



**HAL**  
open science

## Three- and two-point spatial correlations of intergalactic medium at $z \sim 2$ using projected quasar triplets

Soumak Maitra, Raghunathan Srianand, Patrick Petitjean, Hadi Rahmani, Prakash Gaikwad, Tirthankar Roy Choudhury, Christophe Pichon

► **To cite this version:**

Soumak Maitra, Raghunathan Srianand, Patrick Petitjean, Hadi Rahmani, Prakash Gaikwad, et al.. Three- and two-point spatial correlations of intergalactic medium at  $z \sim 2$  using projected quasar triplets. Monthly Notices of the Royal Astronomical Society, 2019, 490 (3), pp.3633-3653. 10.1093/mnras/stz2828 . hal-02350199

**HAL Id: hal-02350199**

**<https://hal.science/hal-02350199>**

Submitted on 22 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Three- and two-point spatial correlations of intergalactic medium at $z \sim 2$ using projected quasar triplets

Soumak Maitra<sup>1</sup>,<sup>\*</sup> Raghunathan Srianand,<sup>1</sup> Patrick Petitjean,<sup>2</sup> Hadi Rahmani,<sup>3</sup>  
Prakash Gaikwad<sup>1</sup>,<sup>4,5</sup> Tirthankar Roy Choudhury<sup>1</sup>,<sup>6</sup> and Christophe Pichon<sup>2,7</sup>

<sup>1</sup>*IUCAA, Postbag 4, Ganeshkhind, Pune 411007, India*

<sup>2</sup>*Institut d'Astrophysique de Paris, CNRS-SU, UMR 7095, 98bis bd Arago, Paris F-75014, France*

<sup>3</sup>*GEPI, Observatoire de Paris, PSL Université, CNRS, 5 Place Jules Janssen, Meudon F-92190, France*

<sup>4</sup>*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

<sup>5</sup>*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

<sup>6</sup>*National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune 411007, India*

<sup>7</sup>*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*

Accepted 2019 October 3. Received 2019 September 30; in original form 2019 July 1

## ABSTRACT

We present analysis of two- and three-point correlation functions of Ly $\alpha$  forest (at  $2 \leq z \leq 2.5$ ) using X-Shooter spectra of three background quasar triplets probing transverse separations of 0.5–1.6 pMpc. We present statistics based on transmitted flux and clouds identified using Voigt profile fitting. We show that the observed two-, three-point correlation functions and reduced three-point correlation (i.e. Q) are well reproduced by our simulations. We assign probabilities for realizing all the observed correlation properties simultaneously using our simulations. Our simulations suggest an increase in correlation amplitudes and Q with increasing  $N_{\text{HI}}$ . We roughly see this trend in the observations too. We identify a concurrent gap of 17 Å (i.e. 14.2 h<sup>-1</sup> cMpc, one of the longest reported) wide along one of the triplets. Such gap is realized only in 14.2 per cent of our simulated sightlines and most of the time belongs to a void in the matter distribution. In the second triplet, we detect Damped Ly $\alpha$  systems (DLAs) along all three sightlines (with spatial separations 0.64 to 1.6 pMpc) within a narrow redshift interval (i.e.  $\Delta z = 0.088$ ). Detection of a foreground quasar ( $\sim 1$  pMpc from the triplet sightlines) and excess partial Lyman Limit systems (with  $N_{\text{HI}}$  in the range  $10^{16-17.2}$  cm<sup>-2</sup>) around these DLAs suggest that we may be probing a large overdense region. We also report positive C IV–C IV correlations up to  $\sim 500$  km s<sup>-1</sup> only in the longitudinal direction. Additionally, we conclude a positive C IV–Ly $\alpha$  correlations for higher  $N_{\text{HI}}$  thresholds ( $> 10^{15}$  cm<sup>-2</sup>) up to a scale of  $\sim 1000$  km s<sup>-1</sup> both in transverse and longitudinal directions.

**Key words:** intergalactic medium – large-scale structure of the universe.

## 1 INTRODUCTION

The Ly $\alpha$  forest absorption seen in the spectra of distant quasars directly probe structures in the intergalactic medium (IGM) and are therefore used to constrain (i) primordial density fluctuations (Bi, Boerner & Chu 1992; McDonald 2003), (ii) cosmic re-ionization (Fan et al. 2006; Worseck et al. 2018), (iii) thermal history of the universe (Hui & Gnedin 1997; Schaye et al. 2000), and (iv) the impact of various feedback processes (such as SNe and AGN driven outflows) on the IGM that operate during the formation and evolution of galaxies over cosmic time (Aguirre et al. 2001; Oppenheimer & Davé 2006). Numerical simulations and analytical

modelling of a warm photoionized IGM in the framework of  $\Lambda$ CDM models successfully reproduce many observational properties of the Ly $\alpha$  forest absorption: the column density ( $N_{\text{HI}}$ ) distribution, the Doppler b-parameter distribution, the flux probability distribution function, power spectrum of transmitted flux, and the redshift evolution of absorption lines above a certain  $N_{\text{HI}}$  threshold (see Cen et al. 1994; Petitjean, Mueket & Kates 1995; Springel 2005; Smith et al. 2011; Bolton et al. 2012; Rudie, Steidel & Pettini 2012; Gaikwad et al. 2017a, b). Through these models we can constrain the H I photoionization rate, mean IGM temperature, and effective equation of state over a large redshift range (see, for example, Schaye et al. 2000; Becker & Bolton 2013).

In these simulations, the H I density fluctuations responsible for the Ly $\alpha$  absorption closely trace those of the underlying dark matter density field on scales larger than the pressure smoothing scales

\* E-mail: soumak93@gmail.com

(>few 100 pKpc; see, for example, Miralda-Escudé et al. 1996; Schaye 2001). Most of the baryons producing absorption having  $N_{\text{HI}} \sim 10^{14}\text{--}10^{15} \text{ cm}^{-2}$  are found to be located in mildly non-linear regimes probed by filaments and sheets at  $z \sim 2$  (Petitjean et al. 1995). Most of the volume is, however, occupied by underdense regions that produce unsaturated Ly $\alpha$  absorption lines if any. Therefore a rich insight into the morphological properties of the cosmic web (filaments, sheets, voids, and connection between them), but also the large-scale ionization and chemical inhomogeneities and their redshift evolution can be gained by simultaneous analysis of Ly $\alpha$  absorption detected along closely spaced sightlines. Virtual experiments performed on simulated data suggest that it is possible to reconstruct a 3D map of the full density field using a dense enough grid of spatially close lines of sight (see Pichon et al. 2001; McDonald 2003; Caucci et al. 2008; Lee et al. 2018).

Using a group of background quasars, we can probe different physical processes at different scales: (i) The HI density and velocity fields at the scales of few 100 pKpc may have thermal memory of cosmic re-ionization in the form of pressure broadening (Peeples et al. 2010a, b; Rorai et al. 2017); (ii) at the scale of  $\sim 1$  pMpc, we can probe matter clustering around massive galaxies (quasar hosts and intervening metal systems), and various feedback processes connecting gas flows between galaxies and the IGM. At present these scales are best probed using quasar pairs (see, for example, Prochaska et al. 2013); (iii) at the scales of one to few Mpc, one is probing the cosmic structure of filaments and voids and the effect of radiative feed back from bright persistent objects like quasars (e.g. Finley et al. 2014; Lee et al. 2018); and (iv) the Baryon Acoustic Oscillations (BAO) at  $\sim 100$  Mpc probes primordial density fluctuations at very large scales (Ata et al. 2018).

Our understanding of small-scale Ly $\alpha$  clustering is mainly dominated by observations of quasar pairs and gravitationally lensed images of quasars (Rauch & Haehnelt 1995; Smette et al. 1995; Petitjean et al. 1998; Aracil et al. 2002; Coppolani et al. 2006; D’Odorico et al. 2006). Cappetta et al. (2010) have studied the 3D distribution of Ly $\alpha$  forest at  $z \sim 2$  using high-resolution spectroscopy of the quasar pairs, one triplet and a sextet. However, such high signal-to-noise ratio (SNR) and high-spectral resolution studies of triplets or multiple quasar sightlines are rare due to lack of bright targets. Lee et al. (2014) showed a tomographic reconstruction of Ly $\alpha$  3D absorption field in the redshift range of  $2.2 \leq z \leq 2.45$  using 24 star-forming galaxies within  $5 \times 15$  arcmin of the COSMOS field with a spatial resolution of  $\sim 3.5 \text{ h}^{-1} \text{ cMpc}$ . Krolewski et al. (2018) presented 3D HI density field reconstruction at  $z \sim 2$  using wiener filtering technique (with an effective smoothing scales of  $2.5 \text{ h}^{-1} \text{ cMpc}$ ) applied to Ly $\alpha$  absorption detected in moderate resolution spectra of high- $z$  star-forming galaxies. Recent approaches in this field at  $z > 2$  include Horowitz et al. (2019), which presents a Tomographic Absorption Reconstruction and Density Inference Scheme (TARDIS) giving a more accurate reconstruction at smaller scales than wiener filtering techniques and Porqueres et al. (2019), which presents a Bayesian way of inferring 3D matter distribution and their dynamics. While we wait for the arrival of extremely large telescopes to probe IGM tomography over a large range of scales with better sensitivity (Skidmore et al. 2015; Evans et al. 2016), SDSS quasar catalogue offers rare possibilities to make some important progress. Recently, Tie et al. (2019) has made a theoretical predication of three-point correlation in Ly $\alpha$  forest at large scales ( $10\text{--}30 \text{ h}^{-1} \text{ Mpc}$ ).

In this paper, we present detailed analysis of spatial correlations of Ly $\alpha$  forest and metal absorption lines using our X-Shooter spectra along the line of sight towards two quasar triplets and a quasar

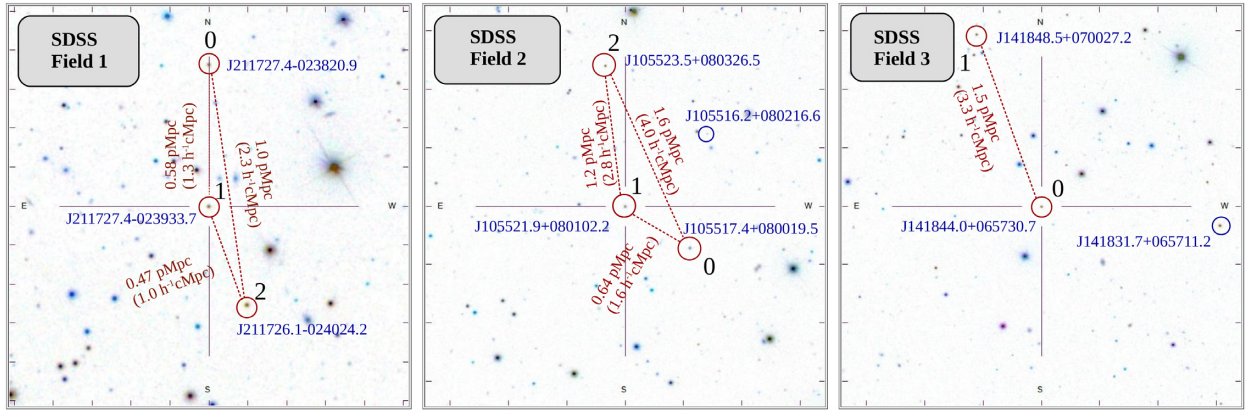
pair (see Fig. 1). The main motivation of our work is to measure the two- and three-point correlations between the Ly $\alpha$  absorption detected along these sightlines. Then quantify the reduced three-point function that will enable us to probe the non-gaussianity in the HI distribution observationally as this is analogous to the skewness of a distribution from normal behaviour. With hydrodynamical simulations at our aid, we can also make a theoretical prediction of this non-gaussianity based on large realizations of the simulated triplets with the observed configurations. Additionally, we will probe the gas distribution around the foreground quasars along the sightlines of the background quasars. Till now this is done using quasar doublets. Presence of additional sightline in the case of triplets allows us to probe the anisotropy of the gas distribution (or ionizing radiation) with additional constraints.

Details of the quasars studied here are summarized in Table 1. The ‘Field 1’ consists of quasars J211727.4–023820.9 ( $z_{em} = 2.323$ ), J211727.4–023933.7 ( $z_{em} = 2.309$ ), and J211726.1–024024.2 ( $z_{em} = 2.309$ ). We will subsequently refer to these as ‘Triplet 1’. The spatial separation between the sightlines ranges from 0.47 to 1.00 pMpc. The ‘Field 2’ consists of quasars J105517.4+080029.5 ( $z_{em} = 2.897$ ), J105521.9+080102.2 ( $z_{em} = 2.709$ ), and J105523.2+080326.5 ( $z_{em} = 2.627$ ) with the spatial separations probed in the range 0.63–1.6 pMpc, which we will subsequently refer as ‘Triplet 2’. In the SDSS-DR12 data base we find a 4th quasar J105516.23+080216.6 ( $z_{em} = 2.320$ ) with a typical separation of 1 pMpc from the other quasars (see Fig. 1). We also study the distribution of HI gas around this quasar along the line of sight to the three background quasars. The ‘Field 3’ consists of quasars J141848.5+070027.2 ( $z_{em} = 2.2305$ ), J141844.03+065730.7 ( $z_{em} = 2.403$ ), and J141831.72+065711.2 ( $z_{em} = 2.389$ ). In this case, we got X-Shooter spectra only for the first two quasars, which we will refer to as the ‘Doublet’. Therefore, our analysis of this triplet is restricted to two-point correlation function of the IGM and gas distribution around the lowest redshift quasar.

This paper is organized as follows. In Section 2, we provide the details of observations, quality of spectra achieved, and properties of the quasars in our sample. In Section 3, we provide details of our simulations and generation of spectra. In Section 4, we validate our simulations by reproducing some observational results. In Section 5, we present the observed transverse and longitudinal two-point and three-point correlations of Ly $\alpha$  forest measured based on transmitted flux as well as using Voigt profile decomposed ‘clouds’. We also quantify the probability of obtaining sightlines in our simulations similar to what we find along the triplets discussed here. In addition to this, we also study the distribution of coherent gaps. In Section 6, we present QSO–Ly $\alpha$ , DLA–Ly $\alpha$ , and C IV transverse correlations. Our main results are summarized in Section 7.

## 2 DETAILS OF OBSERVATIONS

Spectra of quasars were obtained with X-Shooter (Vernet et al. 2011) at the European Southern Observatory (ESO) Very Large Telescope (VLT) in service mode under the programme ID: 096.A-0193 (PI: Petitjean). The X-Shooter spectrograph covers a wavelength range of  $0.3\text{--}2.3 \mu\text{m}$  at medium resolution in a simultaneous use of three arms in UVB, VIS, and NIR. To have a robust sky subtraction, the nodding mode was used following an ABBA scheme. Slit width of 1.2 arcsec was used for all arms of X-Shooter for all our observations. This choice of slit widths results in formal spectral resolutions of 4000, 6700, and 3900 for the UVB, VIS, and NIR, respectively. However, under good seeing conditions, where the QSO point spread function (PSF) is less than slit width, the spectral



**Figure 1.** Fields of quasars used in this study. Red circles mark the position of quasars with X-Shooter spectra. Blue circles mark quasars for which we have only the SDSS spectra. The physical (as well as comoving) distances quoted in the figure are computed at the lowest emission redshift among the triplets.

**Table 1.** Details of quasars studied in this work.

QSO	$z_e$	$\lambda$ range (Å)	SNR	FWHM (km s <sup>-1</sup> )	$L_{912\text{Å}}$ (10 <sup>49</sup> ergs <sup>-1</sup> Å <sup>-1</sup> )	$r_{eq}$ (pMpc)
<b>Field 1</b> ( $r_{01} = 0.579$ pMpc, $r_{12} = 0.472$ pMpc, $\theta = 160^\circ$ )						
J211727.4–023820.9 (X-Shooter)	2.3230	3460–3940	17.3	50	17.1	1.60
J211727.4–023933.7 (X-Shooter)	2.3090	3460–3940	19.7	47	9.5	1.20
J211726.1–024024.2 (X-Shooter)	2.3090	3460–3940	18.0	51	33.6	2.26
<b>Field 2</b> ( $r_{01} = 0.635$ pMpc, $r_{12} = 1.155$ pMpc, $\theta = 130^\circ$ )						
J105517.4+080019.5 (X-Shooter)	2.897	4055–4333	25.3	68	43.9	2.77
J105521.9+080102.2 (X-Shooter)	2.709	4055–4333	13.1	73	43.4	2.72
J105523.5+080326.5 (X-Shooter)	2.627	4055–4333	28.8	73	58.6	3.12
J105516.2+080216.6 (SDSS)	2.320	–	–	–	3.2	0.69
<b>Field 3</b> ( $r_{01} = 1.520$ pMpc)						
J141831.7+065711.2 (SDSS)	2.389	–	–	–	106.4	4.05
J141844.0+065730.7 (X-Shooter)	2.4030	3490–3856	16.7	73	41.0	2.53
J141848.5+070027.2 (X-Shooter)	2.2305	3490–3856	27.7	63	33.8	2.23

The columns 1–7 provide quasar name, emission redshift, wavelength range used for our analysis, median SNR, instrumental resolution, ionizing luminosity, and the radius of influence of the quasar radiation, respectively.

full width at half-maximum (FWHM) is better than the predicted ones. Hence, for such cases we use the method described in Krogager et al. (2017) to obtain the effective spectral resolution.

