The Astrophysical Journal, 835:281 (12pp), 2017 February 1 doi:10.3847/1538-4357/835/2/281

Cosmic Galaxy-IGM H I Relation at $z \sim 2$–3 Probed in the COSMOS/UltraVISTA 1.6 Deg^2 Field

Shiro Mukae[1,2], Masami Ouchi[1,3], Koki Kakiichi[4,5], Nao Suzuki[3], Yoshiaki Ono[1], Zheng Cai[6], Akio K. Inoue[7], Yi-Kuan Chiang[8], Takatoshi Shibuya[1], and Yuichi Matsuda[9,10]

[1] Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan; mukae@icrr.u-tokyo.ac.jp
[2] Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, 113-0033, Japan
[3] Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8583, Japan
[4] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[5] Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85748 Garching bei München, Germany
[6] UCO/Lick Observatory, University of California, 1156 High Street, Santa Cruz, CA 95064, USA
[7] College of General Education, Osaka Sangyo University, 3-1-1, Nakagaito, Daito, Osaka 574-8530, Japan
[8] Department of Astronomy, University of Texas at Austin, 1 University Station C1400, Austin, TX 78712, USA
[9] National Astronomical Observatory of Japan, Osawa 2-21-1, Mitaka, Tokyo 181-8588, Japan
[10] Graduate University for Advanced Studies (SOKENDAI), Osawa 2-21-1, Mitaka, Tokyo 181-8588, Japan
Received 2016 May 1; revised 2016 December 26; accepted 2016 December 29; published 2017 February 2

Abstract

We present spatial correlations of galaxies and IGM neutral hydrogen H I in the COSMOS/UltraVISTA 1.62 deg^2 field. Our data consist of 13,415 photo-z galaxies at $z \sim 2$–3 with $K_s < 23.4$ and the Lyα forest absorption lines in the background quasar spectra selected from SDSS data with no signature of damped Lyα system contamination. We estimate a galaxy overdensity $\delta_{\rm gal}$ in an impact parameter of 2.5 (proper) Mpc, and calculate the Lyα forest fluctuations $\delta_{\langle F\rangle}$ whose negative values correspond to the strong Lyα forest absorption lines. We identify weak evidence of an anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ with a Spearman's rank correlation coefficient of -0.39 suggesting that the galaxy overdensities and the Lyα forest absorption lines positively correlate in space at the $\sim$90% confidence level. This positive correlation indicates that high-z galaxies exist around an excess of H I gas in the Lyα forest. We find four cosmic volumes, dubbed $A_{\rm obs}$, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$, that have extremely large (small) values of $\delta_{\rm gal} \simeq 0.8$ (-1) and $\delta_{\langle F\rangle} \simeq 0.1(-0.4)$, three of which, $B_{\rm obs}$–$D_{\rm obs}$, significantly depart from the $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$ correlation, and weaken the correlation signal. We perform cosmological hydrodynamical simulations and compare with our observational results. Our simulations reproduce the $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$ correlation, agreeing with the observational results. Moreover, our simulations have model counterparts of $A_{\rm obs}$–$D_{\rm obs}$, and suggest that the observations pinpoint, by chance, a galaxy overdensity like a proto-cluster, gas filaments lying on the quasar sightline, a large void, and orthogonal low-density filaments. Our simulations indicate that the significant departures of $B_{\rm obs}$–$D_{\rm obs}$ are produced by the filamentary large-scale structures and the observation sightline effects.

Key words: galaxies: formation – intergalactic medium – large-scale structure of universe – quasars: absorption lines

1. Introduction

The link between baryons and the cosmic web is a clue to understanding both the galaxy formation and the baryonic processes in large-scale structures (LSSs). The processes between galaxies and the intergalactic medium (IGM) are the inflow, which represents gas accretion on to galaxies, and the outflow driven by supernovae and active galactic nuclei. Neutral hydrogen H I in the IGM is probed with the Lyα forest absorption lines in spectra of background quasars (e.g., Faucher-Giguère et al. 2008; Becker et al. 2013; Prochaska et al. 2013) and bright star-forming galaxies (e.g., Steidel et al. 2010; Thomas et al. 2014; Mawatari et al. 2016).

