesting that the adopted cosmological parameters are reasonable and that peculiar velocities do not play a major role in the IGM gas on velocity scales $\gtrsim 100\,\mathrm{km\,s^{-1}}$.

In the next section, we survey the behaviour of the simulated correlation functions when varying three parameters characterizing the physical state of the gas.

6.1 Exploring the parameter space of the simulations

6.1.1 Effective optical depth

As already explained in Section 2.2, the effective optical depth of a simulated spectrum can be rescaled to a new value simply by increasing or decreasing the optical depth of every pixel in the spectrum by a constant factor, which is equivalent to increase or decrease the column density of the absorption lines in the spectrum.

Table 4. Effective optical depth adopted to build the three samples of mock spectra used to compute the correlation functions in Figs 10 and 11.

QSO	z	τ_{min}	τ_{obs}	τ_{max}
Sestet	1.8	0.08	0.12	0.17
Triplet	1.8	0.08	0.10	0.17
Pair U	1.8	0.08	0.10	0.17
Pair A	2.4	0.17	0.21	0.29
Pair Q	2.4	0.17	0.21	0.29

Considering previous observations (Viel et al. 2004b) and the relations computed by Schaye et al. (2003) for the redshift evolution of the effective optical depth, we derived 3σ upper and lower limits for τ_{eff} at the redshifts of the simulated spectra, which are reported in Table 4. Then, we built two new samples of mock spectra starting from the original spectra produced from the simulation, one with the minimum and one with the maximum values of τ_{eff} and recomputed the auto- and cross-correlation functions. Each sample is formed by 50 realizations of the observed sample of QSO pairs and we determined the average value for the correlation function.

Our results are shown in Figs 10 and 11. At the smallest scale, probed only by the auto-correlation function, the agreement between observed and simulated clustering signal improves adopting the larger effective optical depths. However, at the Jeans scale, corresponding to the first bin of $\xi^{\mathrm{f}}_{\times}$ and to the bin centred at $100\,\mathrm{km\,s^{-1}}$ of $\xi^{\mathrm{f}}_{\parallel}$, the larger τ_{eff} possibly overestimates the clustering amplitude.

Since the effective optical depth measured in the original simulation is in very good agreement with the one observationally determined from QSO Lyman α forests (see Section 2.2), we infer that the discrepancies between simulations and observations are not due to a wrong estimate of τ_{eff} but, as already mentioned in Section 4,

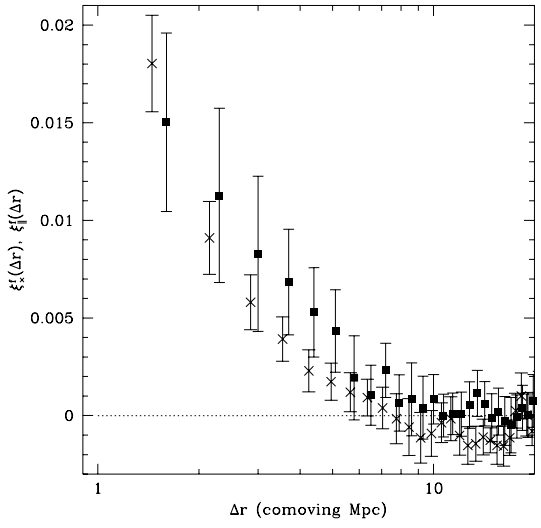

Figure 9. Comparison of the cross-correlation function for our sample of QSO pairs (squares) with the auto-correlation function computed for the LP QSO sample (crosses) as a function of comoving spatial separation across and along the line of sight, respectively. The cross-correlation function is slightly shifted in Δr for clarity.

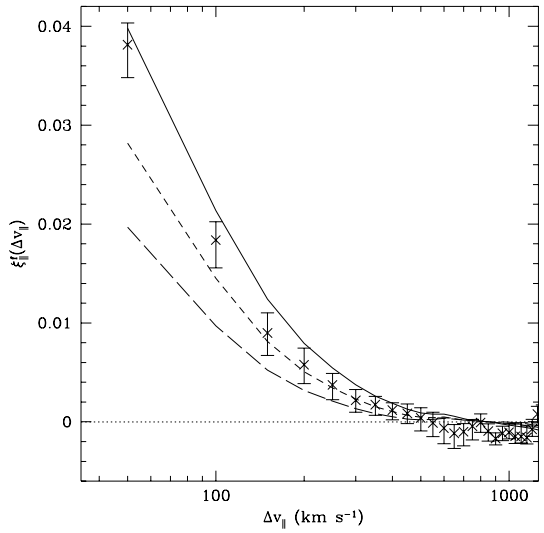

Figure 10. Auto-correlation function for the LP sample (crosses) related to the average auto-correlation functions of three samples of simulated spectra with different τ_{eff}. The long dashed line traces the correlation function for the sample with the minimum optical depths, the dashed line is for the sample with the observed optical depths and the solid line is for the sample with the maximum optical depths (see Table 4).