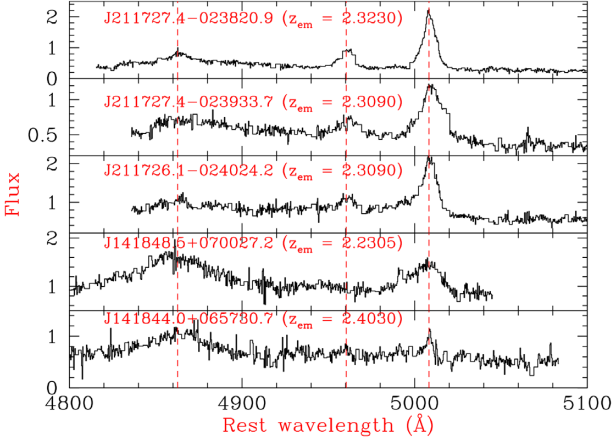
We have used the X-Shooter Common Pipeline Library (Goldoni et al. 2006) release 6.5.1 for reducing the science raw images and produce the final 2D spectra of the QSOs. We first compute an initial guess for the wavelength solution and position of the centre and edges of the orders. Then we trace the accurate position of the centre of each order and follow this step by generating the master flat frame out of five individual lamp flat exposures. Next we find a 2D wavelength solution and modify it by applying a flexure correction to account for the shifts that can be of the order of the size of a pixel. Finally, having generated the required calibration tables we reduce each pair of science frames to obtain the flat-fielded, wavelength calibrated, and sky subtracted 2D spectrum. To extract the 1D flux of the QSO, we follow a spectral point spread function (SPSF) method as described in Rahmani et al. (2016). In summary we model the QSO’s PSF using a Moffat function, which is a smooth function defined by the centroid wavelength and FWHM. We then integrate the light profile at each wavelength pixel to obtain the flux of the QSO.

We fit the continuum to the 1D extracted spectra using lower order polynomial smoothly connecting the identified absorption line free regions. In Table 1, we summarize various details of quasars used in

our study. For each triplet we provide  $r_{01}$  (the projected separation between the first and the second quasar of the triple as listed in column 1),  $r_{12}$  (the projected separation between quasar 2 and 3), and  $\theta$  (the angle between the two pairs). The median SNR obtained and typical FWHM of our SPSF (obtained as discussed above) are given in columns 4 and 5, respectively.

## 2.1 Quasar redshifts using narrow emission lines

The systemic redshift of a quasar is better determined using the low-ionization broad and narrow emission lines (Gaskell 1982). In our X-Shooter NIR spectra [O III], H $\beta$ , and H $\alpha$  emission lines are clearly detected for all the quasars in ‘Triplet 1’ (see left-hand panels in Fig. 2). This together with the detection of [C III] and Mg II emission lines in the VIS spectrum allow us to measure the systemic redshift of the quasars accurately. The second column in Table 1 gives our measurement of the systemic redshift of the quasar. Based on the NIR spectrum both J2117–0240 and J2117–0239 are at the same redshift with a projected separation of 472 pKpc. Thus these quasars could be physically associated with each other. We do find signatures of Ly $\alpha$  extended emission for them and follow-up studies could reveal their association with Ly $\alpha$  halo, similar to Hennawi et al. (2015). All three quasars show associated absorption (see Fig. A1



**Figure 2.** Rest-frame spectrum showing H $\beta$  and [O III] emission lines detected in the spectra of 5 quasars in our sample. Vertical dashed lines mark the locations of emission lines. The systemic redshift is given in each panel.

in the Appendix available online). In the case of J2117–0238 and J2117–0240 narrow C IV associated absorption are seen with both inflowing and outflowing signatures. In the case of J2117–0239 the C IV absorption is consistent with broad absorption lines (BALs) with some of the narrow components at very large ejection velocities (i.e.  $v \sim 4500 \text{ km s}^{-1}$ ) showing signatures of partial coverage.

In the case of ‘Triplet 2’, we do not clearly detect [O III], H $\beta$ , or H $\alpha$  emission lines in the NIR spectrum. Therefore, we base our systemic redshift determination mainly on the Mg II and [C III] broad emission lines. Unlike quasars in ‘Triplet 1’, in this case the redshifts of the quasar are very different. The quasar J1055+0800 shows associated broad C IV absorption with clear signatures of high velocity outflowing components. In the case of J1055+0801 and J1055+0803 narrow associated C IV absorption is clearly detected. In the case of two quasars in ‘Doublet’, in our NIR spectra we clearly detect [O III], H $\beta$ , and H $\alpha$  emission lines (see Fig. 2). These lines were used to determine the systemic redshift of the quasars.

We consider the common redshift range between Ly $\alpha$  and Ly $\beta$  emission lines of the quasars avoiding the proximity regions (i.e. within  $5000 \text{ km s}^{-1}$  to the quasar redshift) for our IGM correlation studies. These wavelength ranges are provided in column 3 of Table 1. We also avoid Ly $\alpha$  of the associated absorption systems.

We identified C IV and Mg II doublets and DLAs in all our spectra. The redshifts of these absorbers are summarized in Table B1 in the Appendix which is available online. We mask the Ly $\alpha$  range that may be contaminated by metal absorption associated with these redshifts during our correlation analysis.

## 2.2 Voigt profile fitting

It is a regular procedure to decompose the Ly $\alpha$  forest into multiple Voigt profile components, i.e. distinct absorbers parametrized by  $z$ ,  $N_{\text{HI}}$ , and  $b$ . We use the automated Voigt profile fitting parallel code VIPER to identify the Ly $\alpha$  absorption lines and obtain the column density and line width of the absorbers through multicomponent Voigt profile fitting (see Gaikwad et al. 2017b, for details regarding VIPER). The code assigns a rigorous significance level (RSL; as described in Keeney et al. 2012) to these fitted absorption features. We consider only Voigt profile components for which the  $\text{RSL} > 3$  to avoid false identifications. For simplicity we denote individual Voigt profile components as ‘clouds’. Voigt profile fits to all the

observed quasar sightlines used in our study are shown in Figs C1, C2, and C3 in the Appendix available online. One of us (HR) fitted the Ly $\alpha$  forest using VPFIT (Carswell & Webb 2014), and we found a good agreement between the decompositions obtained using VIPER and VPFIT.

## 2.3 Ionization sphere of influence of quasars

Assuming an isotropic continuum emission and for a given quasar spectral energy distribution (SED), we can compute the radius of influence of quasar ionization for a given metagalactic UV ionizing background (here we assume the one computed by Khaire & Srianand 2019). The H I photoionization rate due to quasar at distance  $r$  is given as

$$\Gamma(\text{H I}, r) = \int_{100 \text{ \AA}}^{912 \text{ \AA}} \frac{L_{\lambda}/4\pi r^2}{hc/\lambda} \sigma_{\lambda}(\text{H I}) d\lambda, \quad (1)$$

where  $L_{\lambda}$  is the specific H I ionizing luminosity ( $\text{ergs}^{-1} \text{ \AA}^{-1}$ ) of the quasar, and  $\sigma_{\lambda}(\text{H I})$  is the wavelength dependent ionization cross-section for H I by photons with energy above 13.6 eV. We have assumed the UV SED of our quasars as adopted by Khaire & Srianand (2019) for computing their UV background. We have taken the far-UV spectral index  $\alpha = -1.8$  for the flux calculation. The Lyman continuum luminosity inferred for each quasar using the observed flux at rest-frame  $\lambda \sim 1450 \text{ \AA}$  is given in 6th column of Table 1. The ionization radius is then defined to be the radius ( $r_{\text{eq}}$ ) at which  $\Gamma(r, \text{H I})$  from the quasar is equal to the background photoionization rate. The computed  $r_{\text{eq}}$  values for all the quasars are given in the last column of the Table 1. It is clear from the table that all the quasars in the ‘Triplet 1’ and ‘Triplet 2’ will influence the ionization state of the IGM along the other two sightlines. However, this may not be the case for J1055+0802 found close to the sightlines along ‘Triplet-2’ as inferred  $r_{\text{eq}}$  is less than the separation between this quasar and the nearest quasar sightline.

## 3 SIMULATION

We use the smoothed particle hydrodynamical code GADGET-3 (a modified version of the publicly available GADGET-2<sup>1</sup> code; see Springel 2005) for generating  $100 \text{ h}^{-1} \text{ cMpc}$  simulation box with  $2 \times 1024^3$  particles. We use standard flat  $\Lambda$ CDM cosmology with parameters  $(\Omega_{\Lambda}, \Omega_m, \Omega_b, h, n_s, \sigma_8, Y) \equiv (0.69, 0.31, 0.0486, 0.674, 0.96, 0.83, 0.24)$ . The initial conditions are generated at  $z = 99$  using the publicly available 2LPT<sup>2</sup> (Scoccimarro et al. 2012) code. The gravitational softening length has been taken as 1/30th of the mean interparticle separation. The GADGET-3 simulation incorporates radiative heating and cooling of SPH (Smoothed Particle Hydrodynamic) particles internally for a given UV background assuming ionization and thermal equilibrium, but solves time-dependent temperature evolution equation. In our case, we have used the ionization and heating rates as given by Khaire & Srianand (2019) for the assumed far-UV spectral index of  $\alpha = -1.8$ . In order to run the simulation faster, we also use the QUICK\_LYALPHA flag in the simulation, which converts gas particles with  $\Delta > 10^3$  and  $T < 10^5 \text{ K}$  to stars (see Viel, Haehnelt & Springel 2004). The simulation does not include AGN or stellar feedback or galactic outflows. We have stored the simulation outputs between  $z = 6$  and  $z = 1.8$  with

<sup>1</sup><http://wwwmpa.mpa-garching.mpg.de/gadget/>.

<sup>2</sup><https://cosmo.nyu.edu/roman/2LPT/>.

a redshift interval of 0.1. Considering the median redshift intervals, we use simulation box at  $z = 2$  and  $z = 2.5$  for ‘Triplet 1’ and ‘Triplet 2’, respectively. In these redshifts our assumed box size provide a line-of-sight wavelength coverage of  $\sim 122$  and  $151.4 \text{ \AA}$ , respectively. The resolution of the final simulation spectrum we obtain is sufficient to resolve the features in X-Shooter spectrum. As our main focus in this paper is to study the observed correlations, we restrict our self to this simulation only. In our future work, we will provide details of how three-point statistics in the simulations are influenced by the assumed radiation and thermal history, cosmic variance, cosmological parameters, box size, and other convergence issues.

### 3.1 Transmitted flux

To generate triplet (or a doublet) sightlines having configurations similar to the observed one, we place the triplet (or doublet) source configuration along one face of the box and shoot lines of sight (LOS) parallel to the other faces. For the generated LOSs, neutral hydrogen density  $n_{\text{HI}}$ , temperature  $T$ , and the peculiar velocity  $v$  are assigned along the LOS using SPH smoothing of the nearby particle values. We typically sample each line of sight with 2048 equally sampled grids in comoving length of the box. According to the SPH formulation (Monaghan 1992; Springel 2005), the value of a quantity  $f_i$  at the  $i$ th grid point is expressed as

$$f_i = \sum_j f_j \frac{m_j}{\rho_j} W_{ij}, \quad (2)$$

where the summation is done over all the particles.  $f_j$ ,  $m_j$ , and  $\rho_j$  are the values of the quantity, mass, and density of the  $j$ th particle, respectively.  $W_{ij}$  is the SPH kernel which is a window function that depends on the distance between the  $i$ th grid and  $j$ th particle ( $r_{ij}$ ) and the smoothing length  $h_j$ . We use the SPH kernel of Springel (2005):

$$W(r, h) \equiv \frac{8}{\pi h^3} \begin{cases} 1 - 6 \left(\frac{r}{h}\right)^2 + 6 \left(\frac{r}{h}\right)^3, & \text{if } 0 \leq \frac{r}{h} \leq \frac{1}{2} \\ 2 \left(1 - \frac{r}{h}\right)^2, & \text{if } \frac{1}{2} \leq \frac{r}{h} \leq 1 \\ 0 & \text{if } \frac{r}{h} > 1 \end{cases} \quad (3)$$

Next, using the  $n_{\text{HI}}$ , temperature, and velocity fields, we obtain the Ly $\alpha$  optical depth  $\tau$  along the sightlines (see equation 30 of Choudhury, Srianand & Padmanabhan 2001). The Ly $\alpha$  transmitted flux  $F$  is obtained as the negative exponential of the optical depth, i.e.  $F = e^{-\tau}$ .

We add the effects of instrumental resolution and noise to the simulated Ly $\alpha$  transmitted flux skewers. The transmitted flux is convolved with the instrumental LSF (line spread function) which we assume as a Gaussian with FWHM  $\sim 50 \text{ km s}^{-1}$  for the ‘Triplet 1’ and FWHM  $\sim 70 \text{ km s}^{-1}$  for ‘Triplet 2’. The data is then rebinned to  $\sim 15 \text{ km s}^{-1}$  pixels to match the pixel sampling in our X-Shooter spectra. Next, to incorporate the effects of noise, we add a simple Gaussian noise to the skewers corresponding to the SNR values mentioned in columns 4 of Table 1. This simulated spectra are then used for all the statistics that are based on the transmitted flux as well as Voigt profile fitting (using VIPER as discussed before).

## 4 VALIDATION OF OUR SIMULATIONS

In this section we try to reproduce some of the well-known properties of the high- $z$  Ly $\alpha$  forest to validate our simulations before applying them to understand the observed spatial correlations.

### 4.1 Flux probability distribution function

First we compare the probability distribution function of the Ly $\alpha$  transmitted flux (flux PDF) obtained in our simulations with the observed ones. The flux PDF is calculated for 20 bins in  $F$  ranging from 0 to 1. As is usually done, values with  $F < 0$  goes in the first bin, while values having  $F > 1$  goes in the last bin of the distribution. The first column in Fig. 3 shows the observed flux PDF. In the top panel we show the results for three sightlines in ‘Triplet 1’. As the spectral SNR is very similar, we consider only one set of simulated results for comparison. In the case of ‘Triplet 2’ SNR achieved along two sightlines are roughly a factor 2 higher than that along J1055+0801. Therefore, we show comparison with the simulations in different panels (second and third from the top). In the case of the ‘Doublet’, results for the two quasar sightlines are summarized in last two rows. It is clear that, by and large, the simulated distributions are consistent with the observed one within  $1\sigma$  confidence level.

### 4.2 Column density distribution function

Next we compare the HI column density distribution function (CDDF). The CDDF  $f(N_{\text{HI}}, X)$ , is defined as the number of HI absorbers within absorption distance interval  $X$  and  $X + dX$  and column density interval  $N_{\text{HI}}$  and  $N_{\text{HI}} + dN_{\text{HI}}$ . The absorption distance is defined as ,

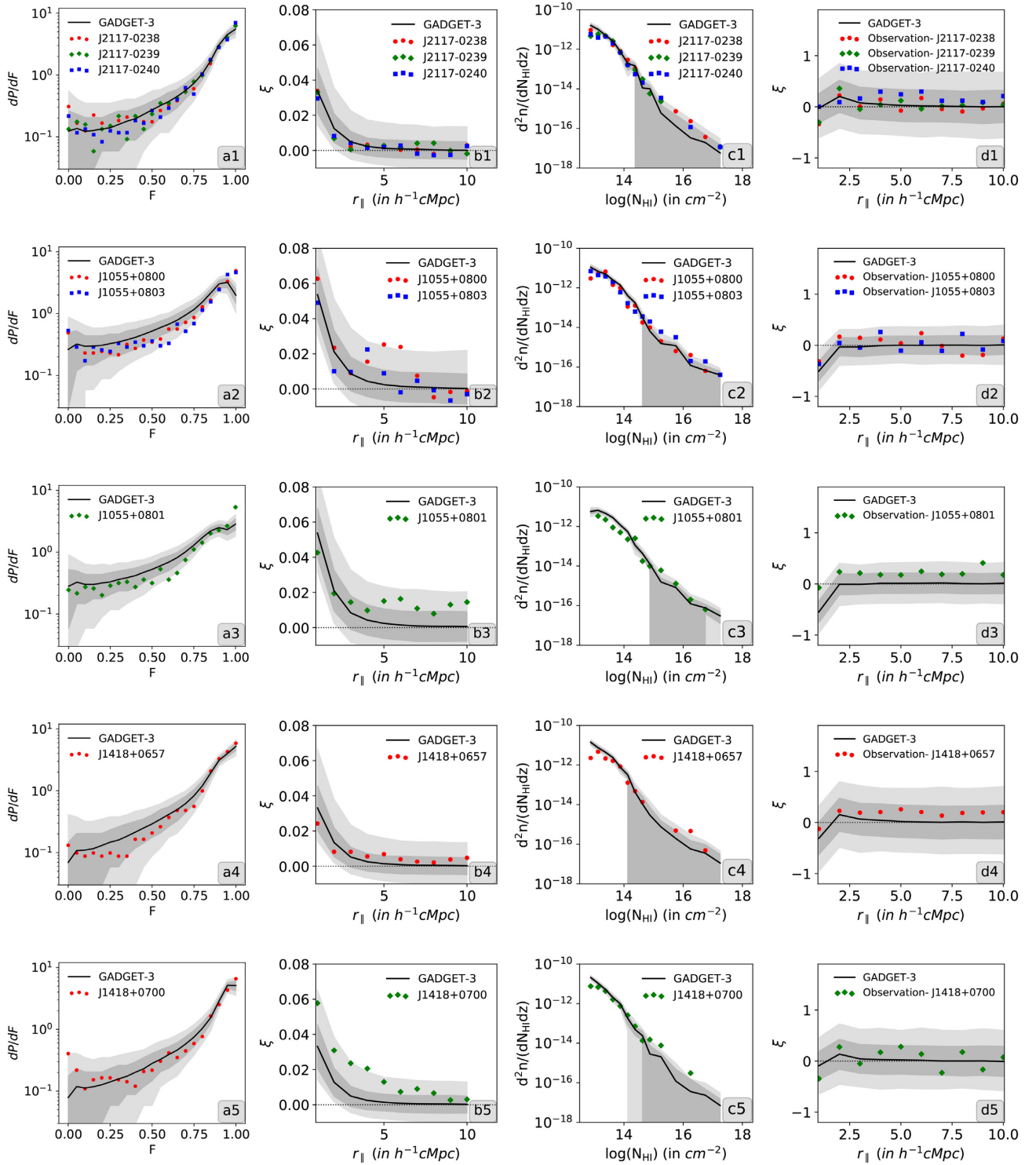
$$X(z) = \int_0^z dz \frac{H_0}{H(z)} (1+z)^2, \quad (4)$$

by Bahcall & Peebles (1969). As a validation of our simulation, we compare the CDDF of our simulated spectra with that observed by Kim et al. (2013) in Fig. 4. The error plotted for Kim et al. (2013) corresponds to  $1\sigma$  range. For simulations, we show the  $1\sigma$  confidence interval due to multiple realizations. Our simulations match well within  $1\sigma$  of the CDDF from Kim et al. (2013).

