High-resolution spectroscopy of the 3D cosmic web with close QSO groups[*]

M. Cappetta,[1,2][†] V. D'Odorico,[2] S. Cristiani,[2,3] F. Saitta[2] and M. Viel[2,3]

[1]*MPE – Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching bei München, Germany*
[2]*INAF – Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, I-34127 Trieste, Italy*
[3]*INFN/National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*

ABSTRACT

We study the 3D distribution of matter at $z \sim 2$ using high-resolution spectra of quasi-stellar object (QSO) pairs and simulated spectra drawn from cosmological hydrodynamical simulations. We present a sample of 15 QSOs, corresponding to 21 baselines of angular separations evenly distributed between $\sim$1 and 14 arcmin, observed with the Ultraviolet and Visual Echelle Spectrograph (UVES) at the European Southern Observatory Very Large Telescope (ESO–VLT). The observed correlation functions of the transmitted flux in the H I Lyman α forest transverse to and along the line of sight are in agreement, implying that the distortions in redshift space due to peculiar velocities are relatively small and – within the relatively large error bars – not significant. The clustering signal is significant up to velocity separations of $\sim$300 km s^{-1}, corresponding to about 5 h^{-1} comoving Mpc. Compatibility at the 2σ level has been found both for the autocorrelation and cross-correlation functions and for the set of the cross-correlation coefficients. The analysis focuses in particular on two QSO groups of the sample, the Sextet and the Triplet. Searching for alignments in the redshift space between Lyman α absorption lines belonging to different lines of sight, it has been possible to discover the presence of a wide H I structure extending over about 10 Mpc in comoving space, and give constraints on the sizes of two cosmic underdense regions in the intergalactic medium, which have been detected with a 91 and 86 per cent significance level, respectively, in the Sextet and the Triplet.

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

1 INTRODUCTION

The understanding of the Lyman α forest has dramatically improved in the recent decade, both on the theoretical and the observational side. Semi-analytical and hydrodynamical simulations have outlined a new picture where the Lyman α forest is due to the fluctuations of the intermediate and low-density intergalactic medium (IGM), arising naturally in the hierarchical process of structure formation. Relatively simple physical processes impact on the thermal state of the gas, which, on scales larger than the Jeans length, effectively traces the underlying distribution of dark matter. Support for this scenario is given by the satisfactory reproduction by semi-analytical and hydro-simulations of many properties of the Lyman α forest (from the column density and the Doppler parameter distribution to the number density and effective opacity evolution) derived from the analysis of high-resolution, high signal-to-noise ratio (S/N) quasi-stellar object (QSO) spectra obtained at 8–10 m class telescopes (e.g. Davè et al. 1999; Brian & Machacek 2000; Kim, Cristiani & D'Odorico 2001; Bianchi et al. 2003; Janknecht et al. 2006). A step forward in the study of the IGM with QSO absorption spectra is represented by the use of multiple lines of sight at close angular separations, which allows information about the transverse direction to be obtained. Common absorption features observed in the spectra of multiply lensed quasars (e.g. Smette et al. 1992) and of close quasar pairs (e.g. D'Odorico et al. 1998; Aracil et al. 2002) have provided evidence that the Lyman α absorbers have dimensions of a few hundred kpc, in agreement with the predictions of simulations. Recently, lensed and more widely separated QSO pairs have been used to recover the kinematics of the gaseous cosmic web (Rauch et al. 2005), confirming that the Hubble expansion and gravitational instability are the main processes influencing the Lyman α forest gas.

---

[*]Based on observations collected at the European Southern Observatory Very Large Telescope, Cerro Paranal, Chile – Programmes 68.A-0216(A), 69.A-0204(A), 69.A-0586(A), 70.A-0031(A) and 077.A-0714(A).
[†]E-mail: cappetta@mpe.mpg.de

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	3–15
	T2	2.05	18.30	1.592–1.999	4–15
	T3	2.053	18.10	1.665–2.002	2.5–7
Sextet	S1	1.907	19.66	1.665–1.859	2–7
	S2	2.387	19.53	1.858–2.331	3–8
	S3	2.102	19.31	1.633–2.051	4–12
	S4	1.849	19.59	1.575–1.802	3–9
	S5	2.121	18.85	1.633–2.069	4–12
	S6	2.068	20.19	1.592–2.017	3–10
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

A critical test of the nature of the Lyman α absorbers, as proposed by simulations, comes from the determination of their spatial distribution properties. This has been done by analysing a great number of uncorrelated QSO lines of sight and computing the flux correlation function (CF) by averaging over many spectra (e.g. Tripp, Limin & Savage 1998; Savaglio et al. 1999; Croft et al. 2002). In this observational approach, however, the 3D information is convolved with distortions in redshift space, due to peculiar motions and thermal broadening. Multiple lines of sight at small angular separations offer an invaluable alternative to address the spatial distribution of the absorbers, enabling a more direct interpretation of the observed correlations.

The final goal of this work is to investigate the distribution properties of matter in the IGM applying the modern interpretation of the Lyman α forest to a sample of close QSO groups. The computation and comparison of the flux CF along and across the lines of sight was carried out in a previous paper (D'Odorico et al. 2006, hereafter

Paper I). In this paper, we update the previous results using higher S/N spectra and, in particular, we investigate the coincidences of absorbers among three or more close lines of sight in order to detect cosmological structures extending to large scales. Finally, we look for extended underdense regions in the IGM using multiple lines of sight.

Note that tomographic studies based on Lyman α lines similar to those presented here will be of great importance in the near future due to the large numbers of quasar spectra that will be collected using low- and medium-resolution spectrographs (Baryon Oscillation Spectroscopic Survey and the X-Shooter instrument for example) that aim at constraining spatial matter correlations in the transverse and longitudinal directions to ultimately measure baryonic acoustic oscillations and perform the Alcock–Paczyinski test at high redshift.

The paper structure is the following. In Section 2 we describe the observed data sample, the reduction procedure and the simulated spectra. Section 3 is devoted to the computation of the auto-CF and cross-CF and the cross-correlation coefficients (CCCs) of both the observed and simulated spectra and to their comparison. In Section 4, we analyse the coincidences in redshift space of the H I and C IV absorbers in order to detect structures extending over several comoving Mpc, while the underdense regions common to multiple lines of sight are studied in Section 5. The conclusions are then drawn in Section 6. Throughout this paper we adopt $\Omega_{0m} = 0.26$, $\Omega_{0\Lambda} = 0.74$ and $h = H_0/(72 \text{ km s}^{-1} \text{ Mpc}^{-1})$.

2 DATA SAMPLE

The QSO pairs and groups forming our sample are the same as Paper I and are described in Table 1. The patterns on the sky of the two QSO groups called the Sextet and the Triplet are reported in Fig. 1.