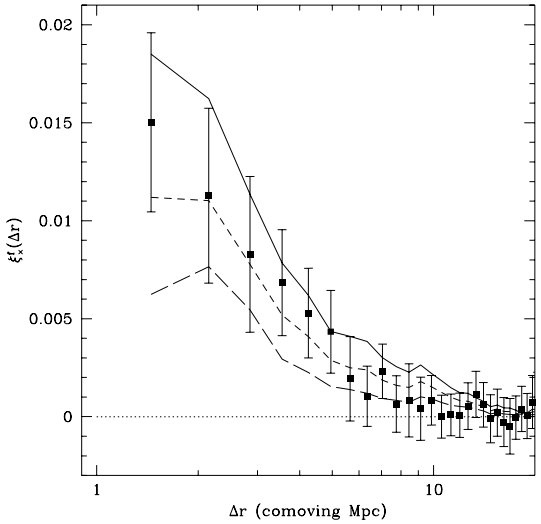

Figure 11. Cross-correlation function for our sample of QSOs (solid squares) compared with the average cross-correlation functions of three samples of simulated spectra with different $\tau_{\rm eff}$. The long dashed line traces the correlation function for the sample with the minimum optical depths, the dashed line is for the sample with the observed optical depths and the solid line is for the sample with the maximum optical depths (see Table 4).

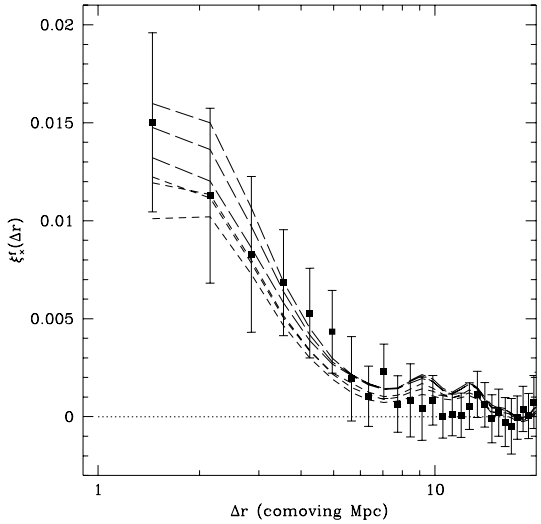

Figure 12. Cross-correlation function for our sample of pairs (solid squares) compared with the average cross-correlation functions of samples of simulated spectra with different temperature and exponent γ of the temperature–density relation (equation 1). Short dashed lines are for simulations in which $\gamma = 1.6$ and temperature is $T = 8000$, 22 000 and 50 000 K going from the upper to the lower curve. The long dashed lines represents simulations with $\gamma = 1.1$ while the temperatures are the same as before but they increase from the lower to the upper curve.

to the deficiency of strong lines in the simulations. In the spectra with the larger $\tau_{\rm eff}$ the column density of all the lines is increased and the average flux is decreased, causing a general increase of the amplitude of the correlation function, partially compensating the small number of strong lines.

In order to verify our statement, a quantitative analysis of the difference in the Lyman α column density distribution function (the number of absorption lines per unit column density per unit redshift) of observed and simulated spectra (as done by Theuns et al. 2002, for small simulated boxes) is necessary. Furthermore, more pairs at small separations (<1 arcmin) would allow to decrease the error bars and test if the behaviour of clustering is the same as along the line of sight.

6.1.2 Temperature and γ

The influence on the flux cross-correlation function of the temperature and of the power-law index of the temperature–density relation adopted for the gas (see equation 1) are explored. The aim of this analysis is to point out the modifications of the simulated correlation function when the parameters are varied, and not to derive precise constraints on the same parameters from the comparison with observations. With a running time of 2–3 weeks for simulation, it is impossible to run an extensive parameter study for different thermal histories, thereby we decided to rescale a posteriori the temperature–density relation at a given output by assuming that all the gas particles obey this relation. We then recomputed the ionization fractions for each gas particle with the proper UV background. This approximation does not account for the effects that the corresponding change in the gas pressure would have on the gas distribution. However, we explicitly checked that the differences in the correlation function for the fiducial run and for a rescaled temperature relation with the same $T - \gamma$ are of the order of 10 per cent (see Viel et al. 2004b, for the effect of the rescaling on the flux power spectrum).