We compute the CDDF along each of our sightlines and compare them with the observations. These are summarized in panels shown in column 3 of Fig. 3. Apart from CDDF measured along the line of sight to J1055+0801 (where we see less number of absorbers compared to the model predictions) the observed distribution along other sightlines are broadly reproduced. Given the small redshift path involved in each sightline such deviations are not unexpected. For the median SNR achieved in our spectra the typical limiting column density for an unresolved absorption component is  $\log(N_{\text{HI}}) = 12.90$  for quasars in ‘Triplet 1’. In the case of ‘Triplet 2’ the limiting column density is  $\log(N_{\text{HI}}) = 13.10$  for J1055+0801 and  $\log(N_{\text{HI}}) = 12.77$  for the other two sightlines. In the case of ‘Doublet’, the limits are  $\log(N_{\text{HI}}) = 12.95$  and  $\log(N_{\text{HI}}) = 12.74$ , respectively, for J1418+0657 and J1418+0700. Therefore, for the full data set, we have a column density completeness close to  $10^{13} \text{ cm}^{-2}$ .

## 5 SPATIAL CORRELATIONS OF LY $\alpha$ ABSORPTION

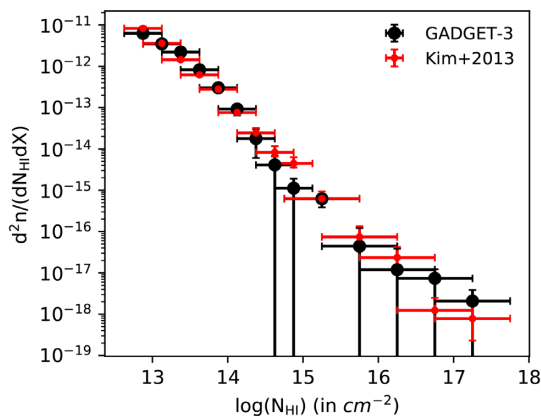
The correlation properties of Ly $\alpha$  forest absorption is usually studied using the statistics of transmitted flux (see Croft et al. 2002; Viel et al. 2002; McDonald 2003; Rollinde et al. 2003, for earlier work). In this work we use statistics based on transmitted flux as well as individual ‘clouds’ identified using Voigt profile fitting. As we are dealing with handful of sightlines, our aim is to mainly quantify how probable are the observed properties in the framework of the



**Figure 3.** Statistics involving single Ly $\alpha$  sightline for the quasar triplets. The deep and light grey regions denote the  $1\sigma$  and  $2\sigma$  confidence interval from the simulations. Column-1: Ly $\alpha$  transmitted flux probability distribution functions. Column-2: Ly $\alpha$  transmitted flux longitudinal correlation function. Column-3: Distribution of  $N_{\text{HI}}$  obtained from the Voigt profile decomposition. Column-4: Longitudinal correlation function for the Voigt profile decomposition of the Ly $\alpha$  forest.

simulations considered and not to match the observations with simulated data by varying model parameters. As we discussed before, unlike observations, our simulated spectra have finite wavelength coverage (limited by our box size). To make realistic comparisons,

in what follows we divide the observed spectrum into random chunks having wavelength intervals similar to that of our simulated spectra (1000 random chunks of the observation). We then compute the mean and probability distribution of the statistical quantity of



**Figure 4.** Comparison of observed (red dots from Kim et al. 2013) and simulated H I CDDF (black dots).

our interest using these random chunks. These values will then be compared with the same obtained from the simulations (typically using 4000 random realizations).

### 5.1 Transmitted flux based statistics

In this section, we investigate two-point and three-point correlation statistics of Ly $\alpha$  transmitted flux as a probe of the spatial clustering of the IGM.

#### 5.1.1 Longitudinal two-point correlation

We define the longitudinal two-point correlation ( $\xi_{\parallel}(\Delta\mathbf{r}_{\parallel})$ ) in transmitted flux decrement,  $D = F - \langle F \rangle$  as,

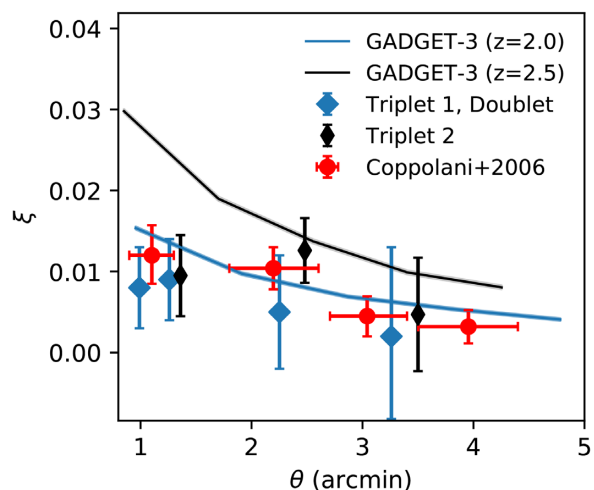
$$\xi_{\parallel}(\Delta\mathbf{r}_{\parallel}) = \langle D(\mathbf{r}'_{\parallel}) D(\mathbf{r}'_{\parallel} + \Delta\mathbf{r}_{\parallel}) \rangle. \quad (5)$$

Here  $\langle F \rangle$  is the mean transmitted flux of the Ly $\alpha$  absorption. Longitudinal two-point correlation measures the flux correlation along the line of sight in the redshift space.

The longitudinal two-point correlation measured as a function of redshift separation are shown in the second column of Fig. 3.

In the top panel we summarize the results for three sightlines in ‘Triplet 1’. The dots are the average  $\xi_{\parallel}$  measured at different values of  $r_{\parallel}$  in observation. The solid curve is the mean value we find from our simulations. Deep and light grey-shaded regions give the  $1\sigma$  and  $2\sigma$  confidence interval measured around the median values from our simulations. It is clear that the observed distribution is well within  $1\sigma$  range of our observations. As the spectral SNR achieved towards all three quasars are similar, we plot them in the same figure.

In the case of ‘Triplet 2’ we notice that the longitudinal two-point correlation function is consistent within  $1\sigma$  range seen in our simulations. The measured longitudinal two-point correlation is also consistent within  $2\sigma$  range predicted by our simulations for J1418+0657. However, for J1418+0700 we find the measured longitudinal two-point correlations deviate by about  $2\sigma$ . From the figure it is also apparent that the longitudinal correlation found for sightlines in ‘Triplet 1’ (that probe similar redshift range) are also lower than what is seen towards J1418+0700. This could imply a strong clustered absorption along this sightline compared to that of J1418+0657.



**Figure 5.** Transverse two-point correlation of transmitted flux as a function of angular separation. The red dots denote the mean measurements by Coppolani et al. (2006) computed from the values given in their table 1. The error reflects the  $1\sigma$  range. Measurements from the present sample are plotted with blue and black diamond symbols. The blue and black curves are from our simulated spectra at  $z \sim 2.0$  and  $z \sim 2.5$ , respectively.

#### 5.1.2 Transverse two-point correlation

Transverse two-point correlation probes the clustering information between spatially separated two sightlines. For a fixed pair separation  $\Delta\mathbf{r}_{\perp}$  between the sightlines, a transverse two-point correlation can be defined as

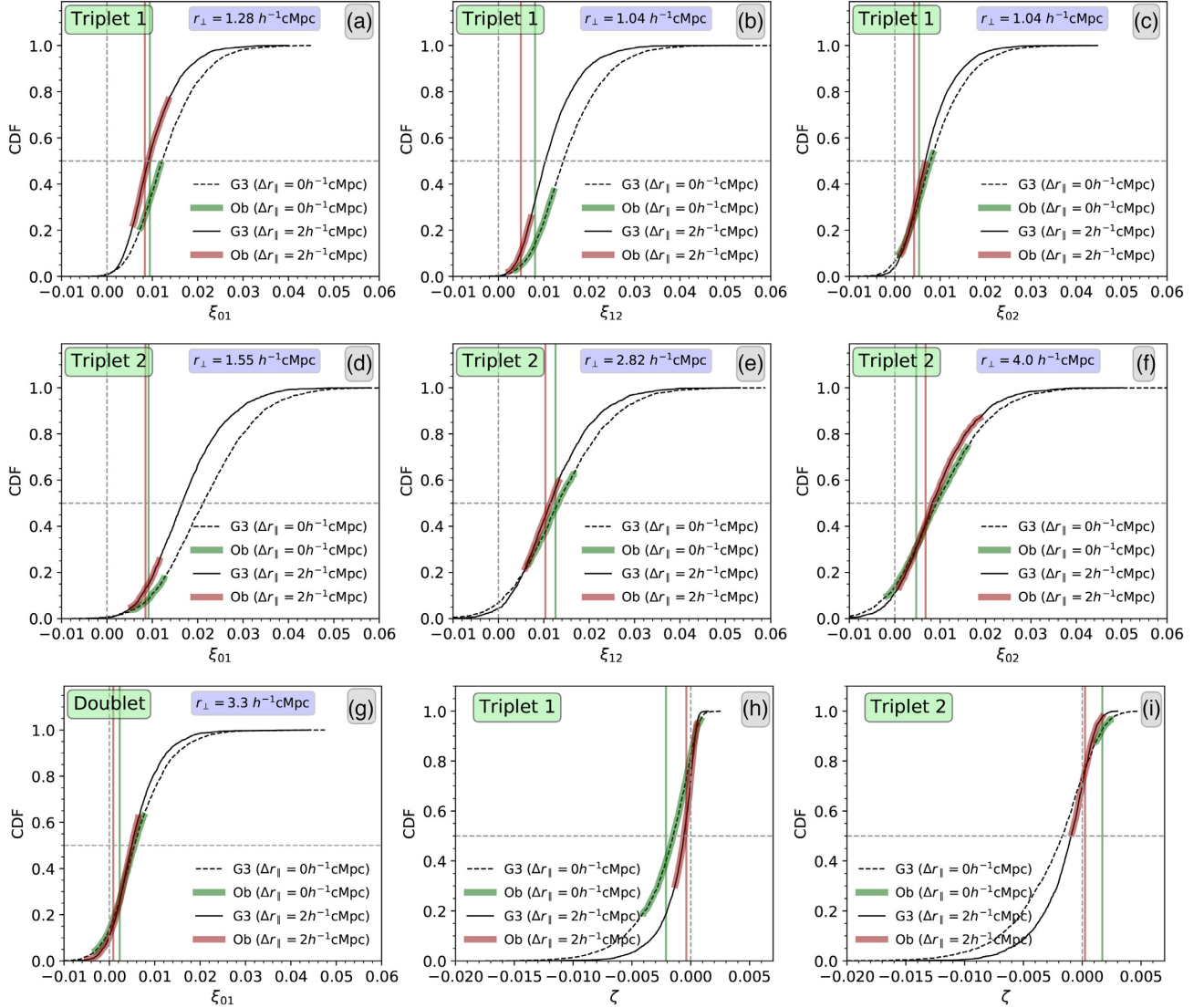
$$\xi(\Delta\mathbf{r}_{\parallel}, \Delta\mathbf{r}_{\perp}) = \langle D_1(\mathbf{r}'_{\parallel}, \mathbf{0}) D_2(\mathbf{r}'_{\parallel} + \Delta\mathbf{r}_{\parallel}, \Delta\mathbf{r}_{\perp}) \rangle. \quad (6)$$

The transverse correlation function can be defined over the actual separation given by  $\Delta r = \sqrt{\Delta r_{\parallel}^2 + \Delta r_{\perp}^2}$ . We will consider transverse two-point correlation functions for  $\Delta r_{\parallel} = 0$  (i.e. correlating flux along two sightlines having same redshift as in Coppolani et al. 2006). In addition to this, we also calculate the transverse two-point correlation by averaging over  $\Delta r_{\parallel} = \pm 2 h^{-1}$  cMpc. This is done to average out the distortions due to peculiar velocity effects along the sightline. Subsequently, we will refer to this as  $\Delta r_{\parallel} = 2 h^{-1}$  cMpc case. Note in the flux based two-point correlation functions both the correlated overdense and underdense regions will have positive correlation amplitudes. However, as the mean transmission is close to the continuum, the mean of the flux two-point correlation will be more influenced by the overdense (i.e. strong absorption) regions. Uncorrelated regions will have negative correlation amplitudes.

We compare the predictions from our simulation with the observed transverse two-point correlation ( $\Delta r_{\parallel} = 0$ ) as a function of angular separations by Coppolani et al. (2006). Note the observations used in Coppolani et al. (2006) were obtained with four times lower resolution (i.e. FWHM  $\sim 220$  km s $^{-1}$ ) than the spectra used in our study and our simulated spectra (see column 5 in Table 1). It is clear from Fig. 5 that our simulated data for  $z = 2$  reproduces the observed trend very well. This is expected as most of the data points with angular separations less than 4 arcmin in Coppolani et al. (2006) sample the Ly $\alpha$  forest in the range  $1.9 \leq z \leq 2.3$ .

Next we consider individual measurements of  $\xi$  using all possible pairs of sightlines from our triplets. These points are also shown in Fig. 5. As explained before, we divided the observed Ly $\alpha$  forest wavelength range into several segments (of size equal to that of the simulated spectrum) and estimated the mean  $\xi$  and its  $1\sigma$





**Figure 6.** Cumulative distribution of two- and three-point transverse correlation functions. ‘G3’ and ‘Ob’ denote GADGET-3 simulation and observation, respectively. The red and green vertical lines denote the mean of the observed correlation. The red and green regions overlaying the simulated curves represent the  $1\sigma$  confidence interval of the observed triplet sightlines.

confidence interval for each doublets. It is clear from Fig. 5 that the observed  $\xi$  values for all the doublets corresponding to ‘Triplet 1’ and ‘Doublet’ are consistent with the simulated results for  $z = 2$ . However, in the case of ‘Triplet 2’ (where we use simulation data for  $z \sim 2.5$ ), while the observed values  $\xi_{12}$  and  $\xi_{02}$  are consistent with the models,  $\xi_{01}$  measured between the sightlines along J1055+0800 and J1055+0801 is lower than the model predictions.

To quantify this further, in Fig. 6 we compare our measurements (mean and  $1\sigma$  confidence interval with thick shaded curves) with the cumulative probability distribution we get from our simulated sightlines (dashed and dotted curves). The solid curves in Fig. 6 are computed for  $\Delta r_{\parallel} = 0$  (as done by Coppolani et al. 2006; and for our data in Fig. 5). The dotted curves in Fig. 6 are obtained by integrating over  $\Delta r_{\parallel}$  between  $\pm 2 h^{-1} \text{cMpc}$ . Vertical dotted lines in each panel shows  $\xi = 0$  and the other two vertical lines provide the observed mean  $\xi$  values for the two cases considered. It is clear from the simulated curves that when we integrate  $\xi$  along  $r_{\perp}$  we notice a decrease in the predicted value of the median and the scatter around the median. The difference between the two cumulative distributions

**Table 2.** Results of flux based correlation analysis for  $\Delta r_{\parallel} = 2 h^{-1} \text{cMpc}$ .

Sample	Correlation	$r_{\perp}$ (cMpc)	Observed values	Probability	Percentile
Triplet 1	$\xi_{01}$	1.28	+0.008	0.54	44
	$\xi_{12}$	1.04	+0.005	0.23	11
	$\xi_{02}$	2.28	+0.004	0.39	28
	$\zeta$		-0.0004	0.65	57
Triplet 2	$\xi_{01}$	1.55	+0.008	0.20	13
	$\xi_{12}$	2.82	+0.010	0.38	44
	$\xi_{02}$	4.00	+0.007	0.73	41
	$\zeta$		+0.0002	0.47	76
Doublet	$\xi_{01}$	3.30	+0.001	0.62	16

are larger for the smaller pair separations. This trend is also roughly evident in the observed distributions as well.

In Table 2 we quantify the comparison between our observations and simulations for the case where  $\xi$  is obtained by integrating over

$r_{ii}$ . Fourth column of this table provides the mean values of  $\xi$  and  $\zeta$  ( $\zeta$  is the transverse three-point correlation defined in Section 5.1.3). Fifth column gives the probability that the observed mean and  $1\sigma$  confidence interval is realized in our simulations. We obtain this by computing the fractional area of the shaded region in the cumulative distributions shown in Fig. 6. The last column in Table 2 gives the percentile of the observed mean  $\xi$  based on the cumulative distribution obtained from the simulations.

In the top row we plot the transverse two-point correlation function for the three possible doublets belonging to the ‘Triplet 1’. It is clear that in the case of ‘Triplet 1’ when we consider  $\xi$  measurement with  $r_{ii} = 0$ , the observed mean and  $1\sigma$  confidence range tend to be in the lower end of the simulated distribution (i.e. typically below the median). This is also the case when we consider integration in  $r_{ii}$  for two of the doublets. However, as can be seen from Table 2, the probability of realizing the observed distributions of  $\xi$  in our simulation box is not very low.

Similarly in the case of ‘Triplet 2’, the measured  $\xi$  is lower than our model prediction for the pair of sightlines towards J1055+0800 and J1055+0801 (i.e.  $\xi_{01}$  in Fig. 6). This doublet also shows the lowest probability in Table 2. For the other two pairs in this triplet the observed  $\xi$  is consistent with the predictions within  $1\sigma$  level. For the only pair for which we have X-shooter data in ‘Doublet’, the measured  $\xi$  is consistent with the model predictions.

In summary, the transverse two-point correlation functions based on transmitted flux measured for all the doublets we could construct in our sample are not abnormal in the framework of our simulations. Next we look at the three-point correlation function based on the transmitted flux.