We were allocated 17 h with Ultraviolet and Visual Echelle Spectrograph (UVES) at the Very Large Telescope (VLT) in order to increase the S/N of the spectra of objects T1 and T2 of the Triplet and S1, S2, S5 and S6 of the Sextet. The new observations were reduced with the UVES pipeline following a standard procedure (Ballester et al. 2000). A pre-filtering of the cosmic rays of the blue
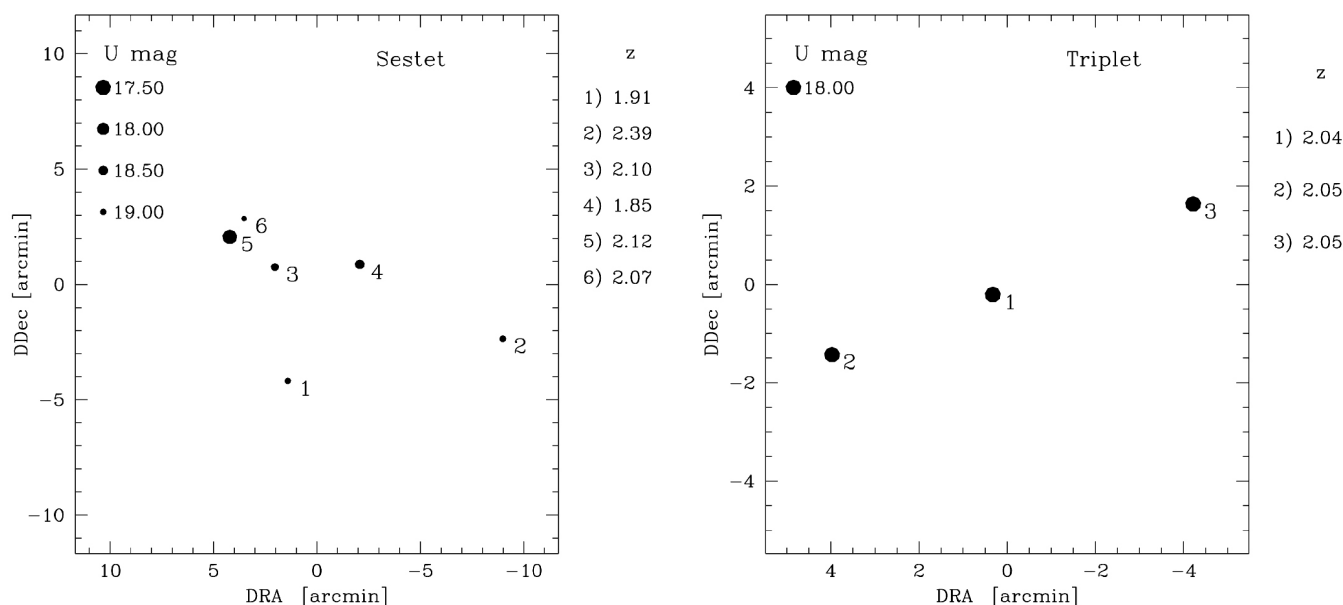
Figure 1. Relative positions, *U* magnitudes and redshifts of the QSOs composing the Sextet (left-hand plot) and the Triplet (right-hand plot) in our sample.

band frames was necessary for the faintest spectra to carry out the optimal extraction properly.

Then, the heliocentric and vacuum wavelength corrections were performed and the new spectra were combined with the old ones. The final spectra have resolution $R \sim 45\,000$, while the S/N per pixel varies on average between 3 and 12 in the Lyman α forest and between 6 and 20 in the C IV forest (see Table 1 for details).

The estimate of the continuum level, in particular in the Lyman α forest region, is a very delicate step in the process of spectra reduction. Procedures realized up to now in order to determine the continuum position through automatic algorithms do not give satisfactory results. Following the same procedure adopted in Paper I, we fitted the regions free from clear absorption with a spline polynomial of third order. The limitations introduced by the uncertainty in the true continuum level should play a minor role in the computation of the cross-CF with respect to the case of single line of sight analysis. This is because the power on scales of the continuum fluctuations is uncorrelated between the different lines of sight. Thus, probing the transverse direction could potentially allow us to measure the matter distribution at scales larger than those probed along the line of sight (see Viel et al. 2002, for a discussion in the case of power spectra).

All the lines falling in the Lyman α forest have been fitted with a Voigt profile via χ^2 minimization. The lines with an equivalent width (EW) lower than three times the related EW uncertainty have been removed from the list of fitted lines. The metal lines have been identified first looking for the most common doublets (e.g. C IV, Si IV, O VI and Mg II). Then, we have searched for other common transitions (e.g. Si III, Si II, C II and Fe II) at the redshift of the previously determined systems.

Our sample provides 21 QSO pairs with angular separations uniformly distributed between ~ 1 and 14 arcmin, corresponding to comoving spatial separations between ~ 1.4 and $21.6\,h^{-1}$ Mpc. The median redshift 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.

2.1 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.

The details of the adopted simulations can be found in Paper I; here we provide only the basic information. 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 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 of 3.3 in order to fit observational constraints on the temperature evolution of the IGM. More details can be found in Viel, Haehnelt & Springel (2004).

The cosmological model corresponds to a 'fiducial' Lambda cold dark matter (ΛCDM) universe with parameters $\Omega_{0m} = 0.26$,

$\Omega_{0\Lambda} = 0.74$, $\Omega_{0b} = 0.0463$ and $H_0 = 72$ km s^{-1} Mpc^{-1} (the B2 series of Viel et al. 2004). 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 and other results obtained by the Lyman α forest community (e.g. Spergel et al. 2003; Viel et al. 2004; Seljak et al. 2005).

The $z = 1.8$ output of the simulated box was pierced by three lines of sight in order to obtain 50 triplets of spectra carefully reproducing the observed Triplet mutual separations and spectral properties. The same was done for 50 sextets of lines of sight reproducing the observed Sextet and 50 pairs of lines of sight at the same angular separation as Pair U. 50 different realizations of Pair A spectra and of Pair Q were obtained from the output box at redshift $z = 2.4$.

Finally, we added to the simulated spectra both instrumental broadening and a Gaussian noise in order to reproduce the observed average S/N (per pixel): S/N = 5 for the Triplet, the Sextet and Pair A; S/N = 9 for Pair U and S/N = 15 for Pair Q.

3 CORRELATION FUNCTIONS OF THE IGM

In this section, we discuss the flux correlations in the absorption spectra of our sample of QSO pairs. The statistical quantities are the same as those already computed in Paper I; here we verify the effect of having an increased S/N.

On the basis of the interpretation of the Lyman α forest as due to a continuous density field with a one-to-one 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 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 contributions 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.

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 et al. 2002; Schaye et al. 2003; Viel et al. 2004; Kirkman et al. 2005; Faucher-Giguère et al. 2008), because we verified that its effect on the CF is negligible. By means of this new field, δ_f, we could then define and compute three useful tools for the investigation of the Lyman α correlation properties: the auto-CF and cross-CF and the set of the CCCs.