We built six sets of simulated spectra each one formed by 50 realizations of the observed sample of pairs. Every set is characterized by a value of γ (1.1 or 1.6) and a value of the temperature (8000, 22 000 and 50 000 K), while the effective optical depth is the one used for the main sample (refer to Section 2.2).

In Fig. 12, we plot together the observed $\xi^{\rm f}_\times$ with the average correlation functions computed from the six groups of simulated spectra. At the smallest separation probed by the observations, the predictions of the different models are well separated and could be tested with a larger sample of close QSO pairs. The models with $\gamma = 1.1$ give a larger clustering signal than those with $\gamma = 1.6$ for all temperatures and are in better agreement with the data at the small scales. An improved concordance between observations and simulations for lower values of γ was observed also in the case of the flux probability distribution function estimated from the Lyman α forest (James Bolton, private communication).

The exponent of the temperature–density relation is determined by the ionization history of the gas. Hui & Gnedin (1997) demonstrated that instant reionizations of H I occurring at redshifts decreasing from $z = 10$ to 5 would imply shallower and shallower γ at redshifts $z \sim 3$. However, also for very late reionizations γ should have increased above values $\sim$1.4 by $z \sim 2$. Ricotti, Gnedin & Shull (2000) studied the evolution with redshift of the equation of state of the gas and showed that a reionization of He II occurring at $z \sim 3$ would cause a sudden increase in the gas temperature and a corresponding decrease in the value of γ (their fig. 13). Observationally, this has been studied mainly by using the distribution of Doppler parameters and column densities of the Lyman α forest lines (e.g. Schaye et al. 2000; Kim et al. 2002), but also with metal absorptions (e.g. Songaila 1998; Vladilo et al. 2003). Our result supports previous observational evidences of a second reheating of the Universe happening at $z \sim 3$, most likely due to the reionization of He II and it requires confirmation by a larger sample of observed spectra and a set of more refined mock spectra.

7 CONCLUSIONS

In this paper, we exploited the capabilities of the largest sample of high-resolution UVES spectra of QSO pairs to study the three-dimensional distribution properties of baryonic matter in the IGM as traced by the transmitted flux in the QSO H I Lyman α forests. Our sample is formed by 21 QSO pairs evenly distributed between angular separations of $\sim$1 and 14 arcmin, with Lyman α forests at a median redshift $z \simeq 1.8$. We selected also eight UVES QSO spectra from the ESO Large Program 'The Cosmic Evolution of the IGM' (Bergeron et al. 2004) to compute the correlation function along the line of sight and to be used as a control sample for the cross-correlation function (see Section 5.1). We compared the observed sample with a set of mock spectra drawn from a cosmological hydrosimulation run in a box of $120\,h^{-1}$ comoving Mpc, adopting the cosmological parameters of the concordance model. The simulated sample reproduces 50 different realizations of the observed sample.

In the following, we resume our main results.

(1) The clustering properties of matter in the IGM are the same in the direction parallel and transverse to the line of sight when using the parameters of the concordance cosmology to transform the angular distance into velocity separation. As an implication, peculiar velocities in the absorbing gas are likely smaller than $\sim$100 km s^{-1}.

(2) Matter in the IGM is clustered on scales smaller than $\sim$200 km s^{-1} or about $3\,h^{-1}$ comoving Mpc. We verified that the clustering signal is significant also for the slightly overdense gas ($\tau_{\rm HI} \gtrsim 2.3$ or $\rho/\bar{\rho} \lesssim 6.5$) although on smaller scales.

(3) The simulated correlation functions are consistent with the observed analogous quantities at the 3σ level, although they systematically predict lower clustering at the smaller scales. The agreement becomes better when only the lower density regions are selected for the computation or when the effective optical depth of the simulated spectra is fixed to a larger value (marginally consistent with previous extensive observational results on the redshift evolution of the effective optical depth). These are indications of a deficiency of strong absorption lines in the simulated spectra that needs further investigation.