### 5.1.3 Three-point correlation in flux

If the matter density field is Gaussian then the two-point correlation would be sufficient to describe the spatial distribution of matter. However, due to non-linear nature of evolution of gravitational instabilities and structure formation, the density field is expected to be non-Gaussian, necessitating the usage of three-point correlation  $\zeta(\Delta\mathbf{r}_{1i}, \Delta\mathbf{r}_{2i}, \Delta\mathbf{r}_{1\perp}, \Delta\mathbf{r}_{2\perp}, \theta)$ . For a fixed point in one of the LOS (say LOS 0),  $\Delta\mathbf{r}_{ii}$  denotes the longitudinal separation between this reference point and points of our interest in  $i$ th ( $i = 1, 2$ ) LOS,  $\Delta\mathbf{r}_{i\perp}$  denotes the transverse separation between the reference point and  $i$ th LOS, and  $\theta$  denotes the angle subtended by the lines joining the reference point and two other lines of sight in the sky plane. Hence,  $\Delta\mathbf{r}_{1\perp}$ ,  $\Delta\mathbf{r}_{2\perp}$ , and  $\theta$  denote the quasar triplet configuration in the sky-plane at the same redshift.

For a given configuration of triplet source in the sky plane ( $\Delta\mathbf{r}_{1\perp}$ ,  $\Delta\mathbf{r}_{2\perp}$ ,  $\theta$ ), we define triple correlation  $\zeta(\Delta\mathbf{r}_{1i}, \Delta\mathbf{r}_{2i})$  as

$$\zeta(\Delta\mathbf{r}_{1i}, \Delta\mathbf{r}_{2i}) = \langle D_0(\mathbf{r}_{ii}) D_1(\mathbf{r}_{ii} + \Delta\mathbf{r}_{1i}) D_2(\mathbf{r}_{ii} + \Delta\mathbf{r}_{2i}) \rangle. \quad (7)$$

This gives the triple correlation as a function of two redshift space axes. It is clear from the above equation that three-point correlation function will be negative when all three-points under consideration have strong absorption or only one of them have strong absorption. In all other cases the three-point function will be positive.

In the last row in Fig. 6, we plot the transverse three-point correlation function ( $\zeta$ ) measured with  $\Delta r_{ii} = 0$  and  $2h^{-1}$  cMpc for the Triplets 1 and 2. As we have seen in the case of  $\xi$ ,  $\zeta$  values also go towards zero when we integrate it along the longitudinal directions. In the case of ‘Triplet 1’, the observed distribution has large spread and hence, the probability of realizing the observed  $\zeta$  distribution is also very high (see Table 2). In the case of ‘Triplet 2’

the observations trace the upper end of the probability distribution for both the cases considered. This indicates a slightly weaker clustering in the regions probed by ‘Triplet 2’ in comparison to the simulations. As we discussed before at least one of the pairs in this triplet also shows weaker two-point correlation function.

Thus when we consider individual transverse two-point and three-point function measurements, we find that the observed values are not statistically significant outliers in the framework of simulations considered here. However, what will be more important is to ask what is the probability for simultaneously reproducing the three  $\xi$ s and  $\zeta$ . This we will explore in Section 5.3.

### 5.1.4 Concurrent gap statistics in flux

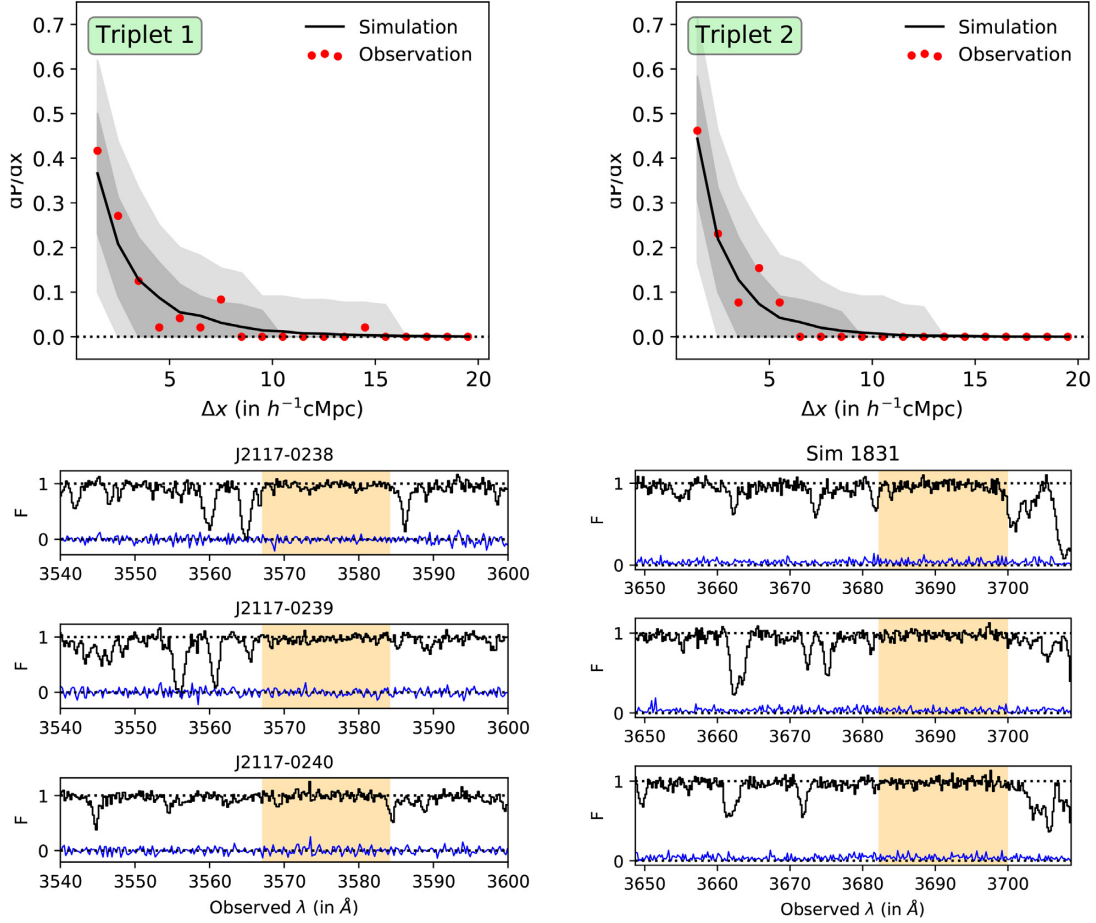
In this section, we probe the presence of large concurrent absorption gaps in Ly $\alpha$  forest along our two triplets. Following Rollinde et al. (2003), for a transmitted flux  $F$  in a wavelength bin to be a gap, it must satisfy the criteria

$$F > \bar{F} - \sigma, \quad (8)$$

where  $\sigma = (\sigma_{\text{noise}}^2 + \sigma_{\text{cont}}^2)^{1/2}$  is the standard deviation due to the noise and error from continuum fitting. We also do a running mean over 3 pixels in the spectra to remove spurious peaks due to noise. The  $\sigma_{\text{noise}}$  is taken to be  $1/\text{SNR}_{\text{min}}$ , where  $\text{SNR}_{\text{min}}$  is the minimum SNR value of the three sightlines. We have performed independent continuum fits to get the residual flux error,  $\sigma_{\text{cont}}$ . A concurrent gap is a wavelength stretch over which our requirement for gap is satisfied along all three sightlines. For simulated sightlines, we simply record the continuous concurrent gap lengths for the three sightlines and then calculate the concurrent gap length probability distribution. For the observed triplets, we select a random wavelength section whose size matches that of the corresponding simulated sightline and calculate the concurrent gap length probability distribution. We do this random selection 1000 times and average over them to get the observed gap length distribution.

In the top panel in Fig. 7, we show the probability distributions of observed (red points) and simulated (solid line) concurrent gaps for ‘Triplet 1’ (top left panel) and ‘Triplet 2’ (top right panel). Thick and light shaded regions give  $1\sigma$  and  $2\sigma$  range seen from the simulated spectra. In the case of ‘Triplet 1’, the largest measured gap has a size of  $17 \text{ \AA}$  ( $14.2 h^{-1}$  cMpc) around the observed wavelength of  $3576 \text{ \AA}$ . In our simulations we find 14.2 per cent of the total triplet sightlines have concurrent gap length  $\geq 14.2 h^{-1}$  cMpc. In this case we have masked an intervening C IV absorption present along the line of sight to J2117–0238 that also falls in the gap. For the individual sightlines, the underdense region corresponds to a gap of length  $17.3 \text{ \AA}$  ( $14.5 h^{-1}$  cMpc) in J2117–0238,  $22.8 \text{ \AA}$  ( $19 h^{-1}$  cMpc) in J2117–0239, and  $29 \text{ \AA}$  ( $24.2 h^{-1}$  cMpc) in J2117–0240. A spectra of a simulated triplet that closely resembles that of ‘Triplet 1’ is also shown in the bottom panel in Fig. 7. In passing we note that even when we consider the full sightline (not the small regions as we use for comparison with simulations), we find the same gap as the largest concurrent gap.

Rollinde et al. (2003) have found a concurrent gap of  $13 \text{ \AA}$  wide at  $z = 1.99$ , between four nearby sightlines that span an angular separations between 2.1 and 9.0 arc min. Cappetta et al. (2010) have reported two concurrent gaps along two triple QSO sightlines with the gap length of 10 and  $15 \text{ \AA}$  with redshifts 1.82 and 1.93, respectively. Thus we report the concurrent gap that is larger than those reported in the literature using more than three sightlines. However, the transverse spatial scales probed by ‘Triplet 1’ is smaller than those probed in the literature.



**Figure 7.** *Top:* Probability distribution of the concurrent gap length in ‘Triplet 1’ (left) and ‘Triplet 2’ (right). *Bottom-left:* Largest concurrent gap identified along ‘Triplet 1’. *Bottom-right:* An example from our simulated spectra that resembles closest to the observed largest concurrent gap.

In case of Triplet 2, such large concurrent gap length is not found (see top right panel in Fig. 7). The largest concurrent gap spreads over  $8.7 \text{\AA}$  corresponding to a length of  $5.9 h^{-1} \text{cMpc}$ . The overall distribution of the gap is consistent with the predictions of our simulations.

## 5.2 Cloud based statistics

In this section, we perform correlation analysis between Voigt profile components obtained along each line of sight using our automated Voigt profile fitting code VIPER (Gaikwad et al. 2017b). Unlike flux based statistics, in this case we will be able to probe the dependence of correlations on the  $N_{\text{HI}}$  (which is known to trace the underlying overdensity). Moreover, there is an inherent degeneracy in flux based three-point correlation. Regions having an absorption in one of the sightlines and gap in the other two sightlines will give a negative three-point correlation in flux, which is degenerate with regions having absorption in all the three sightlines. This degeneracy also makes it difficult to interpret the reduced three-point function (i.e.  $Q$  defined below) measured using flux statistics.

### 5.2.1 Longitudinal two-point correlation

When we treat IGM being constituted by distinct clouds, we follow the standard procedure to compute correlation function (both for the simulated as well as observed spectra) with respect to a random

distribution of clouds. The longitudinal two-point correlation is estimated using *Landy–Szalay* estimator (Landy & Szalay 1993),

$$\xi(\Delta r_{\parallel}) = \frac{DD - 2DR + RR}{RR}. \quad (9)$$

Here, ‘DD’, ‘RR’, and ‘DR’ corresponds to data–data, random–random, and data–random pair counts, respectively, measured at a separation of  $\Delta r_{\parallel}$ . For a given sightline, we construct the random distribution of clouds using Poisson distribution with a mean number of clouds being equal to what is expected based on the observed redshift distribution of clouds having  $N_{\text{HI}}$  above a threshold value. Once we have the observed and random cloud distributions, we consider all possible combinations of pair separations between these clouds. The *DD* pair separation counts are normalized with the total number of pair combinations, (i.e. it is divided with  $n_D(n_D - 1)$ , where  $n_D$  is the number of clouds along a sightline). In a similar fashion, one can normalize the data–random pair separations *DR* and random–random pair separation *RR*. We have used 100 random skewers for every data skewer.

The measured  $\xi(\Delta r_{\parallel})$  along our sightlines are compared with the expectations from the simulations in the last column of Fig. 3. It is clear from this plot that the expected correlation is negative in the first bin corresponding to transverse separations  $r_{\perp} < 2 h^{-1} \text{cMpc}$ . This is mainly because while there is no restriction on the minimum separations between the randomly generated clouds, thermal broadening together with instrumental resolution set a

**Table 3.** Results of cloud based ( $\log(N_{\text{HI}}) > 13$ ) correlation analysis.

Sample	Correlation	$r$ ( $\text{h}^{-1}$ cMpc)	Observed values	Probability	Percentile
Triplet 1	$\xi_{01}$	1.28	0.207	0.55	15.5
	$\xi_{12}$	1.04	0.184	0.20	8.3
	$\xi_{02}$	2.28	0.137	0.33	21.8
	$\zeta$		-0.052	0.39	24.1
	Q		-0.122	0.83	26.6
Triplet 2	$\xi_{01}$	1.55	0.143	0.27	81.7
	$\xi_{12}$	2.82	0.131	0.21	87.4
	$\xi_{02}$	4.00	0.018	0.26	42.7
	$\zeta$		-0.052	0.32	17.1
	Q		-2.246	0.68	31.7
Doublet	$\xi_{01}$	3.30	0.172	0.40	51.0

limit on the lowest measurable separation during the Voigt profile decomposition (i.e. cloud exclusion). It is clear from the figure that expected transverse two-point correlation function of clouds is consistent within  $1\sigma$  range of our model predictions. Thus we do not encounter any abnormal line of sight clustering (with scales larger than  $1 \text{ h}^{-1}$  cMpc) along the sightlines considered in this study.

### 5.2.2 Transverse two-point correlation

To probe the transverse two-point correlation between two data skewers  $D1$  and  $D2$  along two closely spaced sightlines, we generate two random skewers  $R1$  and  $R2$ . The transverse two-point correlation is then defined as

$$\xi(\Delta r_{\parallel}, \Delta r_{\perp}) = \frac{D_1 D_2 - D_1 R_2 - R_1 D_2 + R_1 R_2}{R_1 R_2}. \quad (10)$$

Similar to longitudinal two-point correlation, the transverse two-point correlation has been generated by averaging over 100 random skewers for every data skewers. We express the two-point correlation  $\xi(\Delta r_{\parallel}, \Delta r_{\perp})$  as a function of actual cloud separation  $\Delta r = \sqrt{\Delta r_{\parallel}^2 + \Delta r_{\perp}^2}$  in different plots. In Table 3, we summarize our measurement of transverse two-point correlation, of clouds having  $\log N_{\text{HI}} \geq 13.0$ , measured within the longitudinal separation bin of  $\pm 2 \text{ h}^{-1}$  cMpc (The effect of choosing a longitudinal bin smaller than  $\pm 2 \text{ h}^{-1}$  cMpc is shown in Fig. D1 in the Appendix available online). The entries in this table are similar to that of Table 2. In Fig. 8, we compare the predicted cumulative distributions from our simulated spectra with the observations (as in Fig. 6) for three different  $N_{\text{HI}}$  thresholds.

In the case of ‘Triplet 1’, we find that the observed correlations is slightly less (as suggested by the percentile) than the median predicted correlation for  $\log N_{\text{HI}} \geq 13.0$  (see Table 3 and Fig. 8). This trend is consistent with what we are finding above based on statistics using transmitted flux. Our simulations also predict an increase in  $\xi$  with increasing  $N_{\text{HI}}$ . While the observations confirm this trend, the dependence on  $N_{\text{HI}}$  is usually weaker than the predictions from the simulation in case of the transverse two-point correlations (see the vertical lines in Fig. 8).

In the case of ‘Triplet 2’ the observed  $\xi$  for two of the closest separations (i.e.  $\xi_{01}$  and  $\xi_{12}$ ) are found to be higher than our model predictions (for  $\log N_{\text{HI}} \geq 13$ ). This seems to be the case even when we include lower  $N_{\text{HI}}$  clouds. However, the observed distribution is close to the simulation results when we consider  $\log N_{\text{HI}} \geq 13.3$  for both the pairs. In case of the longest separated pair in ‘Triplet 2’ the measured  $\xi_{02}$  for all three column density thresholds are consistent

with the predictions of our simulations. This seems also to be the case for ‘Doublet’.