The detailed properties of galaxy-IGM H I relations (hereafter galaxy-H I relation) have been studied by spectroscopic observations of the Keck Baryonic Structure Survey (KBSS: Rakic et al. 2012; Rudie et al. 2012; Turner et al. 2014), Very Large Telescope LBG Redshift Survey (VLRS: Crighton et al. 2011; Tummuangpak et al. 2014), and other programs (e.g., Adelberger et al. 2003, 2005). These spectroscopic observations target H I gas of the circumgalactic medium (CGM) around Lyman break galaxies (LBGs) that are high-z star-forming galaxies identified with a bright UV and blue continuum.

These LBG spectroscopy for the galaxy-H I studies alone do not answer to the following two questions. One is the relation between IGM H I and galaxies that are not selected as LBGs. Because LBGs are identified in their dust-poor star-forming phase, dust-rich and old-stellar population galaxies are missing in the past studies. In fact, the average star-formation duty cycle (DC) of LBGs is estimated to be $\sim$30%–60% (Lee et al. 2009; Harikane et al. 2016). A large fraction of galaxies are not investigated in the studies of the galaxy-H I relation. The other question is what the galaxy-H I relation in a large scale is. To date, the previous studies have investigated LBG-H I relations around sightlines of background quasars within $\sim$1 deg^2 corresponding to $\sim$70 $\times$ 70 comoving Mpc^2 at $z \sim 2$–3 (Adelberger et al. 2003; Rudie et al. 2012; Tummuangpak et al. 2014). There has been no study on the galaxy-H I relation on a large scale (>1 deg^2) at $z \sim 2$–3. Only at $z \leqslant 1$, Tejos et al. (2014) conduct spectroscopic surveys to investigate galaxy-H I relations on a large scale (>1 deg^2). Tejos et al. (2014) present the clustering analysis of spectroscopic galaxies and H I absorption line systems. At $z \sim 2$–3, Cai et al. (2016) have studied the galaxy-H I relation focusing on extremely massive overdensities with $\sim$6,000 SDSS quasar spectra by the MApping the Most Massive Overdensity

Through Hydrogen (MAMMOTH) survey, but the galaxy-H I relations have not been systematically explored.

We investigate spatial correlations of K_s-band selected galaxies with no DC dependence and IGM H I at $z \sim 2$–3 in a large 1.62 deg^2 area of the COSMOS/UltraVISTA field, in conjunction with the comparisons with our models of the cosmological hydrodynamical simulations. We probe one large field contiguously covering LSSs. Our study of the galaxy-H I spatial correlation is complementary to the on-going programs of the MAMMOTH and the Lyα forest tomography survey of the COSMOS Lyα Mapping And Tomography Observations (CLAMATO: Lee et al. 2014, 2016), which aims to illustrate the distribution of IGM H I gas in LSSs. In contrast, our study focuses on a spatial relation between galaxies and IGM H I gas.

This paper is organized as follows. We describe the details of our sample galaxies and background quasars in Section 2. In Section 3, our data analysis is presented. We investigate the galaxy-H I relation based on the observational data in Section 4. We introduce our simulations to examine the galaxy-H I relation of our observational results in Section 5. In Section 6, we compare observation and simulation results and interpret our observational findings. Finally, we summarize our results in Section 7.

Throughout this paper, we adopt AB magnitudes (Oke & Gunn 1983). We use a cosmological parameter set of (Ω_m, Ω_Λ, Ω_b, σ_8, n_s, h) = (0.26, 0.74, 0.045, 0.85, 0.95, 0.72), which is consistent with the nine-year *WMAP* result (Hinshaw et al. 2013). We denote pkpc and pMpc (ckpc and cMpc) to indicate distances in proper (comoving) units.