3.1 The autocorrelation function

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

$$\xi_\parallel^f(\Delta v_\parallel) = \langle \delta_f(v_\parallel)\delta_f(v_\parallel + \Delta v_\parallel)\rangle, \tag{1}$$
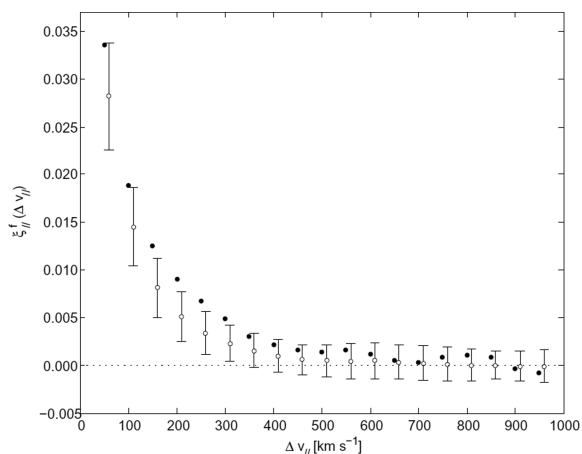
Figure 2. Comparison of the observed and simulated auto-CFs: solid dots refer to the observations and empty dots to the simulations. Simulated data are slightly shifted in velocity for clarity.
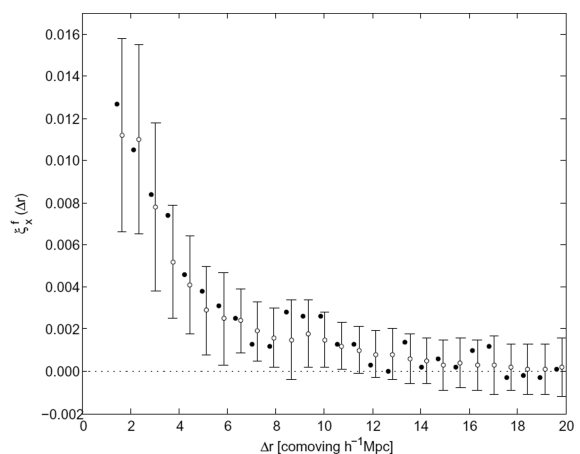

Figure 3. Comparison of cross-CFs: solid dots are related to the observations and empty points to the simulations. Simulation data are slightly shifted in comoving distance for clarity.

following previous studies on the same subject (e.g. Rollinde et al. 2003; Becker, Sargent & Rauch 2004; McDonald et al. 2005). The auto-CF for our sample of QSO spectra was obtained by averaging over all the pixels of all the QSOs. The results were binned in 50 $km\,s^{-1}$ velocity bins. The auto-CF for the simulated spectra was computed as the arithmetic mean of the CFs 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 for the simulated $\xi_{\parallel}^{f}$ are strongly correlated. This is due to the fact that every pixel contributes to the CF in several velocity bins.

In Fig. 2 we show the comparison of the observed and simulated auto-CF in velocity space. With the improved S/N with respect to Paper I, the agreement between the two CFs has increased, weakening any evidence for extra clustering in the regions occupied by QSO pairs.

3.2 The cross-correlation function

In this section we exploit the capabilities 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 CF 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. As a first approach, we computed the cross-CF extending in a natural way the procedure adopted for the auto-CF.

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:

$$r_{\parallel}(z) = \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')},\qquad(2)$$

where $E(z)$ describes the evolution of the Hubble parameter as a function of the redshift.[1]

In the definition of the comoving distance there is the implicit hypothesis that peculiar velocities give a negligible contribution to

the measured redshift in the Lyman α forest (Rauch et al. 2005; Paper I). The cross-CF of the transmitted flux between two lines of sight at angular separation $\Delta\theta$ is defined as

$$\xi_{\times}^{f}(\Delta r) = \langle \delta_f(\theta, r_{\parallel,1}) \delta_f(\theta + \Delta\theta, r_{\parallel,2}) \rangle,\qquad(3)$$

where $\Delta r = \sqrt{(r_{\parallel,1}^2 + r_{\parallel,2}^2 - 2\,r_{\parallel,1}\,r_{\parallel,2}\,\cos\Delta\theta)}$ 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.

We also computed the cross-CF for the sample of mock spectra. The simulated spectra are characterized by the redshift of the output box and a velocity extent. 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-CF and its 1σ standard deviation. The result of our computation is shown in Fig. 3 compared with the observed cross-CF of the pairs. The two functions show very good agreement (at the 1σ level) and a significant clustering signal up to comoving separations of $\sim 4\,h^{-1}$ comoving Mpc.

We also compared the cross-CF with the previously computed auto-CF. As can be seen in Fig. 4 the two data series show a consistency at the 1σ level, supporting the hypothesis that the peculiar velocity along the lines of sight is negligible.

3.3 The cross-correlation coefficient

A measure of the transverse clustering properties of the IGM which is less affected by peculiar velocities is the set of the flux CCCs,

$$\chi_{\times}^{f}(\Delta\theta) = \langle \delta_f(\theta, v_{\parallel}) \delta_f(\theta + \Delta\theta, v_{\parallel}) \rangle,\qquad(4)$$

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 $\chi_{\times}^{f}(\Delta\theta)$, and a sample with several pairs at different separations, as in our sample, gives an estimate of the CF. At a given redshift, the angular separation $\Delta\theta$ corresponds to a velocity separation $\Delta v_{\perp} = c\,F(z)\,\Delta\theta$, where c denotes the speed of light and $F(z)$ is a dimensionless function of redshift that includes all the

---

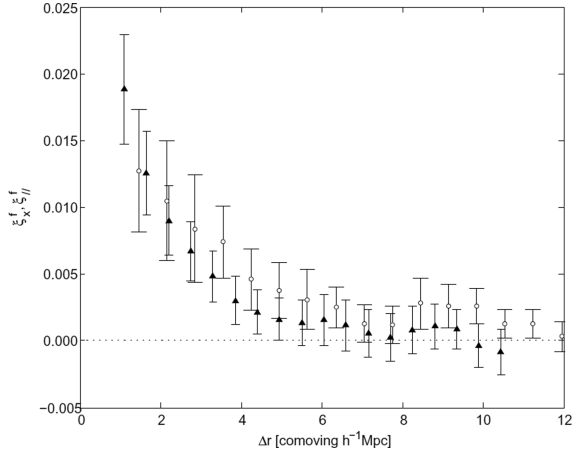[1] $E(z) = \sqrt{\Omega_{0m}\,(1+z)^3 + \Omega_{0\Lambda}}.$

Figure 4. Comparison between auto-CF and cross-CF, shown with triangles and circles, respectively. Error bars show the uncertainties related to the simulated samples.

dependence on the global cosmological metric. In the cosmological model that we have adopted

$$F(z) = \frac{E(z) \int_0^z [\mathrm{d}z'/E(z')]}{1 + z}, \qquad (5)$$

where z is the mean redshift of the overlapping Lyman α region from a pair. We computed the range of velocity separations covered by each of our pairs of spectra; then we grouped the pairs in velocity bins of variable width and computed the average CCC for every group. Given the small number of pairs in each group (a maximum of three QSO pairs) the uncertainties associated with these determinations can be only computed by applying a simple error propagation. However, these uncertainties would account only for the pixel statistics and the noise associated with the transmitted flux, but would not be 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 CCC with its error, which in this case is the standard deviation of the distribution of values.

The results are shown in Fig. 5. The CCC computed from the observations are in excellent agreement, at the 1σ confidence level (CL), with the simulations.

A comparison can be done with the same results obtained in Paper I: thanks to the increased S/N for some of the spectra in the present sample, in particular the S1–S3 pair, the enhanced clustering signal measured in Paper I with the CCC at a transverse velocity separation $\Delta v_\perp \sim 500 \ \mathrm{km \, s^{-1}}$ is no longer significant.

We conclude this section by comparing the CCC with the cross-CF. Both estimators measure the clustering properties of the IGM in the transverse direction with respect to the line of sight. However, the cross-CF could be affected by the peculiar velocities in the z direction while their effect should be negligible for the CCC.