(4) We observed an improved consistency between observations and simulations also when a lower γ is adopted in the equation of state of the gas, $T = T_0\,(\rho/\bar{\rho})^{\gamma-1}$($\gamma = 1.1$ instead of the standard 1.6). This result hints to a late He II reionization epoch whose effects on the IGM could still be measured at the redshifts investigated here.

(5) We measured an enhanced clustering signal for the cross-correlation coefficient at a transverse velocity separation $\Delta v_\perp \sim 500$ km s^{-1}. This velocity distance matches the line split of the C IV doublet. We propose to gather more observational and theoretical material to shed light on this result.

From the analysis carried out in the present study, we evidenced the need of increasing our sample of observed QSO pairs, in particular at the small angular separations ($\sim$1 arcmin), but also at $\Delta v_\perp \simeq 500$ km s^{-1} in order to confirm the overdensity of coincident Lyman α lines at this transverse distance. On the other hand, we realized the necessity of a dedicated effort on the comparison between observed and simulated QSO spectra, in particular as far as the statistics of absorption lines is concerned (number density, Doppler parameter and column density distribution etc.).

ACKNOWLEDGMENTS

We are indebted to an anonymous referee whose comments were extremely helpful to improve the clarity of this paper. We are grateful to Pierluigi Monaco for stimulating comments and a critical reading of the manuscript. It is our pleasure to thank Evan Scannapieco and Emmanuel Rollinde for interesting discussions. SL was partly supported by the Chilean *Centro de Astrofísica* FONDAP No. 15010003, and by FONDECYT grant N$^{\rm o}$1030491. The simulations were run on the COSMOS (SGI Altix 3700) supercomputer at the Department of Applied Mathematics and Theoretical Physics in Cambridge. COSMOS is a UK-CCC facility which is supported by HEFCE and PPARC.

REFERENCES

Alcock P., Paczyński B., 1979, Nat, 358, 359
Ballester P., Modigliari A., Boitquin O., Cristiani S., Hanuschik R., Kaufer A., Wolf S., 2000, The Messenger, 101, 31
Bechtold J., Crotts A. P. S., Duncan R. C., Fang Y., 1994, ApJ, 437, L83
Becker G. D., Sargent W. L. W., Rauch M., 2004, ApJ, 613, 6
Bergeron J. et al., 2004, The Messenger, 118, 40
Bi H.-G., Davidsen A. F., 1997, ApJ, 479, 523
Bolton J. S., Haehnelt M. G., Viel M., Springel V., 2005, MNRAS, 357, 1178
Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, ApJ, 437, L83
Coppolani F. et al., 2006, MNRAS, 370, 1804
Cristiani S., D'Odorico S., D'Odorico V., Fontana A., Giallongo E., Savaglio S., 1997, MNRAS, 285, 209
Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44
Croft R. A. C., Weinberg D. H., Pettini M., Hernquist L., Katz N., 1999, ApJ, 520, 1
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, MNRAS, 349, 1397
Crotts A. P. S., Fang Y., 1998, ApJ, 502, 16
Davé R., Hernquist L., Weinberg D. H., Katz N., 1997, ApJ, 477, 21
Dekker H., D'Odorico S., Kaufer A., Delabre B., Kotzlowski H., 2000, Proc. SPIE, 4008, 534
Dinshaw N., Impey C. D., Foltz C. B., Weymann R. J., Chaffee F. H., 1994, ApJ, 437, L87
Dinshaw N., Foltz C. B., Impey C. D., Weymann R. J., Morris S. L., 1995, Nat, 373, 223
Dinshaw N., Weymann R. J., Impey C. D., Foltz C. B., Morris S. L., Ake T., 1997, ApJ, 491, 45
D'Odorico V., Cristiani S., D'Odorico S., Fontana A., Giallongo E., Shaver P., 1998, A&A, 339, 678
D'Odorico V., Petitjean P., Cristiani S., 2002, A&A, 390, 13
Fang Y., Duncan C., Crotts A. P. S., Bechtold J., 1996, ApJ, 462, 77
Haardt F., Madau P., 1996, ApJ, 461, 20
Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J., 1996, ApJ, 457, L51
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Hui L., Stebbins A., Burles S., 1999, ApJ, 511, L5
Impey C. D., Foltz C. B., Petry C. E., Browne I. W. A., Patnaik A. R., 1996, ApJ, 462, L53
Jakobsen P., Jansen R. A., Wagner S., Reimers D., 2003, A&A, 397, 891
Kim T.-S., Cristiani S., D'Odorico S., 2001, A&A, 373, 757
Kim T.-S., Cristiani S., D'Odorico S., 2002, A&A, 383, 747
Kim T.-S., Viel M., Haehnelt M., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Lopez S., Hagen H.-J., Reimers D., 2000, A&A, 357, 37
Lu L., Sargent W. L. W., Womble D. S., Takada-Hidai M., 1996, ApJ, 472, 509
McDonald P., 2003, ApJ, 585, 34
McDonald P., Miralda-Escudé J., 1999, ApJ, 518, 24
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471