2. Data

2.1. Photometric Galaxy Samples

We investigate galaxy overdensities in the COSMOS/UltraVISTA field. Our photometric galaxy sample is taken from the COSMOS/UltraVISTA catalog, which is a K_s-band selected galaxy catalog (Muzzin et al. 2013a) made in the 1.62 deg^2 area of UltaVISTA DR1 imaging region (McCracken et al. 2012) in the COSMOS field (Scoville et al. 2007).

The COSMOS/UltraVISTA catalog consists of point-spread function matched photometry of 30 photometric bands. These photometric bands cover the wavelength range of 0.15–24 μm that includes the *GALEX* FUV and NUV (Martin et al. 2005), Subaru/SurimeCam (Taniguchi et al. 2007), CFHT/MegaCam (Capak et al. 2007), UltraVISTA (McCracken et al. 2012), and *Spitzer* IRAC+MIPS data (Sanders et al. 2007). Photometric redshifts for all galaxies are computed with the EAZY code (Brammer et al. 2008). The catalog contains ~150,000 galaxies at $z \sim 0$–5. Stellar masses are determined by SED fitting by the FAST code (Kriek et al. 2009) with stellar population synthesis models.

We use the criteria of Chiang et al. (2014) to select $z \sim 2$–3 photo-z galaxies from the catalog. We apply a 90% completeness limit of $K_s < 23.4$ mag that corresponds to a stellar mass limit of $\log_{10}(M_*/M_\odot) > 9.7$ at $z \sim 2.5$. Note that this stellar mass limit of $z = 2.5$ differs only by 3% at the edges of our redshift window, $z = 2$ and 3. This stellar mass limit is as large as ~0.1 M^* at $z \sim 2.5$, where M^* is the characteristic stellar mass of a Schechter function parameter (Schechter 1976) for the stellar mass functions (SMFs) taken from Muzzin et al. (2013b). We remove objects (~4%) whose photometric redshifts show broad and/or multi-modal redshift probability distributions indicating poorly determined redshifts. Here, the photo-z galaxies that we use have the redshift distribution function for which the 80% probability distribution extends no larger than $\Delta z = 0.2$ from the best estimate redshift. Finally, our photometric samples consist of 13,415 photo-z galaxies at $z \sim 2$–3 with $K_s < 23.4$.

2.2. Background Quasar Samples

We search for the Lyα forest absorption lines found in background quasar spectra in the COSMOS/UltraVISTA field. Our background quasar spectra are primarily taken from the BOSS Data Release 9 (DR9) Lyα Forest Catalog (Lee et al. 2013, hereafter L13). L13 has reproduced quasar continua by the technique of mean-flux regulated principal component analysis (MF-PCA) continuum fitting (Lee et al. 2012). Because L13 does not include all quasars identified by the SDSS-III surveys (Eisenstein et al. 2011), our background quasar spectra are also taken from the BOSS Data Release 12 (BOSS DR12) and the SDSS-III Data Release 12 (SDSS DR12) (Alam et al. 2015). The BOSS DR9 and DR12 spectra are covered in the wavelength range of 3600–10400 Å. The SDSS DR12 spectra are obtained in the wavelength range of 3800–9200 Å, which is slightly narrower than the BOSS wavelength range. Both the BOSS and the SDSS spectra have the spectral resolution of $R \equiv \lambda/\Delta\lambda \approx 2000$.

Because L13 compile spectra of quasars with a redshift range of $z_{\rm qso} > 2.15$, we search for background quasars at $z_{\rm qso} > 2.15$ from the BOSS and the SDSS data. We find a total of 26 background quasars in the COSMOS/UltraVISTA field. We remove four background quasars that are located at the edge of the COSMOS/UltraVISTA field, because the cylinder volumes of these background quasars are cut by the COSMOS/UltraVISTA-field border by >50%.