Concerning the CCC, the angular separation $\Delta\theta$ between two QSO lines of sight was transformed into a comoving spatial separation, Δr, with the formula

$$\Delta r = \frac{c \Delta\theta}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')}, \qquad (6)$$

where $E(z)$ was defined above and here z is the mean redshift of the Lyman α interval considered for each pair. Assigning, at both the CCC and the cross-CF observed data points, the error bars computed from the simulations, there is an agreement within 1σ between the
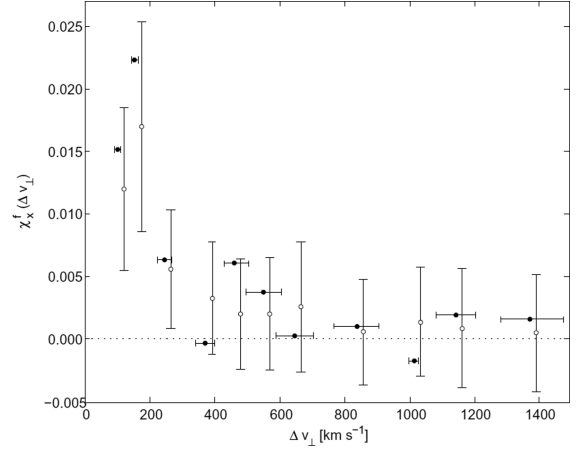

Figure 5. Comparison of the CCCs for our sample of observed spectra (solid dots) and of simulated ones (empty circles) as a function of the velocity separation, $\Delta v_\perp = c \, F(z) \, \Delta\theta$, corresponding to the angular separation, $\Delta\theta$, of the QSO pairs. Simulated data are slightly shifted in velocity for clarity. Error bars on the observed values along the x-axis represent the velocity range covered by the considered pairs.
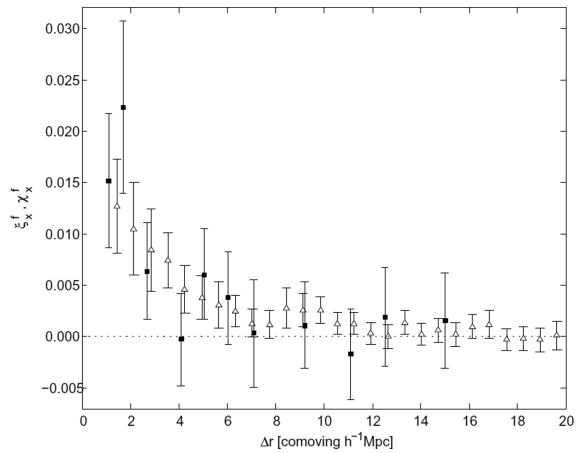

Figure 6. Cross-CF of the transmitted flux, $\xi_\times^f$, for our sample of pairs and groups of QSOs (empty triangle, slightly shifted in Δr) compared with the observed CCCs, $\chi_\times^f$ (squares), as a function of spatial separation (see text). Error bars both on $\chi_\times^f$ and $\xi_\times^f$ have been determined from simulations.

two series of points (see Fig. 6), confirming the hypothesis that the peculiar velocity along the lines of sight are negligible. The large variations from one data point to the other in the CCC 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 $\xi_\times^f$ is artificially increased by the fact that the values in the different bins are not independent.

All the results reported in this section are in agreement with the same quantities derived from other larger sample of QSO pairs (see e.g. Rollinde et al. 2003; Coppolani et al. 2006).

4 COINCIDENCES OF THE ABSORPTION LINES

4.1 The Lyman α line coincidences

The cross-CF and the CCC are two powerful tools which allow us to measure the statistical properties of the IGM distribution in the real

© 2010 The Authors. Journal compilation © 2010 RAS, MNRAS **407**, 1290–1300

3D comoving space. By definition, they account for the correlation between pairs of Lyman α forests in close QSO spectra. What we want to study now is the simultaneous correlation among three or more Lyman α forests. A natural way to extend the cross-CF of the transmitted flux to more than two lines of sight is the addition of a third parameter, related to the third QSO at angular separation $\Delta\theta$ with respect to the first. This operation leads to a function of three variables, the three reciprocal comoving distances between the three considered pixels, which is computation time demanding and difficult to interpret.

In the previous section, we have taken advantage of the interpretation of the Lyman α forest which allows us to map directly the transmitted flux along the line of sight into the IGM density field. A boost in the signal of the cross-CF between two lines of sight is due to the presence in redshift space of two aligned, or very close, Lyman α lines belonging to the two considered spectra. On this basis, we can provide a measure of the cross-correlation between three Lyman α forests by searching for triplets of Lyman α lines, belonging to three different spectra, aligned in redshift space within a given velocity window.

This kind of analysis has been applied to the Triplet and to all the combinations of three QSOs that could be formed with the Sextet. The adopted procedure has been the following.

(1) The lists of H I Lyman α lines compiled for the QSOs in our sample were considered in the redshift range between the Lyman β emission (or the shortest observed wavelength, when the Lyman β was not included in the spectrum) and 5000 km s^{-1} from the Lyman α emission (to avoid proximity effect due to the QSO).

(2) Each pair of lines with a velocity separation $\Delta v \leq 100$ km s^{-1} has been replaced by a single line with central wavelength equal to the average value of the parent lines weighted on the EW. This velocity threshold has been chosen on the basis of the characteristic width of Lyman α lines, $\sim$25–30 km s^{-1} (see e.g. Kim et al. 2002). Furthermore, this is also the velocity scale corresponding to the Jeans length, which sets the characteristic dimension of H I absorbers.

(3) Triplets of lines, each one belonging to a different line of sight, have been considered and the velocity difference between the largest and smallest redshift has been computed. This operation has been done for the Lyman α lines in the three lines of sight of the Triplet and in all the triplets of lines of sight (20 possible combinations) provided by the Sextet. Then, all the measures of velocity difference lower than 1000 km s^{-1} have been divided into velocity bins of 100 km s^{-1} and the related histogram with the number of occurrences for each bin has been computed.

(4) Next, the previous three steps have been repeated for a sample of 10^3 mock lists of lines built in the following way. In order to take into account the varying number density of detectable lines along the Lyman α forests, due to the varying S/N, each forest has been simulated in chunks of about 200 Å. In each mock chunk, the number of simulated lines has been determined from a Poissonian distribution centred on the number of observed lines in that chunk, while the positions of the mock lines have been randomly generated following a uniform distribution within the related wavelength range of each chunk. The redshift intervals masked in the observed spectra were masked also in the simulated ones. The EWs of the mock lines have been randomly chosen among all the EWs measured by the fit of the lines in the observed spectra. In this way it has been possible, for each velocity bin, to compute the mean and the standard deviation of the number of occurrences for synthetic lists of lines.
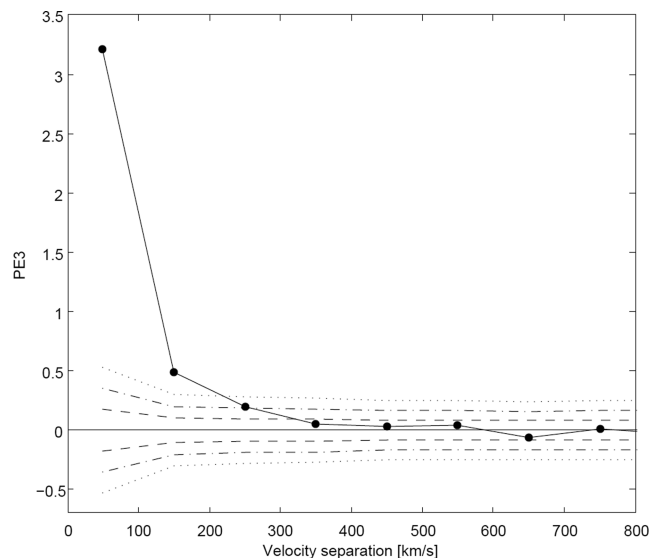

Figure 7. Three-point probability excess (PE3) as a function of the velocity difference of the three Lyman α lines (dots). Values referring to a significance level of 1, 2 and 3σ are reported too (dashed, dot–dashed and dotted lines, respectively).