### 5.2.3 Three-point correlation in clouds

Here, we study the three-point correlation function between three data skewers  $D1$ ,  $D2$ , and  $D3$  using three random skewers  $R1$ ,  $R2$ , and  $R3$  generated by the same approach discussed above. For a fixed sightline configuration (i.e.  $\Delta r_{1\perp}$ ,  $\Delta r_{2\perp}$ ,  $\theta$ ), three-point correlation is a function of two redshift space separations along two sightlines with respect to the reference sightline, i.e.  $\zeta = \zeta(\Delta r_{1\parallel}, \Delta r_{2\parallel})$ . Hence, we generate data–data–data triplet separations in 2D logarithmic bins since there are two redshift space separation axes involved. This gives us  $D_1 D_2 D_3$ , which we normalize by dividing with  $n_D(n_D - 1)(n_D - 2)$ . Similarly, we generate the data–data–random, data–random–random, and random–random–random triplet separations. The three-point correlation is then given, following Szapudi & Szalay (1998), as

$$\zeta(\Delta r_{1\parallel}, \Delta r_{2\parallel}) = \frac{D_1 D_2 D_3 - DDR_{(123)} + DRR_{(123)} - R_1 R_2 R_3}{R_1 R_2 R_3}, \quad (11)$$

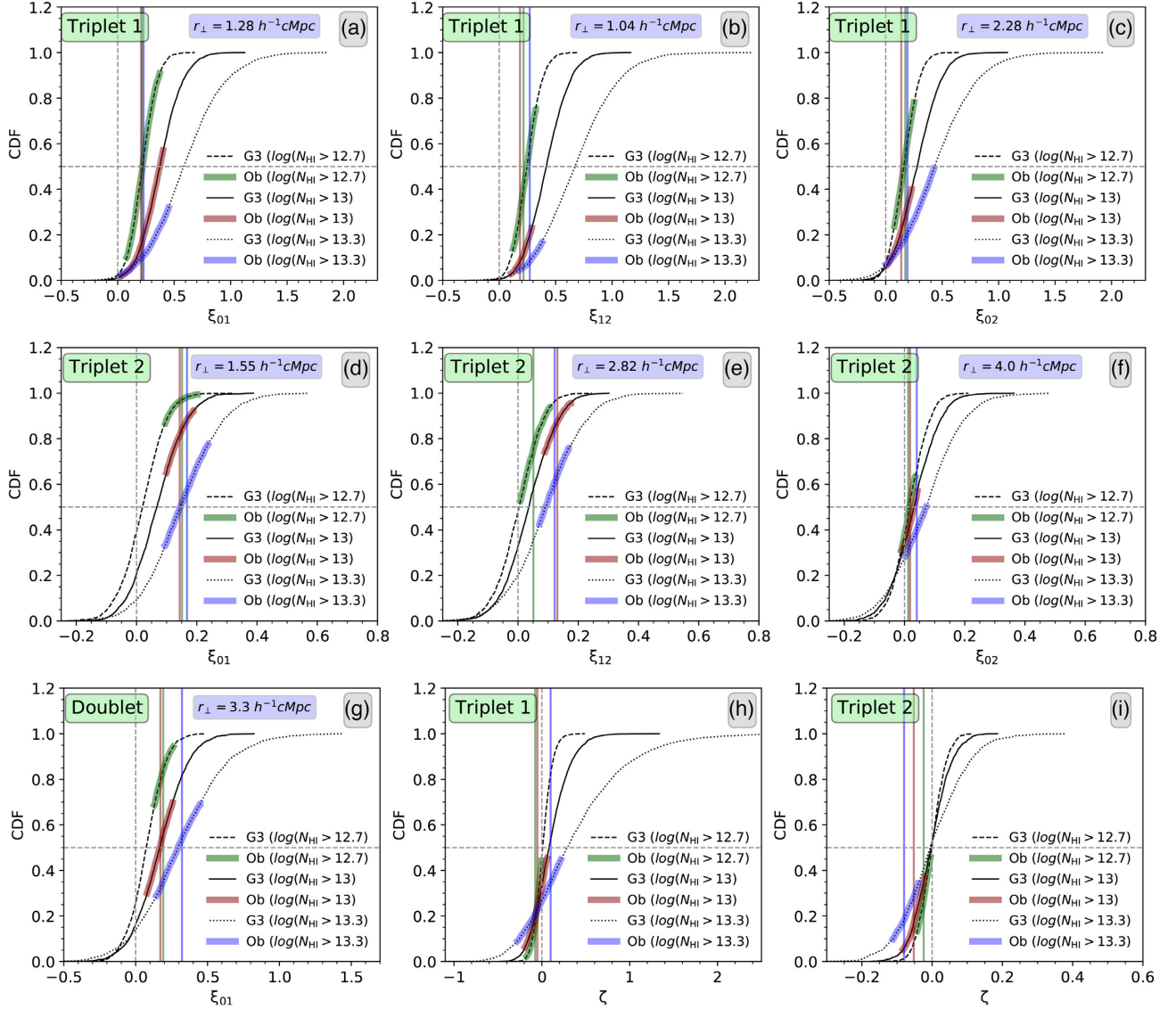
where  $DDR_{(123)} = D_1 D_2 R_3 + D_1 R_2 D_3 + R_1 D_2 D_3$  and  $DRR_{(123)} = D_1 R_2 R_3 + R_1 D_2 R_3 + R_1 R_2 D_3$ . Here also, we use 100 random skewers for every data skewers.

In Fig. 9 we plot the triple correlation measured in our simulations for configuration similar to the observed triplets among clouds along LOS 1 and LOS 2 with respect to those along LOS 0 (see Fig. 1 for reference). Strong correlations are seen up to the longitudinal separations of  $2 \text{ h}^{-1}$  cMpc along both these sightlines in ‘Triplet 1’. So while computing the three-point function, we integrate the correlations over  $2 \text{ h}^{-1}$  cMpc along each sightlines. However, in the case of ‘Triplet 2’ we do not see strong correlations. This is not surprising given the larger angular separation between the sightlines. The measured values of the three-point correlations (summed over  $\pm 2 \text{ h}^{-1}$  cMpc along each of the longitudinal directions) for both the triplets for  $\log N_{\text{HI}} \geq 13$  are given in Table 3. The cumulative distribution are shown in the bottom panels of Fig. 8.

In the case of configurations similar to ‘Triplet 1’, our simulations predict positive values of  $\zeta$  for cases  $\log(N_{\text{HI}}) \geq 13$  and 13.3. However, when we include weak  $N_{\text{HI}}$  clouds  $\zeta$  becomes close to zero. The measured values of  $\zeta$  and its  $1\sigma$  range from the observations are lower than the median values obtained in the simulations. Even though the increase in  $\zeta$  with the increase in  $N_{\text{HI}}$  is evident for the observations the actual  $\zeta$  value obtained even for the higher  $N_{\text{HI}}$  clouds are lower than those realized in the simulations. In the case of configuration similar to ‘Triplet 2’, our simulations predict median  $\zeta$  close to zero for the three  $N_{\text{HI}}$  thresholds considered. The observed values are less than the median from the simulations and mostly negative, but follow the expected trend of the cumulative distributions (see Fig. 8), that is,  $\zeta$  becomes more negative with increasing  $N_{\text{HI}}$  in the lower end of the cumulative distribution.

In the case of galaxies, it is usual procedure to define the reduced three-point function Q through the *hierarchical ansatz* suggested by Peebles (1980). In this case the observed three-point correlation function can be written in term of the cyclic combination of respective two-point correlation function as

$$\zeta(\mathbf{r}_{01\perp}, \mathbf{r}_{12\perp}, \mathbf{r}_{02\perp}) = Q [\xi(\mathbf{r}_{01\perp})\xi(\mathbf{r}_{12\perp}) + \xi(\mathbf{r}_{12\perp})\xi(\mathbf{r}_{02\perp}) + \xi(\mathbf{r}_{01\perp})\xi(\mathbf{r}_{02\perp})]. \quad (12)$$



**Figure 8.** Cloud based transverse two-point and three-point correlation cumulative distribution for the three triplets. ‘G3’ and ‘Ob’ denote GADGET-3 simulation and observation, respectively. The dashed, solid, and dotted curves are the cumulative distribution functions for different neutral hydrogen column density cut-offs. The red, green, and blue regions overlaying the simulated curve represent the  $1\sigma$  confidence interval obtained from sub-sampling of the observed triplet sightlines for the two different longitudinal grids. The vertical lines with colours denote the mean of the corresponding sub-sampled observed correlation.

For clouds with  $N_{\text{HI}}$  above a threshold, one can construct a similar reduced three-point correlation function  $Q$ . In case of galaxies where the two-point correlation is high, one expects  $Q$  to be a positive quantity (see Peebles 1980). This need not be the case for IGM, where the clustering is weak. In Fig. 10 we plot the cumulative distribution of  $Q$  for different  $N_{\text{HI}}$  thresholds. The probabilities and percentiles associated with the observed  $Q$  values are given in Table 3. Note that the observed value of  $Q$  mentioned in Table 3 is taken to be the median value out of all the 1000 randomly sub-sampled skewers. This is done since  $Q$  can diverge when two-point correlations approach zero and thereby artificially boost the mean value.

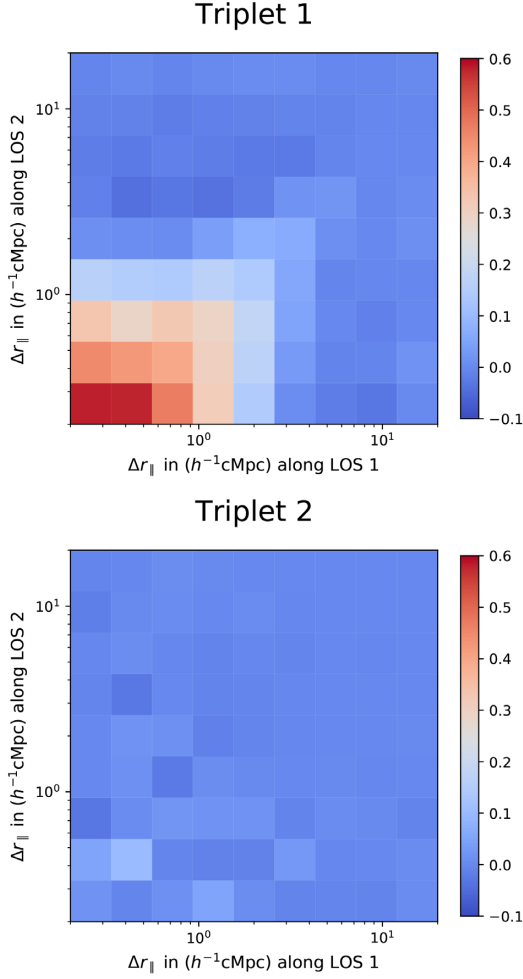
In the case of configuration similar to ‘Triplet 1’, our simulations predict median  $Q$  value to be close to zero for clouds having  $\log(N_{\text{HI}}) > 12.7$  with a large scatter. It is also evident that median  $Q$  value

increases with increasing  $N_{\text{HI}}$ . Also, cumulative distribution for high  $N_{\text{HI}}$  clouds show smaller scatter. Increasing  $Q$  values with increasing  $N_{\text{HI}}$  is clearly evident even in observations. It is also evident from the Table 3 that the observed  $Q$  distributions can be realized with high probability in our observations.

As expected, based on the  $\zeta$  values obtained above the median  $Q$ , values are zero in the case of ‘Triplet 2’ irrespective of whatever threshold column density we adopt. Also, the observed points are most of the time negative underlying the lack of strong correlation across the line of sights considered.

#### 5.2.4 Concurrent gap statistics in clouds

In the cloud treatment of IGM, the definition of gap is more straightforward. The region between two clouds in a sightline is



**Figure 9.** Cloud based (for  $\log(N_{\text{HI}}) > 13$ ) three-point correlation function from simulation as a function of redshift space separation of clouds along 2nd and 3rd sightline of the triplet from the 1st sightline. The colour of a bin represents triple correlation for that bin.

considered to be a gap. We searched for concurrent gaps as all the common regions in the three neighbouring sightlines devoid of any clouds.

Fig. 11 shows the concurrent gap distribution between clouds obtained using 4000 simulated triplet sightlines for Triplets 1 and 2. These figures also show the gap probability distribution from our simulations for three different  $N_{\text{HI}}$  threshold values. We also plot the largest concurrent gap sizes using vertical lines for different  $N_{\text{HI}}$  thresholds. As expected, the distribution shifts towards larger gap sizes with increasing  $N_{\text{HI}}$  thresholds and one picks up larger concurrent gaps more and smaller ones less. One can also see that Triplet 1 (which has smaller spatial separation in the projected space) picks up larger concurrent gaps more and smaller ones less as compared to Triplet 2. We also mention the wavelength range of the largest concurrent gaps that one gets for different  $N_{\text{HI}}$  thresholds. In Table 4, we summarize the details of the largest gaps and associated probability of occurrence. The probability given in the last of the table is the percentage of simulated triplet sightlines, which have at least one concurrent gap greater than the largest concurrent gap in the observations. In case of Triplet 1, it is that the largest gap that one picks up for  $\log(N_{\text{HI}}) > 13$  ( $11.1 \text{ h}^{-1} \text{ cMpc}$ ) is a subset of the gap that one picks up for  $\log(N_{\text{HI}}) > 13.5$  ( $25.4 \text{ h}^{-1} \text{ cMpc}$ ). For

$\log(N_{\text{HI}}) > 14$ , we pick the largest gap ( $36.4 \text{ h}^{-1} \text{ cMpc}$ ) just adjacent to the previous gap. In case of Triplet 2, the largest concurrent gap that one picks up with  $\log(N_{\text{HI}}) > 13$  ( $5.6 \text{ h}^{-1} \text{ cMpc}$ ) is a subset of the gap that one picks up with  $\log(N_{\text{HI}}) > 14$  ( $17.4 \text{ h}^{-1} \text{ cMpc}$ ). For  $\log(N_{\text{HI}}) > 13.5$ , we pick the largest gap ( $7.6 \text{ h}^{-1} \text{ cMpc}$ ) roughly  $80 \text{ \AA}$  redward of the previous gap.

### 5.3 Correlation statistics

Till now comparison of the three transverse two-point and one transverse three-point correlations for each triplet with simulations were done independent of each other. Next, we explore how closely our simulations can simultaneously reproduce all the three  $\xi_s$  and  $\zeta$  of a triplet. In this exercise, instead of assigning average statistics for the entire observed triplet sightlines, we assign statistics to definite sections. These sections are obtained by uniformly selecting 10 regions from the observed sightline with redshift path-length similar to the simulation. For each section, we measure all the four correlation parameters ( $\zeta$ ,  $\xi_{01}$ ,  $\xi_{12}$ ,  $\xi_{02}$ ) which we denote by  $x^i$ , where  $i = 0, 1, 2$ , and 3.

We identify sightlines in the simulation which are closest to the observed sightlines in terms of these correlation parameters. For this, we define a *closeness factor*  $C_4$  of the simulated sightlines with the observed sightlines using all the four correlation parameters as,

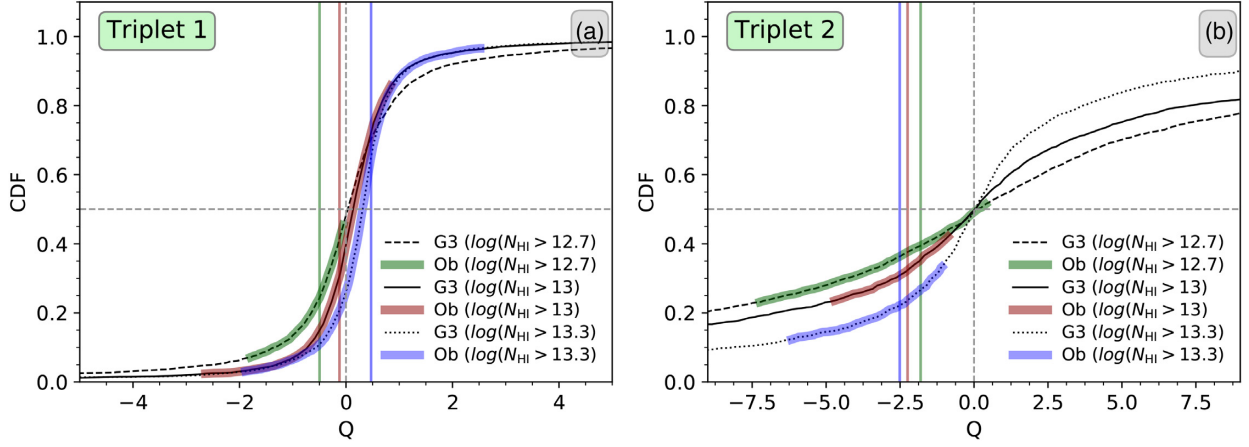
$$C_4 = \left( \sum_{i=0}^{i=3} \frac{(x^i - x_{\text{obs}}^i)^2}{4\sigma_i^2} \right)^{\frac{1}{2}}, \quad (13)$$

where  $\sigma_i$  is the standard deviation of  $i$ th parameter in the simulation. The simulated triplet which gives the smallest  $C_4$  with respect to the observations can be identified. We also define *closeness factor*  $C_3$  based only on the three transverse two-point correlation functions, i.e.

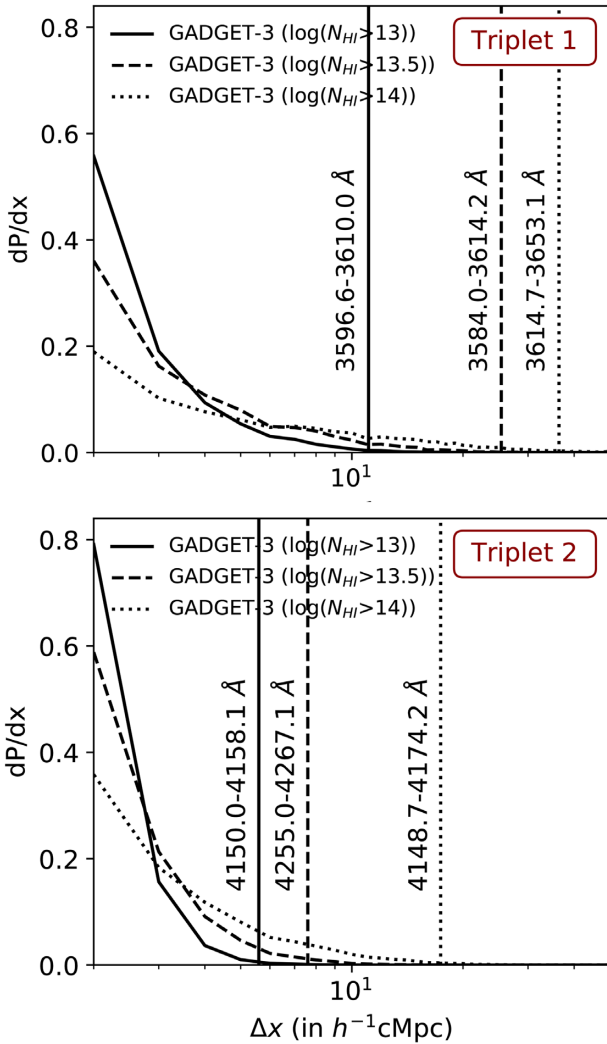
$$C_3 = \left( \sum_{i=1}^{i=3} \frac{(x^i - x_{\text{obs}}^i)^2}{3\sigma_i^2} \right)^{\frac{1}{2}}. \quad (14)$$

For each ‘Triplet’, we identify the sections (based on flux- and cloud based statistics) with highest and lowest three-point function in our observations (given by id O-Max or O-Min in column 1 of Table 5). The wavelength range for these observed sections are given in the second column of Table 5. In columns 3–6 of Table 5, we provide the transverse three- and two-point correlation values. Next, we identify simulated sightlines with smallest  $C_3$  and  $C_4$  values (column 7 in Table 5) separately. These simulated sightlines are identified as Triplet-x-S3 and Triplet-x-S4, respectively, in the table. We also estimate the percentage of simulated sightlines having statistics similar to (i.e. within  $1\sigma_i$  for all statistics considered) the observations in the chosen section. These are provided for two cases: one considering only two-point statistics and the other considering both two- and three-point statistics (columns 9 and 10 in Table 5).