To identify the Lyα forest absorption lines, we adopt the Lyα forest wavelength range of 1041–1185 Å in the quasar rest frame. With the speed of light c, this wavelength range is defined as

$$\lambda_{{\rm Ly}\beta}\left(1+\frac{5000\ {\rm km\ s^{-1}}}{c}\right) < \lambda < \lambda_{{\rm Ly}\alpha}\left(1-\frac{8000\ {\rm km\ s^{-1}}}{c}\right), \quad (1)$$

where $\lambda_{{\rm Ly}\beta}$ and $\lambda_{{\rm Ly}\alpha}$ are the rest-frame wavelengths of the hydrogen Lyβ (1025.72 Å) and Lyα (1215.67 Å) lines, respectively. Here, we include the velocity offsets of 5000 and 8000 km s^{-1}, avoiding the Lyβ forest contamination and the quasar proximity effect, respectively.

In the observed spectra, the Lyα forest wavelength range shifts by a factor of $(1 + z_{\rm qso})$. We search for background quasars whose Lyα forest wavelength ranges cover Lyα absorption lines. Because we investigate Lyα absorption lines at $2 \leqslant z_{{\rm Ly}\alpha} \leqslant 3$ of the COSMOS/UltraVISTA field, Lyα absorption wavelength range is 3647 Å $\leqslant \lambda_{{\rm Ly}\alpha}(1 + z_{{\rm Ly}\alpha}) \leqslant$ 4862 Å in the observed frame that requires $2.08 \leqslant z_{\rm qso} \leqslant 3.67$. Limiting $z_{\rm qso} > 2.15$ given by the L13 spectra, we select 21 background quasars at $2.15 < z_{\rm qso} \leqslant 3.67$ further removing 1 spectrum at $z_{\rm qso} > 3.67$.

Then, we investigate qualities of background quasar spectra. We define S/N$_{{\rm Ly}\alpha}$ as the median signal-to-noise ratio (S/N) per pixel over the Lyα forest wavelength range (1041–1185 Å). Because we find that the absorption signals are not reasonably obtained in seven background quasar spectra with S/N$_{{\rm Ly}\alpha} < 2$

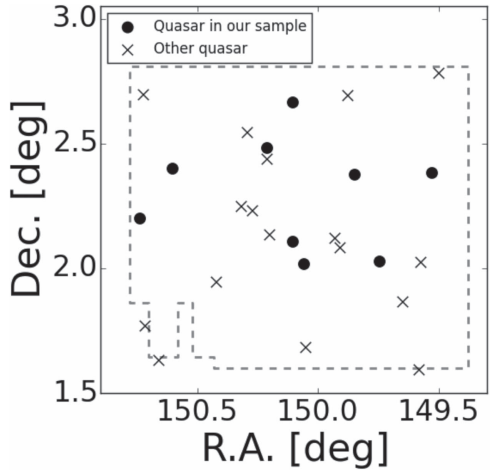

Figure 1. Sky distribution of the 26 background quasars found in the COSMOS/UltraVISTA field. The circles indicate background quasars at $z_{\rm qso} > 2.15$ analyzed in our study, while the crosses are those removed from our analysis. The gray dashed lines show the coverage of the COSMOS/UltraVISTA.

by visual inspection, we remove these seven quasar spectra in our analysis.

We discard four broad absorption line (BAL) quasar spectra referring to the SDSS database. In addition, we check background quasar spectra by visual inspection, and remove one spectra with large flux fluctuations originated from unknown systematics.

Finally, we use nine background quasar spectra in the COSMOS/UltraVISTA field. Figure 1 shows the distribution of the background quasars in the COSMOS/UltraVISTA field. The mean $S/N_{Ly\alpha}$ in our quasar samples is $S/N_{Ly\alpha} \sim 5$.