(5) Finally, we have defined the three-point probability excess (PE3) as a function of the velocity difference, Δv, according to the following formula:

$$\text{PE}(v) = N_{\text{obs}}(\Delta v)/N_{\text{sim}}(\Delta v) - 1. \qquad (7)$$

The resulting PE3 is reported in Fig. 7, together with the 1, 2 and 3σ CLs. The PE3 is non-zero at a 2σ level up to a velocity difference of $\sim$250 km s^{-1}. Most of the signal of the PE3 is due to the large number (26) of coincidences produced by the S3–S5–S6 QSOs triplet which is also the closest triplet (mean angular separation of 2.02 arcmin corresponding to $\sim$2 h^{-1} comoving Mpc).

Fig. 8 shows the probability excess considering quadruplets of Lyman α lines. A significant signal at more than 3σ level is
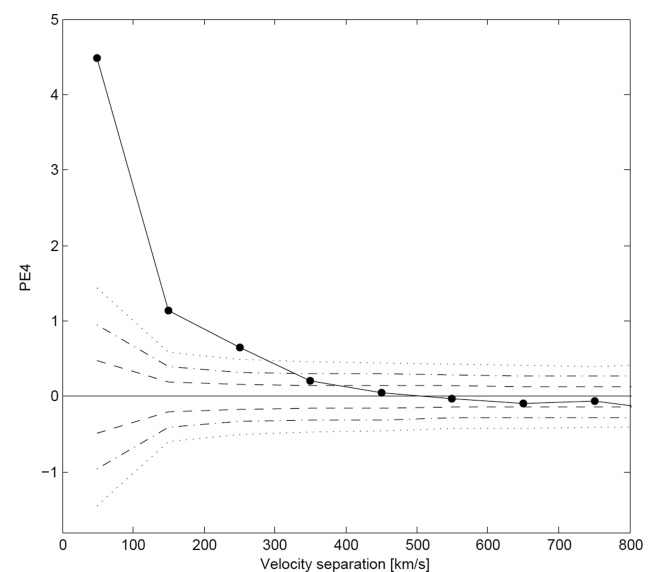

Figure 8. Four-point probability excess (PE4) as a function of the velocity difference of the three Lyman α lines (dots). Values referring to a significance level of 1, 2 and 3σ are reported too (dashed, dot–dashed and dotted lines, respectively).
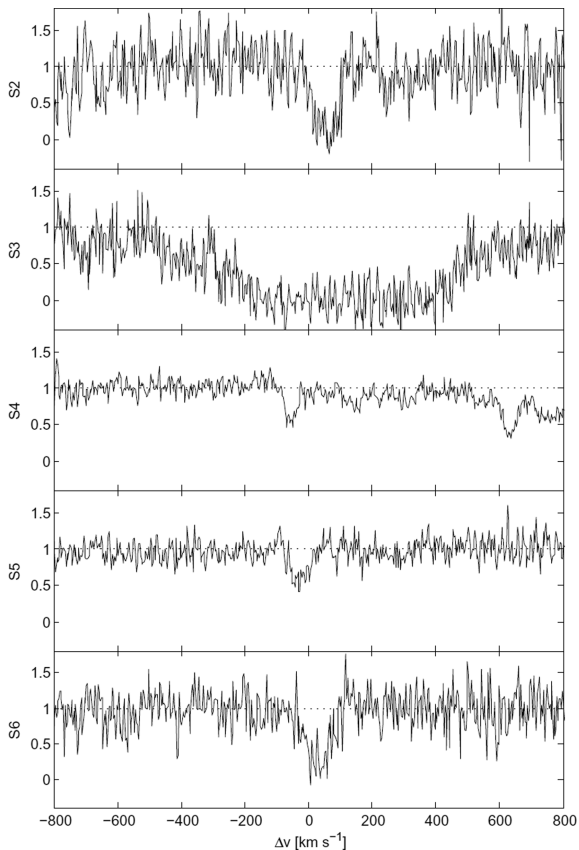
Figure 9. Portions of the spectra of the QSOs S2, S3, S4, S5 and S6 centred at 3435 Å showing the aligned Lyman α lines.

measured up to a velocity difference of $\sim$250 km s^{-1}. Besides, one group of five coincident lines within 100 km s^{-1} in the S2–S3–S4–S5–S6 QSOs is observed at a mean redshift of 1.825, an occurrence that has a probability $P = 0.013$ to arise from a random distribution of lines. The portion of spectra where these five lines fall are reported in Fig. 9.

In particular, it is possible to observe in the spectrum of the S3 QSO the presence of a damped Lyman α system (DLA): the fitted Voigt profile gives a column density value of $\log N = 20.6$. This DLA is associated with several metallic ion absorption lines found in the redder part of the spectrum. Indeed at the same redshift we have found evidence of C IV, Fe II, Si IV, Si III, Si II, Al III and Al II. This correlated H I absorption across lines of sight separated by $\sim$14 h^{-1} comoving Mpc could be interpreted as due to a gas filament, but other explanations are possible such as detected H I haloes of clustered galaxies.

4.2 The C IV line coincidences

We have also looked for correlation across two lines of sight for C IV absorbers. Given the characteristics of our sample, which allow us to investigate comoving scales larger than $\sim$1 h^{-1} comoving Mpc, we do not expect C IV coincidences due to the same C IV cloud, whose typical size should be of the order of tens of kpc (Rauch et al. 2001). Our lines of sight could pierce C IV gas connected with different galaxies in groups or clusters (Francis et al. 2001).

The procedure to identify C IV coincidences is similar but simpler than that previously employed for the Lyman α lines. The first two steps are the same, with the difference that here lines closer

than 500 km s^{-1} have been merged together. Then, we searched for coincidences of C IV merged absorbers lying along the redshift space in windows smaller than 500 km s^{-1}. According to Scannapieco et al. (2006) this velocity scale corresponds to the characteristic size of coherent C IV structures.

In order to establish the significance of our observation, we have applied our algorithm for coincidences to 10^4 groups of synthetic list of lines. The method for creating the synthetic list of lines is the same as adopted before in the case of the mock list of Lyman α lines. With the total set of simulated results, we could find the number $N(n \geq n_0)$ of cases in which, in a random distribution of C IV absorbers along the lines of sight, the number of chance coincidences n is equal to or greater than the number of coincidences observed in our sample, n_0. Then the related probability P was simply computed using the ratio between $N(n \geq n_0)$ and the total number of simulations, that is 10^4. Finally, we computed the CL of our observation simply with the formula $CL = 1 - P(n \geq n_0)$.