First, we consider the flux based statistics for the segment with minimum  $\zeta$  value (i.e. having strong transverse three-point correlation). It is evident that  $C_3$  and  $C_4$  based selection picked completely different simulated sightlines. We find that about 22.6 per cent of 4000 simulated triplet sightlines produce all the three  $\xi$  values that match with the observed values within their corresponding  $1\sigma$  levels. When we demand similar matching also for  $\zeta$ , the percentage falls to 12.3 per cent. Even when we consider the segment with maximum  $\zeta$  (or weak transverse three-point correlation), the simulated sightline picked by  $C_3$  and  $C_4$  are not the same. However, the probability of sightlines producing the  $\xi$  and  $\zeta$



**Figure 10.** Reduced three-point correlation ( $Q$ ) for the two triplets. The dashed, solid, and dotted black lines represent the cumulative distribution function for different  $N_{\text{HI}}$  column density cut-offs in simulations. The red, green, and blue regions overlaying the simulated curve represent the  $1\sigma$  confidence interval obtained from sub-sampling of the observed triplet sightlines. The vertical lines with colours denote the median observed  $Q$  value.



**Figure 11.** Concurrent gap distribution between clouds for different  $N_{\text{HI}}$  thresholds. The vertical lines with linestyle same as the gap distribution curves represent the largest concurrent gap (its wavelength range is provided) in the observed triplet for that corresponding  $N_{\text{HI}}$ .

**Table 4.** Results of concurrent gap statistics in cloud.

Sample	$\log(N_{\text{HI}})$ threshold	Largest gap ( $h^{-1}$ cMpc)	Wavelength range ( $\text{\AA}$ )	% of simulated triplets
Triplet 1	13	11.1	3596.6–3610.0	20.7
	13.5	25.4	3584.0–3614.2	5.8
	14	36.4	3614.7–3653.1	20.6
Triplet 2	13	5.6	4150.0–4158.1	9.1
	13.5	7.6	4255.0–4267.1	41.1
	14	17.4	4148.7–4174.2	25.0

close to the observed values are higher than what we found for the segment with strong  $\zeta$ .

In the case of ‘Triplet 2’ the segment with the strongest  $\zeta$  is between 4167.5 and 4318.9  $\text{\AA}$  (see Fig. C2 in the Appendix, which is available online). It is interesting to note from the table that  $\xi_{02}$  (measured between the largest angular separation in the configuration) is larger than the other two  $\xi$ s. The simulated triplet sightline picked based on  $C_4$  is also picked by  $C_3$ . It is clear in our simulations that this configuration is very rare (i.e. only 2 per cent of the simulations produce all the correlations within their corresponding  $1\sigma$ ). As before, when we consider the maximum  $\zeta$  (i.e. less correlated) segment  $C_3$  and  $C_4$  pick different simulated sightlines. We also find the probability of producing the high- $\zeta$  sightlines are higher than low- $\zeta$  sightlines.

Next, we repeat the analysis for the cloud based (for  $\log(N_{\text{HI}}) \geq 13$ ) statistics. In the case of ‘Triplet 1’ same segment has strongest transverse three-point correlation function, whether we use flux or cloud based statistics. However, unlike flux based statistics the probability of realizing all four correlation function values within  $1\sigma$  of the observed values in our simulation turns out to be very low (i.e.  $\sim 3$  per cent). The segment with lowest transverse three-point correlation function values is slightly shifted with respect to what we found based on flux statistics. However, there is a considerable overlap between the two segments. Here also we find that  $C_3$  and  $C_4$  pick different simulated sightlines. Moreover, the probabilities are less than what we find when we use flux statistics.

In the case of ‘Triplet 2’, we do not obtain similar segments between flux and cloud approach based on our criteria of strongest transverse three-point correlation. This is interesting since the segment which has the highest negative transverse three-point

**Table 5.** Correlation statistics.

Triplet Id	Wavelength (Å)	$\zeta$	$\xi_{01}$	$\xi_{12}$	$\xi_{02}$	$C_3$	$C_4$	Per cent with	
								$1\sigma_{i=1,2,3}$	$1\sigma_{i=0,1,2,3}$
<b>Flux based</b>									
Triplet1-O-Min	3778.2–3900.2	−0.0018	0.013	0.006	0.004	–	–	22.6	12.375
Triplet1-S4	–	−0.0017	0.014	0.004	0.003	0.18	<b>0.16</b>	–	–
Triplet1-S3	–	0.0005	0.013	0.006	0.004	<b>0.09</b>	0.71	–	–
Triplet1-O-Max	3539.5–3661.5	0.0007	0.009	0.005	0.004	–	–	37.5	33.5
Triplet1-S4	–	0.0006	0.009	0.005	0.003	0.13	<b>0.11</b>	–	–
Triplet1-S3	–	−0.0003	0.009	0.004	0.005	<b>0.03</b>	0.31	–	–
Triplet2-O-Min	4167.5–4318.9	−0.0014	0.003	0.003	0.019	–	–	2.0	1.9
Triplet2-S4-S3	–	−0.0017	0.003	0.007	0.019	<b>0.30</b>	<b>0.27</b>	–	–
Triplet2-O-Max	4055.0–4206.4	0.0012	0.009	0.013	0.005	–	–	31.6	24.2
Triplet2-S4	–	0.0009	0.008	0.013	0.004	0.07	<b>0.09</b>	–	–
Triplet2-S3	–	−0.0012	0.010	0.013	0.005	<b>0.06</b>	0.52	–	–
<b>Cloud based</b>									
Triplet1-O-Max	3778.2–3900.2	0.05	0.11	0.10	−0.04	–	–	3.4	2.825
Triplet1-S4-S3	–	0.05	0.17	0.11	−0.05	<b>0.19</b>	<b>0.17</b>	–	–
Triplet1-O-Min	3500–3622	−0.24	0.30	0.37	0.29	–	–	33.2	9.825
Triplet1-S4	–	−0.21	0.31	0.39	0.27	0.12	<b>0.13</b>	–	–
Triplet1-S3	–	0.05	0.29	0.37	0.30	<b>0.05</b>	0.70	–	–
Triplet2-O-Max	4069–4220.4	−0.01	0.12	0.15	0.00	–	–	12.3	8.5
Triplet2-S4-S3	–	−0.01	0.12	0.16	−0.01	<b>0.05</b>	<b>0.12</b>	–	–
Triplet2-O-Min	4139.4–4290.8	−0.08	0.15	0.14	0.05	–	–	14.0	4.15
Triplet2-S4	–	−0.07	0.16	0.15	0.07	0.20	<b>0.22</b>	–	–
Triplet2-S3	–	0.04	0.14	0.13	0.05	<b>0.08</b>	1.23	–	–

correlation in flux (4167.5–4318 Å) is supposed to be strongly correlated. But, when we go to cloud based analysis, we do not pick that region to have the strongest transverse three-point correlation. Rather, in cloud based analysis, this particular region has a transverse three-point correlation of  $-0.07$ . This discrepancy is probably seen because of a strong absorption present in one of the sightlines ( $\lambda \sim 4290\text{Å}$ ) of ‘Triplet 2’. In case of flux, presence of a strong absorption in one of the sightlines will make the transverse three-point correlation largely negative when it correlates with either gaps or absorption in both the adjacent sightlines. This is not true for clouds as it gives positive transverse three-point correlation only with coherent absorption in all the three sightlines.

## 6 CORRELATIONS BETWEEN $\text{Ly}\alpha$ , C IV, DLAS, AND QUASARS

### 6.1 Quasar– $\text{Ly}\alpha$ absorption transverse correlations

In this section, we consider the correlation between the foreground quasars and  $\text{Ly}\alpha$  forest along the LOS to the background quasars. In the framework of standard proximity effect, one expects radiation induced voids in the transverse directions as  $r_{eq}$  for most of our quasars are larger than the transverse separations probed (see Table 1). However, studies of quasar pairs have revealed excess  $\text{Ly}\alpha$  absorption in the transverse direction (Guimarães et al. 2007; Kirkman & Tytler 2008; Prochaska et al. 2013; Rollinde et al. 2013; Lau, Prochaska & Hennawi 2016; Jalan, Chand & Srikanand 2018). Triplets allow us to probe the gas around QSOs through more than one sightlines.

When we consider spectra of all the three triplets we have (including the BOSS spectrum of J141831.7+065711.2), we cover a redshift path-length  $\Delta z$  of  $\sim 6.9$  for detecting damped  $\text{Ly}\alpha$  systems or DLAs ( $N_{\text{HI}} > 2 \times 10^{20} \text{cm}^{-2}$ ). From the DLA frequency distribution (i.e.  $dN/dz$ ) given in the table 2 of Noterdaeme et al.

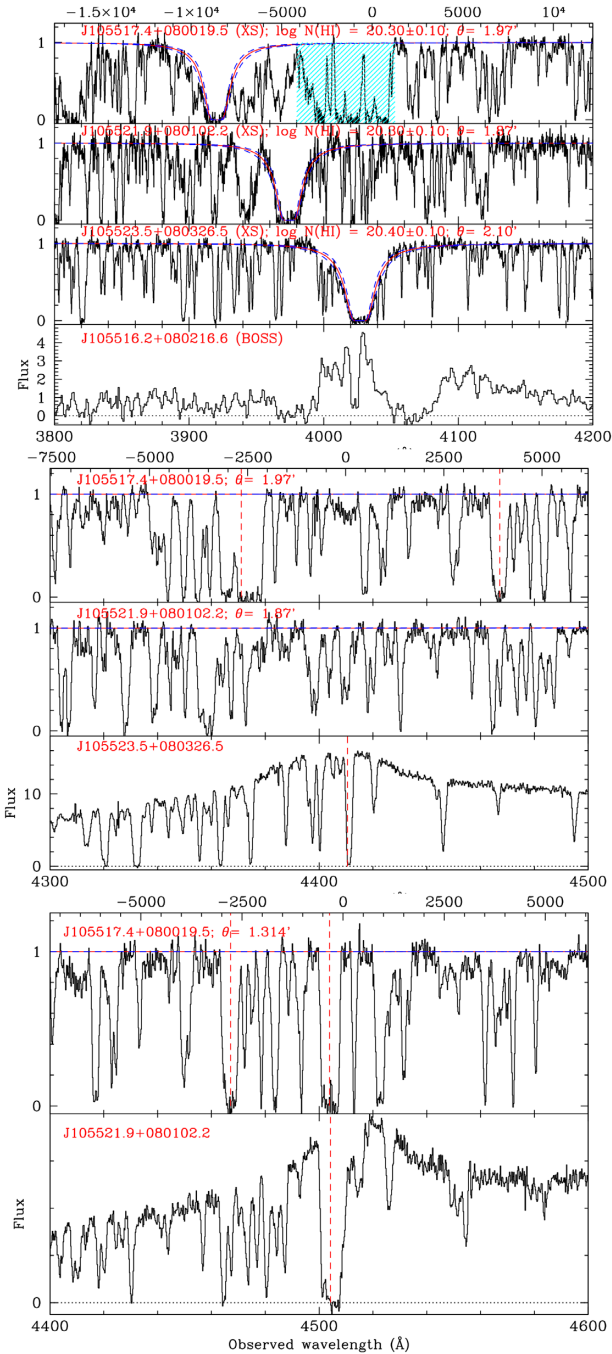
(2012), we expect 1.3 DLAs while we detect 3 DLAs in Triplet 2. Along each sightline in ‘Triplet 2’, we cover a redshift path-length of  $\Delta z = 1$ , where the expected number of DLAs is 0.2. Detection of a DLA along each sightline correspond to a factor 5 excess compared to the expected number of randomly distributed DLAs. Interestingly, the redshifts of the DLAs at the maximum differ by  $\Delta z = 0.088$  (or  $\Delta v = 6000 \text{km s}^{-1}$ ; See Fig. 12). Given the presence of a DLA the expected number of DLAs in the other two sightline having a redshift difference within 0.088 is  $3 \times 10^{-4}$  [i.e.  $(0.088 \times 0.2)^2$ ]. This implies that the chance coincidence of three DLAs with close redshift separation is less probable.

What is making this case more interesting is the fact that the fourth quasar (i.e. J105516.2+080216.6 with  $z_{em} = 2.320$ ) we identified in this field roughly has equal angular separation with respect to the quasars in ‘Triplet 2’. The  $\text{Ly}\alpha$  emission from this quasar coincides with the DLA seen along the line of sight to J1055+0803 (see Fig. 12). This quasar is really faint and its sphere of influence for HI ionization (see last column of Table 1) is much smaller than separation of the other three sightlines from the quasar. Note while the presence of quasar and DLAs within a small redshift range and angular separation is interesting, the redshift space separation of  $\sim 37 \text{pMpc}$  may seem too large. It will be interesting to perform deep imaging in this regions to search for a possible presence of large-scale density enhancements.

Based on the  $r_{eq}$  provided in Table 1, we identify the wavelength range in the transverse direction that will be affected by isotropic ionizing radiation from the foreground quasars. For each quasar pairs (listed in first two columns) these are listed in the third column of the Table 6. The number of clouds having  $N_{\text{HI}}$  above a given threshold found within these wavelength range are given in last column of this table. In this column we also provide average number of clouds and associated errors found for this similar wavelength range in our spectrum far away from quasars.

Two background quasar sightlines [J1055+0800 (with a separation of 1.6 pMpc) and J1055+0801 (with a separation of 1.2 pMpc)]





**Figure 12.** *Top:* Ly $\alpha$  forest towards quasars constituting ‘Triplet 2’ around the foreground quasar J105516.2+080216.6. DLAs are detected along all the three sightlines with redshifts close to the redshift of the foreground quasar. The fit to the DLAs are shown, the measured HI column densities and the angular separation of the sightline with respect to the foreground quasars are given in each panel. Shaded region is the contamination by associated O VI absorption. *Middle:* Ly $\alpha$  forest towards J105517.4+080019.5 and J105521.9+080102.2 around the foreground quasar J105523.5+080326.5. *Lower:* Ly $\alpha$  forest J105517.4+080019.5 around the foreground quasar J105521.9+080102.2. In all cases the vertical dotted line gives the location Ly $\alpha$  lines having associated the C IV absorption.

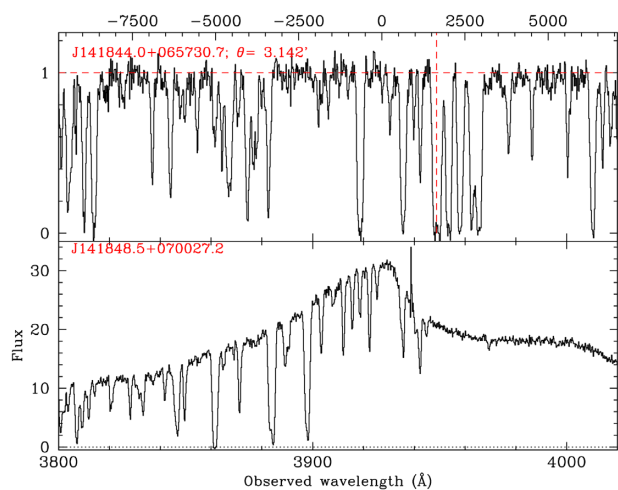
**Table 6.** Quasar–Ly $\alpha$  transverse correlation.

Quasar LOS	Ly $\alpha$ LOS	Proximity region( $\text{\AA}$ )	No. of absorbers
<b><math>\log(N_{\text{HI}}) &gt; 13</math></b>			
J1055+0803	J1055+0801	$4490.2 \pm 11.3$	$7(3.7 \pm 1.9)$
J1055+0803	J1055+0800	$4490.2 \pm 10.4$	$8(4.5 \pm 1.8)$
J1055+0801	J1055+0800	$4508.9 \pm 10.9$	$4(4.8 \pm 1.9)$
J1418+0700	J1418+0657b	$3927.2 \pm 4.7$	$2(1.4 \pm 1.1)$
<b><math>\log(N_{\text{HI}}) &gt; 14</math></b>			
J1055+0803	J1055+0801	$4490.2 \pm 11.3$	$3(1.8 \pm 1.3)$
J1055+0803	J1055+0800	$4490.2 \pm 10.4$	$3(1.1 \pm 1.1)$
J1055+0801	J1055+0800	$4508.9 \pm 10.9$	$4(1.2 \pm 1.2)$
J1418+0700	J1418+0657b	$3927.2 \pm 4.7$	$0(0.8 \pm 1.0)$

probe the proximity region of the foreground quasar J1055+0803. The Ly $\alpha$  absorption along these sightlines are shown in the middle panel of Fig. 12. The vertical dashed lines mark the locations of Ly $\alpha$  clouds with associated C IV absorption. In the case of J1055+0803, we find a factor 1.8 times more absorption along J1055+0800 and J1055+0801 when we consider clouds having  $\log(N_{\text{HI}}) > 13$ . While based on standard proximity effect we would have expected deficit of absorption. When we consider  $\log(N_{\text{HI}}) > 14$ , the observed excess is 1.7 and 2.7 times the expected value towards J1055+0801 and J1055+0800, respectively. Given the small wavelength range probed these excess are not statistically significant.

In the bottom panel of Fig. 12, we show the Ly $\alpha$  absorption towards J1055+0800 in the proximity of the foreground quasar J1055+0801 (at a separation of 640 pkpc). It is clear from the figure that the foreground quasar has a strong associated absorption at  $z_{\text{abs}} = 2.7051$ . The associated C IV absorption shows three components with absorption redshifts of 2.7030, 2.7043, and 2.7052. While we detect Ly $\beta$  and Si IV absorption, neither low ions like Si II nor high ions like O VI or NV are clearly seen in absorption at the redshifts of the C IV components. We see Ly $\alpha$  absorption and associated C IV at the same redshift along the line of sight to J1055+0800. While Ly $\alpha$  absorption profile matching is very good, the C IV absorption shows three components (at  $z_{\text{abs}} = 2.7051$ , 2.7064, and 2.7075) that are slightly redshifted with respect to absorption seen towards J1055+0801. Interestingly, we detect C II, Si II, Si III, and Si IV absorption in addition to possible O VI. These are the closest sightlines (with a physical separation of 635 pkpc) in ‘Triplet 2’. From Table 6 we notice that within the proximity region considering the redshift of the quasar and not the above discussed associated C IV system, we find Ly $\alpha$  absorption for  $\log N_{\text{HI}} > 13$  to be as per expectation and an excess of 3.3 times the expected absorption for  $\log N_{\text{HI}} > 14$ .