Machacek M. E., Bryan G. L., Meiksin A., Anninos P., Thayer D., Norman M., Zhang Y., 2000, ApJ, 532, 118
Maza J., Ruiz M. T., Gonzalez L. E., Wischnjewsky M., Antezana R., 1993, Rev. Mex. Astron. Astrophys., 25, 51
Maza J., Ortiz P. F., Wischnjewsky M., Antezana R., Gonzalez L. E., 1995, Rev. Mex. Astron. Astrophys., 31, 159
Miralda-Escudé J., Rees M. J., 1993, MNRAS, 260, 617
Miralda-Escudé J., Cen R., Ostriker J. P., Rauch M., 1996, ApJ, 471, 582
Nusser A., Haehnelt M. G., 1999, MNRAS, 303, 179
Nusser A., Haehnelt M. G., 2000, MNRAS, 313, 364
Petitjean P., Surdej J., Smette A., Shaver P., Mücket J., Remy M., 1998, A&A, 334, L45
Rauch M., Sargent W. L. W., Barlow T. A., Carswell R. F., 2001, ApJ, 562, 76
Rauch M., Becker G. D., Viel M., Sargent W. L. W., Smette A., Simcoe R. A., Barlow T. A., Haehnelt M. G., 2005, ApJ, 632, 58
Ricotti M., Gnedin N. Y., Shull J. M., 2000, ApJ, 534, 41
Rollinde E., Petitjean P., Pichon C., Colombi S., Aracil B., D'Odorico V., Haehnelt M. G., 2003, MNRAS, 341, 1279
Sargent W. L. W., Young P., Boksenberg A., Tytler D., 1980, ApJS, 42, 41
Scannapieco E., Pichon C., Aracil B., Petitjean P., Thacker R. J., Pogosyan D., Bergeron J., Couchman H. M. P., 2006, MNRAS, 365, 615
Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, MNRAS, 318, 817
Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768
Seljak U. et al., 2005, Phys. Rev. D, 71, 3511
Smette A., Surdej J., Shaver P. A., Foltz C. B., Chaffee F. H., Weymann R. J., Williams R. E., Magain P., 1992, ApJ, 389, 39
Smette A., Robertson J. G., Shaver P. A., Reimers D., Wisotzki L., Koehler T., 1995, A&AS, 113, 199
Songaila A., 1998, AJ, 115, 2184
Spergel D. N. et al., 2003, ApJS, 148, 175
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2002, MNRAS, 333, 649
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Theuns T., Leonard A., Efstathiou G., 1998a, MNRAS, 297, L49
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998b, MNRAS, 301, 478
Theuns T., Viel M., Kay S., Schaye J., Carswell R. F., Tzanavaris P., 2002, ApJ, 578, L5
Viel M., Matarrese S., Mo H. J., Haehnelt M., Theuns T., 2002, MNRAS, 329, 848
Viel M., Haehnelt M. G., Carswell R. F., Kim T.-S., 2004a, MNRAS, 349, L33
Viel M., Haehnelt M. G., Springel V., 2004b, MNRAS, 354, 684
Vladilo G., Centurión M., D'Odorico V., Pèroux C., 2003, A&A, 402, 487
Weinberg D. et al., 1999, in Banday A. J., Sheth R. K., da Costa L. N., eds, Evolution of Large Scale Structure: From Recombination to Garching. Print Partners Ipskamp, Enschede, p. 346
Worseck G., Wisotzki L., 2006, A&A, 450, 495
Young P. A., Impey C. D., Foltz C. B., 2001, ApJ, 549, 76
Zhang Y., Anninos P., Norman M. L., 1995, ApJ, 453, L57
Zhang Y., Anninos P., Norman M. L., Meiksin A., 1997, ApJ, 485, 496

This paper has been typeset from a TEX/LATEX file prepared by the author.