We have found only two coincidences, in the T1–T2 and S3–S6 pairs, at a redshift of 2.062 and 1.486, respectively, and separated by a comoving distance of 4.4 and 2.8 h^{-1} comoving Mpc. The statistical significance of these detections are 90 and 85 per cent, respectively. Coincidences have also been found between the T2 QSO redshift and a T3 absorber (with a separation of 9.6 h^{-1} comoving Mpc) and between the S6 QSO redshift and a S3 absorber (with a separation of 2.8 h^{-1} comoving Mpc). These coincidences are in agreement with the fact that QSOs reside in overdense regions of the Universe (e.g. D'Odorico et al. 2008). No coincidences among absorbers have been found considering three or more lines of sight. The portions of spectra showing the two coincidences are reported in Fig. 10.

5 COMMON UNDERDENSE REGIONS

The two groups of QSOs in our sample are well suited to search for large underdense regions in the IGM. Previous attempts to detect underdense regions in the Lyman α forest have searched for regions without absorption lines along a single line of sight. They looked for portions of spectra where the observed number of absorption lines was significantly smaller than the expected number from synthetic spectra (Carswell & Rees 1987; Crotts 1987; Ostriker, Bajtlik & Duncan 1988; Cristiani et al. 1995; Kim, Cristiani & D'Odorico 2001).

A strong improvement of the statistical significance of each detection can be provided by the simultaneous presence of an underdense region along several close lines of sight. This would allow the detection of smaller underdense regions maintaining a high CL. As done in the previous section, we focused our attention on searching for common underdense regions among the lines of sight of the Triplet and among all the possible groups of lines of sight provided by the Sextet.

Following Rollinde et al. (2003), we can define an underdense region along a single line of sight as a portion of the spectrum between λ_i and λ_e, where the normalized flux f is higher than the specific threshold ($\bar{f} - \sigma_f$), where $\bar{f}$ is the average flux in the Lyman α forest and σ_f is the value of the flux uncertainty. Note also that at $z \sim 2$ regions above the mean density correspond on average to regions below the mean flux level of high-resolution spectra (Viel, Colberg & Kim 2008). Since the S/N of the spectra analysed by Rollinde et al. (2003) is much larger than that of our spectra (15–70 versus 4–15, respectively), this simple procedure for the detection of underdense regions cannot be adopted in the present case. Indeed, the fluctuations due to the flux noise cause the splitting of single
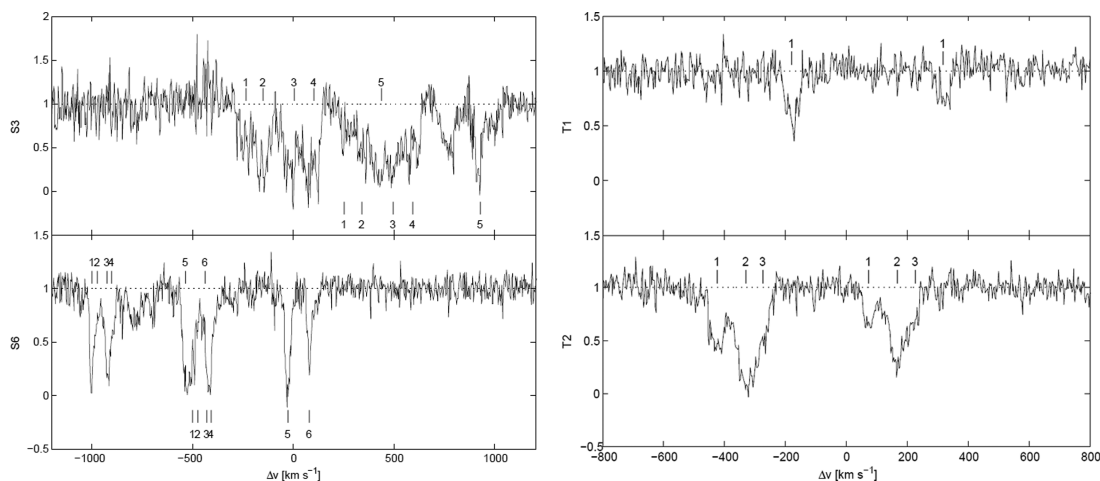
Figure 10. Portions of spectra representing the C IV systems found to be in coincidence between the S3 and S6 QSOs (left-hand plots) and the T1 and T2 QSOs (right-hand plots). The single C IV components are marked with increasing number. The same number is related to both the C IV lines.

large underdense regions into several smaller regions, preventing their detection.

In order to reveal the presence of underdense regions in our spectra, we have developed the following method. The spectra are smoothed with a median filter characterized by a very short kernel length, only 3 pixels. After the smoothing, common underdense regions are looked for with the method described above (always adopting the same average flux and flux uncertainty). Then, the spectra are smoothed again and again searched for common voids. This process is iterated several times. The results are shown in Fig. 11 where we have reported the size of the bigger common underdense region found in the Triplet and in the S3–S5–S6 group of Sextet QSOs as a function of the number of times the spectra have been smoothed. We report only the result concerning the S3–S5–S6 group because it is the only one in which a significant underdense region has been detected.

We can see how, after an initial linear growth of the size of the common underdense region, a step behaviour occurs in both the series of data. Steps occur whenever two separate underdense regions are merged together, that is when the absorption line (both real and spurious) lying between them is smoothed out by the nth action
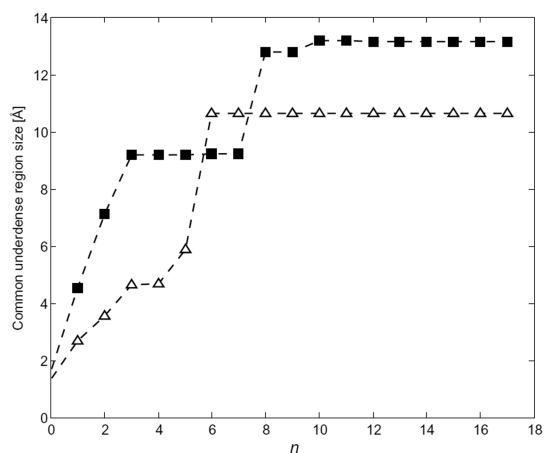
of the median filter. We can note also how, after $\sim$10 smoothings, the size of the underdense regions in both the Triplet and the S3–S5–S6 group reaches an 'asymptotic' value. Each asymptotic rate corresponds to the size of a common underdense region enclosed between two strong lines (not necessarily belonging to the same spectrum) requiring a large number of smoothings to be cancelled.

A series of simulations has been carried out in order to evaluate the minimum value of the column density $N(H\,I)$ of the Lyman α lines surviving after six and eight smoothing processes in the Triplet and the S3–S5–S6 group, respectively. First, a mock spectrum has been created placing several lines characterized by an increasing $\log N(H\,I)$, from 12 to 15 by steps of 0.05, and central wavelengths separated by 10 Å. Line profiles have been modelled with a Gaussian since we were interested only in the centre of the lines and not in reproducing correctly their wings. The Doppler parameter b was set to a value of 20 km s^{-1} (Kim et al. 2001); the dependence of the simulation results from this parameter is negligible.