In Fig. 13, we plot the Ly $\alpha$  forest region along the line of sight to J141844.0+065730.7 around the foreground quasar J141848.5+070027.2. The statistics of the Ly $\alpha$  absorption within the expected proximity along the foreground quasar are summarized in Table 6. While there is excess Ly $\alpha$  suggested from this table, it is evident from the figure that there is a strong clustered Ly $\alpha$  absorption within the velocity range  $-1000$  to  $+3000$  km s $^{-1}$ . One of this components at  $z = 2.2483$  (identified with red dashed line) shows C IV and Mg II absorption. While we do not have X-Shooter spectrum of the third quasar J141831.7+065711.2 in ‘Field 3’, the available SDSS spectrum shows strong Ly $\alpha$  absorption in the same identified velocity range. Thus there appears to be an overdense region at slightly higher redshift around this quasar.



**Figure 13.** Ly $\alpha$  forest towards quasar J141844.0+065730.7 around the foreground quasar J141848.5+070027.2.

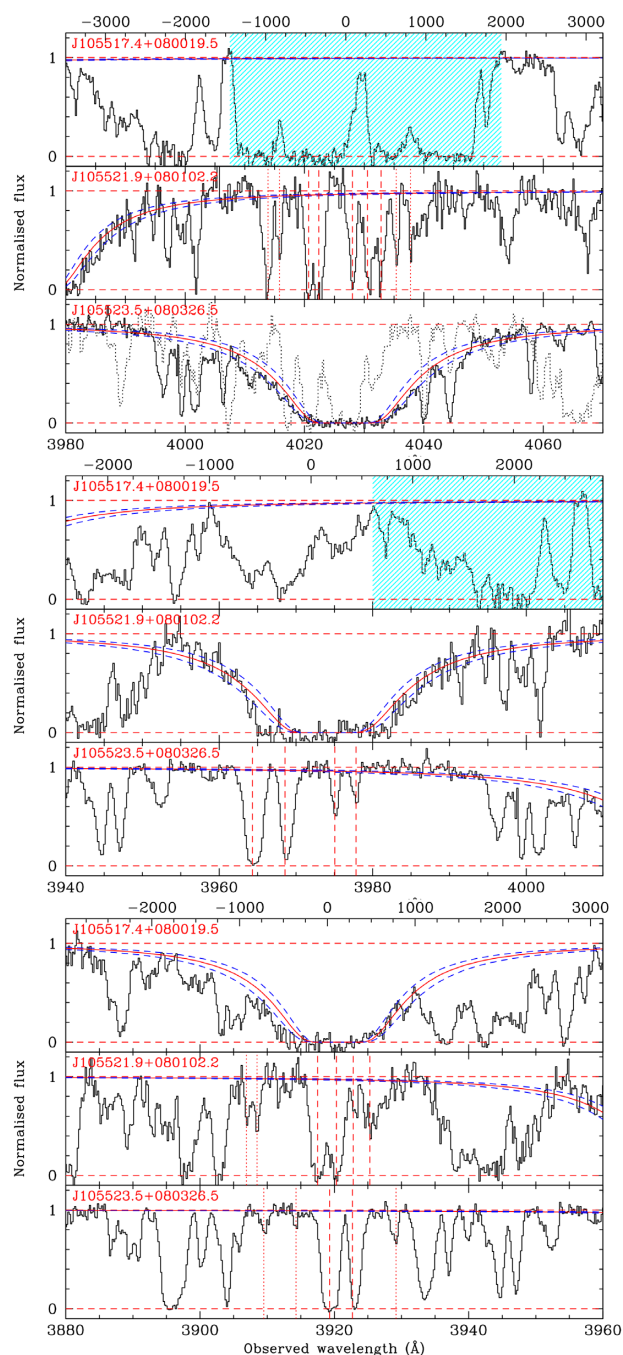
In summary, in all the triplet sightlines considered here we do find signature of excess Ly $\alpha$  absorption around quasars. Previous studies have shown excess absorption at the redshift of the foreground quasars in the spectrum of background quasars. Here we are able to probe the quasar environment using more than one quasar sightlines. Such studies using large number of triplets (or multiple sightlines) will allow us to probe either the geometry of the gas distribution or the nature of the ionizing radiation (isotropic or an-isotropic). In principle one will be able to constrain the time-scale related to the quasar activities much better than what one could do with doublets.

## 6.2 DLA–Ly $\alpha$ absorption transverse correlations

Rubin et al. (2015) have studied the gas distribution around high- $z$  DLAs using absorption lines detected in QSO sightlines within a transverse separations of 300 kpc to 40  $z \sim 2$  DLAs. They found optically thick H I absorption (i.e.  $\log N_{\text{HI}} \geq 17.3$ ) up to 200 kpc with a covering fraction of  $\geq 30$  per cent. Low ionization metals traced by Si II  $\lambda$  1526 are found to have a covering fraction of 20 per cent within 100 kpc, while high ions traced by C IV absorption seem to have higher covering factor ( $\sim 57$  per cent) and kinematically coupled to the DLA to a larger transverse separations ( $\sim 200$  kpc) compared to the low ions.

In this section we study the DLA–Ly $\alpha$  transverse correlation considering the three DLAs detected along the lines of sight towards quasars in ‘Triplet 2’. Note the separations we probe are at least a factor two higher than that of Rubin et al. (2015); however, we have a unique opportunity to probe the gas distribution around 3 DLAs simultaneously using two sightlines.

In the bottom panel of Fig. 14, we plot the Ly $\alpha$  absorption towards J105521.9+080102.2 (at a projected separation of  $\sim 640$  pkpc) and J105523.5+080326.5 (at a projected separation of 1.6 pMpc) at the redshift of the DLA detected along the line of sight to J105517.4+080019.5. We fit the Ly $\alpha$  lines within  $\pm 1000$  km s $^{-1}$  to the DLA in the redshift space with multiple component Voigt profiles. Along J105521.9+080102.2, we find one component having  $10^{16} \leq N_{\text{HI}} [\text{cm}^{-2}] \leq 10^{17}$  within  $\pm 600$  km s $^{-1}$  to the DLA. In the case of J105523.5+080326.5 we have three such components detected within  $\pm 1000$  km s $^{-1}$ . We do not detect Si II  $\lambda$  1526 or C IV associated with these absorbers or any other Ly $\alpha$  absorption within  $\pm 1000$  km s $^{-1}$  to the DLA.



**Figure 14.** Transverse correlations between DLAs and Ly $\alpha$  absorption in ‘Triplet 2’. Vertical dashed and dotted lines identify Ly $\alpha$  absorption within a velocity separations of 600 km s $^{-1}$  and between 600 and 1000 km s $^{-1}$ , respectively. The coincident absorption from other absorption systems are masked with shaded regions.

In Table 7 we present our results for number of clouds observed and predicted by the random distribution for two different velocity intervals centred around the DLAs. First two columns in this table identify the DLA and Ly $\alpha$  sightlines. Third and fourth columns give the number of cloud detected within  $\pm 600$  km s $^{-1}$  and  $\pm 1000$  km s $^{-1}$ , respectively. The numbers in the bracket are the expected number of Ly $\alpha$  absorption within the specified velocity range if we go to a random place in the observed Ly $\alpha$  forest. It is clear that the observed numbers are within  $1.5\sigma$  to the expected

**Table 7.** DLA–Ly $\alpha$  transverse correlation.

DLA LOS	Ly $\alpha$ LOS	Number of absorbers	
		$ v  < 600 \text{ km s}^{-1}$	$ v  < 1000 \text{ km s}^{-1}$
<b><math>\log(N_{\text{HI}}) &gt; 13</math></b>			
J1055+0800	J1055+0801	4(2.7 $\pm$ 1.6)	7(6.0 $\pm$ 1.9)
J1055+0800	J1055+0803	4(3.7 $\pm$ 1.8)	7(4.5 $\pm$ 2.2)
J1055+0801	J1055+0803	4(3.7 $\pm$ 1.8)	6(4.5 $\pm$ 2.2)
J1055+0803	J1055+0801	3(2.7 $\pm$ 1.6)	8(6.0 $\pm$ 1.9)
<b><math>\log(N_{\text{HI}}) &gt; 14</math></b>			
J1055+0800	J1055+0801	2(1.3 $\pm$ 1.1)	3(2.1 $\pm$ 1.4)
J1055+0800	J1055+0803	3(0.9 $\pm$ 1.1)	3(1.5 $\pm$ 1.3)
J1055+0801	J1055+0803	2(0.9 $\pm$ 1.1)	3(1.5 $\pm$ 1.3)
J1055+0803	J1055+0801	3(1.3 $\pm$ 1.1)	7(2.1 $\pm$ 1.4)

number. However, detecting partial Lyman limit systems (i.e.  $N_{\text{HI}} > 10^{16} \text{ cm}^{-2}$ ), along both the sightlines, is interesting. We come back to this after looking at Ly $\alpha$  absorption around other two DLAs.

In the middle panels in Fig. 14, we plot the Ly $\alpha$  absorption towards J105517.4+080019.5 (transverse separation of 640 pkpc) and J105523.5+080326.5 (transverse separation of 1.2 pMpc) at the redshift of the DLA along the lines of sight to J105521.9+080102.2. The shaded region gives the wavelength affected by Ly $\beta$  of the strong associated absorption seen along the line of sight to J105517.4+080019.5. Within  $\pm 1000 \text{ km s}^{-1}$  we notice broad absorption with three visible components. We tentatively identify this to O VI  $\lambda$  1037 associated to a broad feature we identify based on C IV absorption at  $z_{\text{abs}} = 2.8225$ . We call this tentative as profiles of O VI doublets are affected by narrow intervening Ly $\alpha$  absorption. While we see the two profiles do not match perfectly, they do have consistent optical depths as expected. This means that we will not be able to get a handle on number of Ly $\alpha$  absorptions within our velocity range of interest. However, we can conclude that there is no cloud with  $N_{\text{HI}} > 10^{16} \text{ cm}^{-2}$ . Along the line of sight to J105523.5+080326.5, we detect 6 Ly $\alpha$  components within  $\pm 1000 \text{ km s}^{-1}$  with  $N_{\text{HI}} > 10^{13} \text{ cm}^{-2}$  (there is one component with  $N_{\text{HI}} = 10^{16.9} \text{ cm}^{-2}$  within  $\pm 1000 \text{ km s}^{-1}$ , but none with  $N_{\text{HI}} > 10^{16} \text{ cm}^{-2}$  within  $\pm 600 \text{ km s}^{-1}$ ). However, as can be seen from the Table 7 the observed number are consistent with the expectations.

In the top panel in Fig. 14 we plot the Ly $\alpha$  absorption towards J105517.4+080019.5 (transverse separation of 1.6 pMpc) and J105521.9+080102.2 (transverse separation of 1.2 pMpc) at the redshift of the DLA detected along the line of sight to J105523.5+080326.5. Unfortunately strong associated O VI absorption (see the cyan shaded region) completely blanket the expected Ly $\alpha$  wavelength range along the line of sight to J105517.4+080019.5. The Ly $\beta$  range is also in the low-SNR region, so we will not be able to probe the DLA–Ly $\alpha$  cross-correlation along this sightline. Along the line of sight towards J105521.9+080102.2, we find several Ly $\alpha$  absorption within a velocity range of  $\pm 1000 \text{ km s}^{-1}$ . In this case also we do not find any metal line associated to the Ly $\alpha$  absorption. The highest HI column density detected is  $N_{\text{HI}} = 10^{16.6} \text{ cm}^{-2}$ . When we consider clouds with  $\log N_{\text{HI}} \geq 13.0$  the observed number of absorption is consistent with the expected number. However, if we consider only clouds with  $\log N_{\text{HI}} \geq 14.0$  the observed number is consistent with the expected value within 1.5  $\sigma$  level for  $\delta_v < 600 \text{ km s}^{-1}$  and at  $3.5\sigma$  for  $\delta_v < 600 \text{ km s}^{-1}$ .

The detection of partial LLS close to the DLAs is intriguing. We have detected five such absorbers within  $\pm 600 \text{ km s}^{-1}$  and seven absorbers within  $\pm 1000 \text{ km s}^{-1}$  to three DLAs along five LOS

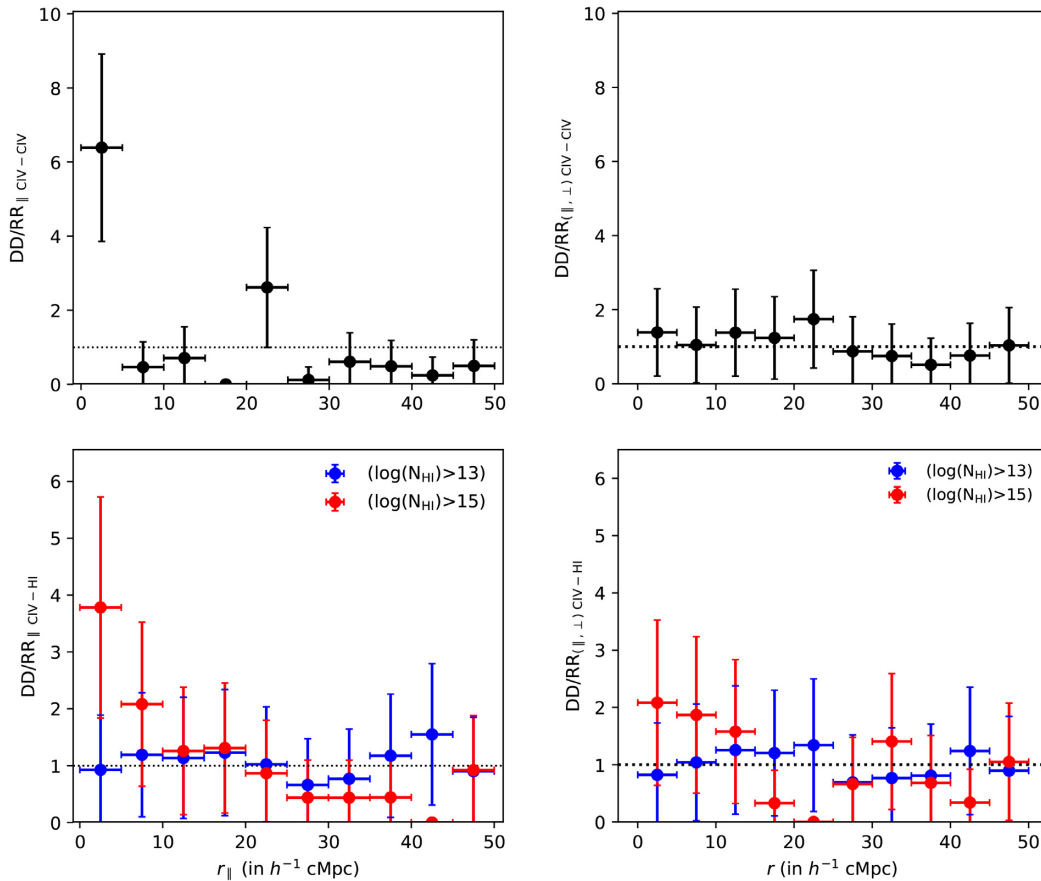
probed. Taken on the face value we detect at least one cloud with  $N_{\text{HI}} \geq 10^{16} \text{ cm}^{-2}$  in 60 per cent of DLAs within  $\pm 600 \text{ km s}^{-1}$  and in 80 per cent of DLAs within  $1000 \text{ km s}^{-1}$  within a project separations of 0.6–1.6 pMpc. Based on the redshift distribution of HI absorbers (given in Kim et al. 2013), we expect  $\sim 0.1$  such absorbers in the redshift range searched within  $\pm 600 \text{ km s}^{-1}$  and  $\sim 0.17$  within  $\pm 1000 \text{ km s}^{-1}$ . This together with the fact that the absorption occur close to a chosen DLA favours excess clustering of high-column density Ly $\alpha$  absorbers around DLAs over the projected separations 0.6–1.6 Mpc. Given the length-scale probed these absorbers may not belong to the circumgalactic medium of the quasars as studied by Rubin et al. (2015), but may reflect biased clustering of the IGM around DLAs (see, for example, Pérez-Ràfols et al. 2018, based on SDSS BOSS data).

### 6.3 C IV correlations

In this section, we discuss the clustering properties of these C IV absorbers (given in Table B1 in the Appendix available online), and how their clustering is associated with the underlying  $N_{\text{HI}}$  field. Individual C IV absorbers were fitted with multiple component Voigt profiles using VPFIT. For C IV–C IV correlations, we consider C IV clouds having observed wavelength greater than the Ly $\alpha$  emission wavelength of that sightline and having redshift less than the Ly $\alpha$  emission redshift by  $5000 \text{ km s}^{-1}$  to avoid the proximity regions of the quasar. To estimate the C IV–C IV correlations, we first calculate the data–data pair counts ‘DD’. We then generate thousand random skewers to calculate the random–random pair counts ‘RR’ keeping the number of clouds in the random skewers same as observed. We use DD/RR as the estimator for C IV–C IV correlation, with DD/RR = 1 means no correlation.

The longitudinal (transverse) correlations have been calculated by summing DD along all quasar sightlines (pairs of sightlines) and then normalizing it by the total RR. In the row 1 of Fig. 15, we have shown the C IV–C IV correlations both along the line of sight (left-hand panel) as well as between two lines of sight (right-hand panel). We have also associated a Poissonian error based on the number of the DD pair counts. For the longitudinal correlation, a sharp peak is seen in the first bin (0–5  $\text{h}^{-1}$  cMpc,  $\sim 0$ –500  $\text{km s}^{-1}$ ). This is consistent with previous works based on C IV–C IV longitudinal correlations that have suggested a correlation length-scales of about  $1000 \text{ km s}^{-1}$  (Rauch et al. 1996; Boksenberg, Sargent & Rauch 2003; Pichon et al. 2003; Scannapieco et al. 2006). Our data also shows a peak at the 20–25  $\text{h}^{-1}$  cMpc bin (2000–2500  $\text{km s}^{-1}$ ) in the case of longitudinal correlation. Upon investigation, the longitudinal clustering at this scale, though not significant beyond  $1\sigma$  level, primarily comes from the sightlines J2117–0238 and J1055+0800. In the case of transverse correlation of the C IV absorbers, we do not see significant clustering at any scale. The correlation signal is consistent with random distribution of absorbers within  $1\sigma$  of the Poissonian error. This is similar to what has been reported by Coppolani et al. (2006) excluding a clustering signal at  $\sim 20 \text{ 000 km s}^{-1}$  observed due to C IV groups in their quasar quartet. In passing we note that the number of detected C IV absorbers are too low to attempt a three-point correlation function measurement.

Next, we probe the C IV–HI clustering using C IV absorbers detected in the redshift range of the observed Ly $\alpha$  forest. The locations of the C IV absorbers are then correlated with those of the Ly $\alpha$  absorbers for two different  $N_{\text{HI}}$  cut-offs to probe its dependence on  $N_{\text{HI}}$ . In the second row of Fig. 15, the blue and red points correspond to C IV–HI cross-correlation for  $\log(N_{\text{HI}}) > 13$  and 15,



**Figure 15.** Row1: two-point correlation using Voigt profile fits of the C IV metals lines. Row2: two-point cross-correlation using Voigt profile fits of the C IV metals lines with Voigt profile fits of H I absorbers with different column density cut-offs. The x-axis is the physical separation of the clouds in comoving units.

respectively. It is seen that there is no significant clustering of low  $N_{\text{HI}}$  clouds with the C IV absorbers. Rather, higher  $N_{\text{HI}}$  Ly $\alpha$  absorbers seem to cluster around C IV absorbers up to 10  $h^{-1}$  cMpc in the longitudinal direction. Similar clustering behaviour is also seen in the transverse correlations albeit with lower amplitude and significant levels. There is no observable correlation between the lower column density H I and C IV absorption, but for higher column densities there is a slightly higher correlation (though with lower statistical significance) up to 10–15  $h^{-1}$  cMpc. Due to small number of sightlines involved the results presented in this sub-section can not be confirmed at high level of statistical significance. With the existing Echelle spectroscopic data of high- $z$  quasars, it will be possible to confirm the trend with better significant level for the longitudinal correlation. This we will do in our forthcoming paper.

## 7 SUMMARY

We present X-Shooter observations of eight quasars that are part of two projected triplets and a doublet (with pair separations spanning 0.5–1.6 pMpc). We used the absorption lines detected to study correlation properties: (a) autocorrelation of Ly $\alpha$  absorption in the longitudinal and transverse directions using flux and cloud based statistics, (b) transverse cross-correlation between Ly $\alpha$  and foreground QSOs and DLAs, and (c) autocorrelation of C IV and cross-correlation of C IV and Ly $\alpha$  along longitudinal and transverse directions. To understand the results of our spatial correlation studies of Ly $\alpha$  forest, we use hydrodynamical simulations that

are validated using observed flux probability distribution function, distribution of  $N_{\text{HI}}$ , and observed relationship between two-point flux correlation function and radial separations. Our main findings are

(i) Based on the transmitted flux, we derive two-point and three-point correlation functions for sightlines of our quasars. The observed transverse two-point correlation as a function of projected separation is found to be consistent with previous studies by Coppolani et al. (2006) and with our simulated results.

(ii) We compute the probability of realizing observed values of two-point and three-point functions and their  $1\sigma$  range in our simulations. It is found that the average clustering properties of the Ly $\alpha$  forest seen along these two triplets are reproduced with the probability ranging from 20 to 75 per cent. Therefore the observed triplets do not seem to probe any abnormal regions of the IGM.

(iii) The conclusions we derived based on the transmitted flux statistics are also confirmed using cloud based statistics obtained using Voigt profile decomposition of Ly $\alpha$  forest. Even though this approach requires more resources it has distinct advantages: (a) as the measured correlation strengths are directly related to real correlations, this allows straight forward interpretations of  $\xi$ ,  $\zeta$ , and  $Q$  and (b) it also allows us to study the dependence of clustering on  $N_{\text{HI}}$ . Cloud based correlation function studies are now possible thanks to the availability of automatic Voigt profile fitting codes like VIPER and high-performance computing.

(iv) Our simulations show that two-point and three-point correlation functions and the reduced three-point correlation (i.e.  $Q$ )

depend strongly of  $N_{\text{HI}}$  for a given spatial separations probed. We observe this trend in the observation too, though the actual dependence of  $N_{\text{HI}}$  seems weaker than what we find in our simulations in case of the two-point correlations.

(v) We searched for large void regions using concurrent gaps found among the three sightlines. The largest concurrent flux gap of  $17 \text{ \AA}$  is observed in ‘Triplet 1’, which corresponds to a length-scale of  $14.2 h^{-1} \text{ cMpc}$ . This is bigger than such gaps reported in the literature (Rollinde et al. 2003; Cappetta et al. 2010). A search for similar gap in the simulations have yields 14.2 per cent sightlines showing a concurrent gap similar to or larger than what we have observed. We also probe the gap statistics using cloud distributions. The above-mentioned void is also picked by this statistics.

(vi) We study the clustering of  $\text{Ly}\alpha$  absorption around quasars using ‘Triplet 2’ and ‘Doublet’ sightlines. As all three quasars are at very similar redshift in the case of ‘Triplet 1’ we did not use them for this study. Based on the inferred Lyman continuum luminosity of the quasars and UV-background intensity we expect all the background quasar sightlines to be affected by the transverse proximity effect. However, consistent with previous studies of quasar pairs, we do find excess HI absorption at the redshifts of the foreground quasars within the expected proximity regions.

(vii) In the case of ‘Triplet 2’ we detect DLAs along all three sightlines within a redshift interval of  $2.22 \leq z \leq 2.32$ . The small probability associated with such a coincidence and the presence of a fourth quasar with a projected separation of  $\sim 1 \text{ Mpc}$  to all the three sightlines suggests a possible presence of large overdense region. It will be interesting to confirm this by deep imaging or integral field spectroscopic observations of this region.

(viii) For the DLAs detected in the spectra of quasars in ‘Triplet 2’, we study the transverse clustering of  $\text{Ly}\alpha$  absorption along other two sightlines. While we do not find any significant excess number of  $\text{Ly}\alpha$  absorption within  $\pm 1000 \text{ km s}^{-1}$  to the DLA, we do find the presence of significant excess  $\text{Ly}\alpha$  absorption with  $N_{\text{HI}} \sim$  few times  $10^{16} \text{ cm}^{-2}$ . Given the length-scale probed these absorbers may not belong to the circumgalactic medium of the quasars as studied by Rubin et al. (2015), but may reflect biased clustering of the IGM around DLAs (see, for example, Pérez-Ràfols et al. 2018). As there are indications that these DLAs may be in the overdense regions, it will be interesting to check whether occurrence of partial LLS at  $\sim 1 \text{ pMpc}$  is special to this region or generic to high- $z$  DLAs.

(ix) Consistent with previous studies we detect correlation signal within  $5 h^{-1} \text{ cMpc}$  (or velocity scale of  $\sim 500 \text{ km s}^{-1}$ ) for longitudinal two-point correlation between C IV absorbers. We do not find any excess clustering in the transverse direction between the C IV absorbers over the projected separations probed in this study.

(x) We have studied the C IV– $\text{Ly}\alpha$  cross-correlation along longitudinal and transverse directions using pairs of sightlines. We find this correlation to be higher (up to  $10 h^{-1} \text{ cMpc}$ ) when we consider higher column density (i.e.  $N_{\text{HI}} > 10^{15} \text{ cm}^{-2}$ )  $\text{Ly}\alpha$  absorbers. It will be important to confirm these findings as through this one can get independent constraints on the objects hosting C IV absorbers.

Using our simulations, we estimate the total redshift path-length required to detect  $\zeta$  at  $5\sigma$  level for two triplets considered here. We found that for flux based statistics, we need a redshift path-length  $\geq 16$  (for  $\Delta r_{\parallel} = 0 h^{-1} \text{ cMpc}$ ) and 24 (for  $\Delta r_{\parallel} = 2 h^{-1} \text{ cMpc}$ ) for resolving the three-point correlation with  $5\sigma$  detectability for ‘Triplet 1’. The corresponding values are  $\geq 55$  (for  $\Delta r_{\parallel} = 0 h^{-1} \text{ cMpc}$ ) and 65 (for  $\Delta r_{\parallel} = 2 h^{-1} \text{ cMpc}$ ) for ‘Triplet 2’. For the cloud based statistics with  $N_{\text{HI}} > 13$ , the required redshift path-length is 44 for ‘Triplet 1’. In case of ‘Triplet 2’ configuration, the cloud based

statistics essentially gives 0 three-point correlation up to 2 decimal places, thus making it difficult to resolve with  $5\sigma$  detectability.

## ACKNOWLEDGEMENTS

RS and PPJ gratefully acknowledge the support of the Indo-French Centre for the Promotion of Advanced Research (Centre Franco-Indien pour la Promotion de la Recherche Avancée) under contract no. 5504–2. We acknowledge the use of High performance computing facility PERSEUS at IUCAA, Pune and thank Kandaswamy Subramanian and Aseem Paranjape for useful discussions. We also thank the anonymous referee for his/her comments which helped improve this work.

## REFERENCES

- Aguirre A., Hernquist L., Schaye J., Katz N., Weinberg D. H., Gardner J., 2001, *ApJ*, 561, 521
- Aracil B., Petitjean P., Smette A., Surdej J., Mückel J. P., Cristiani S., 2002, *A&A*, 391, 1
- Ata M. et al., 2018, *MNRAS*, 473, 4773
- Bahcall J. N., Peebles P. J. E., 1969, *ApJ*, 156, L7
- Becker G. D., Bolton J. S., 2013, *MNRAS*, 436, 1023
- Bi H. G., Boerner G., Chu Y., 1992, *A&A*, 266, 1
- Boksenberg A., Sargent W. L. W., Rauch M., 2003, in Perez E., Gonzalez Delgado R. M., Tenorio-Tagle G., eds, ASP Conf. Ser. Vol. 297, Star Formation Through Time. Astron. Soc. Pac., San Francisco, p. 447
- Bolton J. S., Becker G. D., Raskutti S., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2012, *MNRAS*, 419, 2880
- Cappetta M., D’Odorico V., Cristiani S., Saitta F., Viel M., 2010, *MNRAS*, 407, 1290
- Carswell R. F., Webb J. K., 2014, Astrophysics Source Code Library, record ascl:1408.015
- Caucci S., Colombi S., Pichon C., Rollinde E., Petitjean P., Sousbie T., 2008, *MNRAS*, 386, 211
- Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, *ApJ*, 437, L9
- Choudhury T. R., Srianand R., Padmanabhan T., 2001, *ApJ*, 559, 29
- Coppolani F. et al., 2006, *MNRAS*, 370, 1804
- Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, *ApJ*, 581, 20
- D’Odorico V. et al., 2006, *MNRAS*, 372, 1333
- Evans C. J. et al., 2016, *Proc. SPIE Conf. Ser. Vol. 9908, Science requirements and trade-offs for the MOSAIC instrument for the European ELT*. SPIE, Bellingham, p. 99089J
- Fan X. et al., 2006, *AJ*, 132, 117
- Finley H., Petitjean P., Noterdaeme P., Pâris I., 2014, *A&A*, 572, A31
- Gaikwad P., Khaire V., Choudhury T. R., Srianand R., 2017a, *MNRAS*, 466, 838
- Gaikwad P., Srianand R., Choudhury T. R., Khaire V., 2017b, *MNRAS*, 467, 3172
- Gaskell C. M., 1982, *ApJ*, 263, 79
- Goldoni P., Royer F., François P., Horrobin M., Blanc G., Vernet J., Modigliani A., Larsen J., 2006, in McLean I. S., Iye M., eds, *Proc. SPIE Conf. Ser. Vol. 6269, Data reduction software of the X-shooter spectrograph*. SPIE, Bellingham, p. 62692K
- Guimarães R., Petitjean P., Rollinde E., de Carvalho R. R., Djorgovski S. G., Srianand R., Aghaee A., Castro S., 2007, *MNRAS*, 377, 657
- Hennawi J. F., Prochaska J. X., Cantalupo S., Arrigoni-Battaia F., 2015, *Science*, 348, 779
- Horowitz B., Lee K.-G., White M., Krolewski A., Ata M., 2019, preprint ([arXiv:1903.09049](https://arxiv.org/abs/1903.09049))
- Hui L., Gnedin N. Y., 1997, *MNRAS*, 292, 27
- Jalan P., Chand H., Srianand R., 2019, *ApJ*, 884, 151
- Keeney B. A., Danforth C. W., Stocke J. T., France K., Green J. C., 2012, *PASP*, 124, 830
- Khaire V., Srianand R., 2019, *MNRAS*, 484, 4174

- Kim T.-S., Partl A. M., Carswell R. F., Müller V., 2013, *A&A*, 552, A77
- Kirkman D., Tytler D., 2008, *MNRAS*, 391, 1457
- Krogager J.-K., Møller P., Fynbo J. P. U., Noterdaeme P., 2017, *MNRAS*, 469, 2959
- Krolewski A. et al., 2018, *ApJ*, 861, 60
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Lau M. W., Prochaska J. X., Hennawi J. F., 2016, *ApJS*, 226, 25
- Lee K.-G. et al., 2014, *ApJ*, 795, L12
- Lee K.-G. et al., 2018, *ApJS*, 237, 31
- McDonald P., 2003, *ApJ*, 585, 34
- Miralda-Escudé J., Cen R., Ostriker J. P., Rauch M., 1996, *ApJ*, 471, 582
- Monaghan J. J., 1992, *ARA&A*, 30, 543
- Noterdaeme P. et al., 2012, *A&A*, 547, L1
- Oppenheimer B. D., Davé R., 2006, *MNRAS*, 373, 1265
- Peebles P. J. E., 1980, *The large-scale structure of the universe*. Princeton Univ. Press, Princeton
- Peebles M. S., Weinberg D. H., Davé R., Fardal M. A., Katz N., 2010a, *MNRAS*, 404, 1281
- Peebles M. S., Weinberg D. H., Davé R., Fardal M. A., Katz N., 2010b, *MNRAS*, 404, 1295
- Pérez-Ràfols I. et al., 2018, *MNRAS*, 473, 3019
- Petitjean P., Mueket J. P., Kates R. E., 1995, *A&A*, 295, L9
- Petitjean P., Surdej J., Smette A., Shaver P., Muecket J., Remy M., 1998, *A&A*, 334, L45
- Pichon C., Vergely J. L., Rollinde E., Colombi S., Petitjean P., 2001, *MNRAS*, 326, 597
- Pichon C., Scannapieco E., Aracil B., Petitjean P., Aubert D., Bergeron J., Colombi S., 2003, *ApJ*, 597, L97
- Porqueres N., Jasche J., Lavaux G., Enßlin T., 2019, *A&A*, 630, A151
- Prochaska J. X. et al., 2013, *ApJ*, 776, 136
- Rahmani H. et al., 2016, *MNRAS*, 463, 980
- Rauch M., Haehnelt M. G., 1995, *MNRAS*, 275, L76
- Rauch M., Sargent W. L. W., Womble D. S., Barlow T. A., 1996, *ApJ*, 467, L5
- Rollinde E., Petitjean P., Pichon C., Colombi S., Aracil B., D’Odorico V., Haehnelt M. G., 2003, *MNRAS*, 341, 1279
- Rollinde E., Theuns T., Schaye J., Pâris I., Petitjean P., 2013, *MNRAS*, 428, 540
- Rorai A. et al., 2017, *Science*, 356, 418
- Rubin K. H. R., Hennawi J. F., Prochaska J. X., Simcoe R. A., Myers A., Lau M. W., 2015, *ApJ*, 808, 38
- Rudie G. C., Steidel C. C., Pettini M., 2012, *ApJ*, 757, L30
- Scannapieco E., Pichon C., Aracil B., Petitjean P., Thacker R. J., Pogossyan D., Bergeron J., Couchman H. M. P., 2006, *MNRAS*, 365, 615
- Schaye J., 2001, *ApJ*, 559, 507
- Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, *MNRAS*, 318, 817
- Scoccimarro R., Hui L., Manera M., Chan K. C., 2012, *Phys. Rev. D*, 85, 083002
- Skidmore W., TMT International Science Development Teams, Science Advisory Committee TMT, 2015, *Res. Astron. Astrophys.*, 15, 1945
- Smette A., Robertson J. G., Shaver P. A., Reimers D., Wisotzki L., Koehler T., 1995, *A&AS*, 113, 199
- Smith B. D., Hallman E. J., Shull J. M., O’Shea B. W., 2011, *ApJ*, 731, 6
- Springel V., 2005, *MNRAS*, 364, 1105
- Szapudi I., Szalay A. S., 1998, *ApJ*, 494, L41
- Tie S. S., Weinberg D. H., Martini P., Zhu W., Peirani S., Suarez T., Colombi S., 2019, *MNRAS*, 487, 5346
- Vernet J. et al., 2011, *A&A*, 536, A105
- Viel M., Matarrese S., Mo H. J., Theuns T., Haehnelt M. G., 2002, *MNRAS*, 336, 685
- Viel M., Haehnelt M. G., Springel V., 2004, *MNRAS*, 354, 684
- Worseck G., Davies F. B., Hennawi J. F., Prochaska J. X., 2019, *ApJ*, 875, 111

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://academic.oup.com/mnras/article/490/3/3633/5585410) online.

**Appendix A.** Quasar associated absorption.

**Appendix B.** Metal lines in the observed sightlines.

**Appendix C.** Observed spectra.

**Appendix D.** Dependence of transverse cloud correlation on redshift space binning.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.