White noise has been added to the spectrum in order to obtain S/N = 7, roughly consistent with the observed S/N values of the Triplet, and of the S3, S5 and S6 spectra in the segments where the common underdense regions are situated. Then, the length of the void has been measured, namely the portion of mock spectrum where $f \geq (\bar{f} - \sigma_f)$, after each smoothing. Since the lines are ordered by increasing column density and separated by 10 Å, each increase in void length of $\sim$10 Å corresponds to the vanishing of a specific line with known value of $N(H\,I)$. Repeating this procedure 10^3 times, we have computed the values, and the related uncertainties, of the column density of the lines washed away after a given number of smoothings. The effect of having different S/N in the QSO spectra (Fig. 12) instead of a fixed value of S/N = 7 is smaller than the error bars reported in Fig. 13.

The same kind of simulations have been used to compare the action of the short-scale filter iterated n times with respect to that of a median filter of length $(2n+1)$ pixel employed once. Both the series of data are shown in Fig. 13, with the related error bar, as a function of n. The trend followed by the two series at large n is different: for the short-scale filter iterated n times the smoothing effect is relatively lower and the growth of the column density of the lines smoothed away is flatter. Furthermore, the related uncertainties are much smaller. This behaviour helps us to estimate more accurately the lower column density of the surviving lines because the range
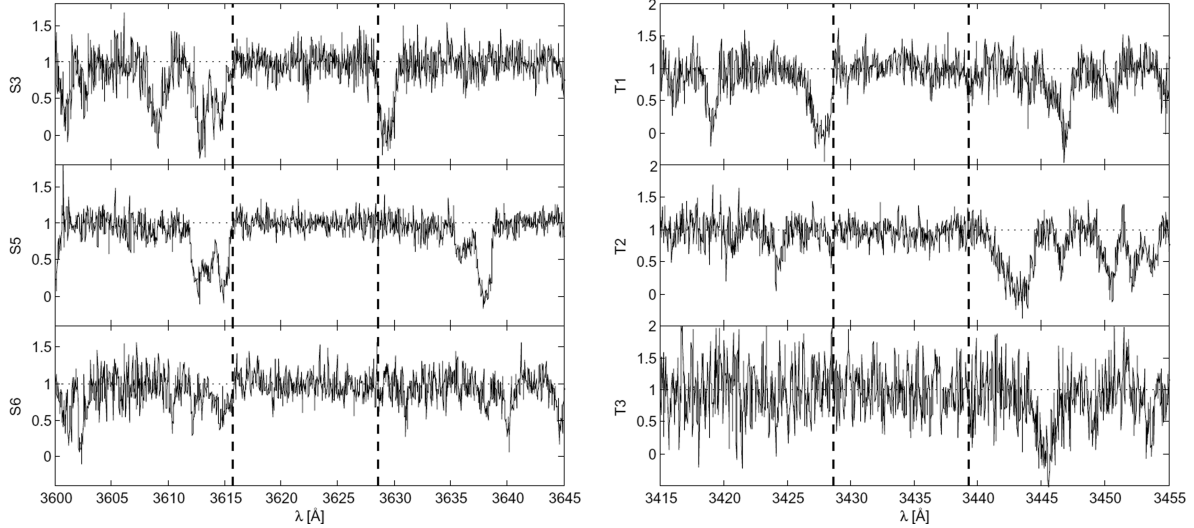

Figure 11. Size of the common underdense regions of Triplet group (empty triangle) and the S3–S5–S6 group (filled square) as a function of the number of times (n) the spectra have been smoothed.

Figure 12. Spectra of the two close groups of quasars (S3–S5–S6 on the left-hand plot and the Triplet on the right-hand plot) centred on the common underdense regions. The extension of the common underdense regions are shown with dashed vertical lines.
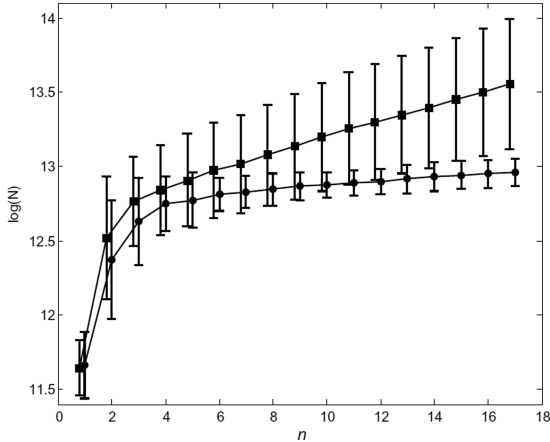

Figure 13. Minimum values of the H I column density of the Lyman α lines left after a given number n of smoothing with a 3 pixel long filter (circle) and after one smoothing with a $(2n+1)$ pixel wide filter (square, slightly shifted along the x-axis for clarity).

of $\log N(\text{H\,I})$ between 12.5 and 13 is sampled more accurately with respect to the upper curve.

Fig. 13 (filled circles) shows that the asymptotic size of the voids observed in the Triplet and in the S3–S5–S6 group are referred to spectra where the absorption lines characterized by $\log N(\text{H\,I}) \lesssim 12.8$ have been smoothed away.

The two largest underdense regions common to the S3, S5 and S6 lines of sight and to the Triplet lines of sight are 12.8 Å (corresponding to 10.7 h^{-1} comoving Mpc) and 10.6 Å (or 8.8 h^{-1} comoving Mpc) in size, respectively. They are shown in Fig. 12. The sizes of the related underdense regions along the single lines of sight are: 13.4, 19.8 and 15.2 Å for the S3, S5 and S6 lines of sight; 10.6, 12.4 and 14.2 Å for the lines of sight of the Triplet. As already said, no common underdense regions have been found considering any other combination of lines of sight in the Sextet.

Assuming that the underdense regions have a spherical geometry (again following Rollinde et al. 2003), a void is then defined as the largest sphere that may be included inside one connected region.

We searched for the largest spherical region included in the three individual underdense regions. This is equivalent to the requirement that the surface of the underdense region matches the six edges of the three segments. We allow an uncertainty of few Å in the position of the surface along each line of sight. This takes into account the fact that observations are in redshift space, so that the border of the underdense region may have a peculiar velocity that shifts its position in wavelength, and that a real underdense region probably does not have a simple boundary surface. This set of equations can constrain a sphere that has four degrees of freedom. Inside the sphere, the IGM will be underdense along each line of sight and on its surface it will be overdense at six positions identifying the edges of the individual underdense regions. When we applied this procedure to the S3–S5–S6 group, the underdense region can be parametrized with a sphere of 6.75 h^{-1} comoving Mpc radius. On the other hand, due to the positions of the members of Triplet in the sky, a spherical parametrization of the void is not possible in this case.

In order to assign a significance level to the detection of these two underdense regions, we performed 10^3 simulations of mock spectra reproducing the two triplets of QSO spectra. First, the lists of central wavelengths were created following the same procedure as defined in Section 4.1. Then, the line profiles were created with the observed column density and Doppler parameter distribution (e.g. Kim et al. 2002). Random white noise was added, according to the observed S/N of each spectrum in each chunk, and then the spectra were smoothed eight times by the median filter. We verified that the results of the simulations do not change significantly for a number of iterations larger than six. Finally, we computed the distribution function of the probability of having one common underdense region of given dimension along three lines of sight. Computing the random probability of having a common underdense region of size equal to or greater than ours, we obtained that the final significance levels of our detections are 91 and 86 per cent for the Triplet and for the S3–S5–S6 group, respectively. These two values are smaller than the significance levels of the common underdense region detected by Rollinde et al. (2003). We expect that this is due to the fact that we are dealing with only three lines of sight and that our spectra have a lower S/N.

6 CONCLUSIONS

In this paper, we exploited the capabilities of high-resolution UVES spectra of QSO pairs to study the 3D distribution properties of baryonic matter in the IGM as traced by the transmitted flux in the QSO 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 $z \simeq 1.8$. By calculating the CFs we compared the observed sample with a set of mock spectra drawn from a cosmological hydro-simulation 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 (see Section 2 for details). Furthermore, particular emphasis has been given to the search for alignments and other particular features of the Lyman α forests in the two QSO groups present in our sample.

In the following, we summarize our main results.

Two-point statistics

(i) The computed CFs are in substantial agreement with those obtained in our previous paper (Paper I). There is consistency between the clustering properties of matter in the IGM calculated in the direction parallel and transverse to the line of sight using the parameters of the concordance cosmology to map the angular distance into velocity separation. This is also due to the relatively large error bars of the computed quantities. As an implication, peculiar velocities in the absorbing gas are likely to be smaller than $\sim$100 km s^{-1}. Matter in the IGM is clustered on scales smaller than $\sim$300 km s^{-1} or about $4\,h^{-1}$ comoving Mpc. The simulated CFs are consistent with the observed analogous quantities at the 1–2σ level for this particular sample.

(ii) Thanks to the increased S/N for some of the spectra in our sample the enhanced clustering signal measured in Paper I with the CCC at a transverse velocity separation $\Delta v_{\perp} \sim 500$ km s^{-1} is no longer significant.

Three or more point statistics

(iii) Significant coincidences of Lyman α absorptions have been detected among the lines of sight forming the Sextet implying the presence of coherent gas structures extending $\sim$14 h^{-1} comoving Mpc. In particular an excess of triplets and quadruplets of lines within $\Delta v = 100$ km s^{-1} has been measured at a significance of 16 and 9σ. Besides, one group of five coincident lines in the S2–S3–S4–S5–S6 QSOs is observed, an occurrence that has a probability $P = 0.013$ to arise from a random distribution of lines.

(iv) A method for the detection of underdense regions in relatively low-S/N spectra has been developed. One cosmic common underdense region has been detected in each QSO group; in the Sextet the underdense region has a dimension of 10.7 h^{-1} comoving Mpc and a significance level of 91 per cent, while in the Triplet it has a dimension of 8.8 h^{-1} comoving Mpc and a significance level of 86 per cent. These values are significantly smaller than those typical of underdense regions detected along single lines of sight. The underdense region common to the lines of sight of the Sextet can be parametrized by a sphere of radius 6.75 h^{-1} comoving Mpc.

This study of the cosmic web environment at $z \sim 2$ will soon be extended with many more QSO spectra either at medium resolution with the X-Shooter spectrograph (Kaper et al. 2009) or/and with the $R \sim 10^5$ QSO spectra provided by SDSS-III (Schlegel et al. 2007). The ultimate goal is the characterization of the topology of the IGM in real space in a variety of environments by using Lyman α and metal lines. On this basis, both hydrodynamical simulations of structure formation and high-resolution spectroscopic samples, like the one presented here, will provide the link between observational quantities and the underlying density field and shed light on the impact of systematic effects.

REFERENCES

Aracil B., Petitjean P., Smette A., Surdej J., Mucket J. P., Cristiani S., 2002, A&A, 391, 1
Ballester P., Modigliani A., Boitquin O., Cristiani S., Hanuschik R., Kaufer A., Wolf S., 2000, The Messenger, 101, 31
Becker G. D., Sargent W. L. W., Rauch M., 2004, ApJ, 613, 61
Bianchi S., Cristiani S., Kim T. S., D'Odorico S., 2003, in Shaver P. A., DiLella L., Giménez A., eds, Proc. ESO/CERN/ESA Symp., Astronomy, Cosmology and Fundamental Physics. Springer, Berlin, p. 417
Bryan G. L., Machacek M. E., 2000, ApJ, 534, 57
Carswell R. F., Rees M. J., 1987, MNRAS, 224, 13
Coppolani F. et al., 2006, MNRAS, 370, 1804
Cristiani S., D'Odorico S., Fontana A., Giallongo E., Savaglio S., 1995, MNRAS, 273, 1016
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Crotts A. P. S., 1987, MNRAS, 228, 41
Davè R., Hernquist L., Katz N., Weinberg D. H., 1999, ApJ, 511, 521
D'Odorico V., Cristiani S., D'Odorico S., Fontana A., Giallongo E., Shaver P., 1998, A&A, 339, 678
D'Odorico V., Viel M., Saitta F., Cristiani S., Bianchi S., Boyle B., Lopez S., Maza J., Outram P., 2006, MNRAS, 372, 1333 (Paper I)
D'Odorico V., Bruscoli M., Saitta F., Fontanot F., Viel M., Cristiani S., Monaco P., 2008, MNRAS, 379, 1727
Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 681, 831
Francis P. J., Catherine L. D., Matthew T. W., Michael J. W., Rachel L. W., 2001, Publ. Astron. Soc. Australia, 18, 221
Haardt F., Madau P., 1996, ApJ, 461, 20
Janknecht E., Reimers D., Lopez S., Tytler D., 2006, A&A, 458, 427
Kaper L. et al., 2009, in Moorwood A., ed., Proc. ESO Workshop on 'Science with the VLT in the E-ELT Era'. Springer, Berlin, p. 319
Kim T.-S., Cristiani S., D'Odorico S., 2001, A&A, 373, 757
Kim T.-S., Carswell R. F., Cristiani S., D'Odorico S., Giallongo E., 2002, MNRAS, 335, 555
Kim T. S., Viel M., Haehnelt M., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Kirkman D. et al., 2005, MNRAS, 360, 1373
McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471
Ostriker J. P., Bajtlik S., Duncan R. C., 1988, ApJ, 327, 35
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
Rollinde E., Petitjean P., Pichon C., Colombi S., Aracil B., D'Odorico V., Haehnelt M. G., 2003, MNRAS, 341, 1279
Savaglio S. et al., 1999, ApJ, 515, L5
Scannapieco E., Pichon C., Aracil B., Petitjean P., Thacker R. J., Pogosyan D., Bergeron J., Couchman H. M. P., 2006, MNRAS, 3365, 615

Schaye J., Aguirre A., Kim T. S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768

Schlegel D. J. et al., 2007, BAAS, 38, 966

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

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

Tripp T. M., Limin Lu, Savage B. D., 1998, ApJ, 508, 200

Viel M., Matarrese S., Mo H. J., Haehnelt M., Theuns T., 2002, MNRAS, 329, 848

Viel M., Haehnelt M. G., Springel V., 2004, MNRAS, 354, 684

Viel M., Colberg J. M., Kim T. S., 2008, MNRAS, 386, 1285

This paper has been typeset from a TeX/LaTeX file prepared by the author.