3. Galaxy Overdensity and H I Absorption

3.1. Galaxy Overdensity

We estimate galaxy overdensities around the quasar sightlines where the Lyα forest absorption lines are observed. The galaxy overdensities are calculated with the COSMOS/UltraVISTA catalog (Section 2.1). The galaxy overdensity $\delta_{\rm gal}$ is defined as

$$\delta_{\rm gal} \equiv \frac{n_{\rm gal}(z)}{\overline{n}_{\rm gal}(z)} - 1, \quad (2)$$

where $n_{\rm gal}$ ($\overline{n}_{\rm gal}$) is the galaxy (average) number density in a cylinder at the redshift z of the cylinder center. The redshift range of $n_{\rm gal}$ and $\overline{n}_{\rm gal}$ are defined by the redshift range of the cylinder length. The base area of the cylinder is defined by a radius of $r = 5'$ corresponding to an impact parameter of ~2.5 pMpc at $z \sim 2.5$. The length of the cylinder along a line of sight is given by an average photometric redshift uncertainty that corresponds to 25 pMpc. The estimated average photometric redshift uncertainty is $\sigma_z = 0.025(1+z)$ for galaxies with $K_s < 23.4$ mag at $2 \leqslant z \leqslant 3$ (Scoville et al. 2013). The error of $\delta_{\rm gal}$ is estimated with the combination of the photo-z uncertainties and the Poisson errors.

3.2. Lyα Forest Absorption Lines

To investigate the Lyα forest absorption lines, we do not use spectra in the wavelength range where damped Lyα systems (DLAs) contaminate the spectra. Here, we search for DLAs in our spectra, performing DLA catalog matching and visual inspection. As explained in Section 2.2, our spectra are taken from the three data sets of L13, BOSS DR12, and SDSS DR12. For the data set of L13, DLAs are already removed based on the DLA catalog of Noterdaeme et al. (2012) who identify DLAs in the BOSS spectra by Voigt-profile fitting. For the data set of BOSS DR12, we find no DLAs in the Lyα forest wavelength range in the Noterdaeme et al.'s DLA catalog. For the data set of SDSS DR12, we perform visual inspection, and identify no DLAs.

Quasar host galaxies cause intrinsic strong metal absorption lines of SIV λ1062.7, CIII λ1175.7, NII λ1084.0, and NI λ1134.4 in the Lyα forest wavelength range. We mask out the sufficient wavelength width ± 5 Å around these metal absorption lines.

The study of L13 has reproduced quasar continua by the MF-PCA continuum fitting technique (Lee et al. 2012). The MF-PCA continuum fitting technique is essentially composed of two steps: (1) an initial PCA fit to the redward of the Lyα emission line to reproduce the Lyα forest continuum and (2) tuning the Lyα forest continuum amplitude extrapolated to the blueward of the Lyα emission line with the cosmic Lyα forest mean transmission of Faucher-Giguère et al. (2008). We obtain quasar continua of the BOSS DR12 and the SDSS DR12 spectra by the MF-PCA technique with the code used in L13. We include estimated median r.m.s, continuum fitting errors that are (7, 5.5, 4.5, 4%) for spectra with $S/N_{Ly\alpha} =$ (2–4, 4–6, 6–10, 10–15) at $z \sim 2.5$–3 (Lee et al. 2012). We calculate the Lyα forest transmission $F(z)$ at each pixel:

$$F(z) = f(z)/C(z), \quad (3)$$

where $f(z)$ and $C(z)$ are the observed flux and the quasar continuum in the Lyα forest wavelength range, respectively.

We investigate the Lyα forest absorption lines in the cylinders used by the galaxy overdensity calculation (Section 3.1). We carry out binning for our spectra with the redshift range of $dz = 0.025(1+z)$ that corresponds to the length of the cylinder and obtain an average Lyα forest transmission $\langle F \rangle_{dz}$ and its error $\langle \sigma_F \rangle_{dz}$. Here, $\langle \sigma_F \rangle_{dz}$ is estimated with pixel noises in the spectra and continuum fitting errors. The absorption of the Lyα forest is defined as $\mathrm{DA} \equiv 1 - \langle F \rangle_{dz}$. We refer to the signal-to-noise ratio of the DA detection $S/N_{\langle F \rangle}$ as $\mathrm{DA}/\langle \sigma_F \rangle_{dz}$. We calculate the $S/N_{\langle F \rangle}$ in the Lyα forest wavelength range, and determine z_h, where z_h is the redshift of the highest $S/N_{\langle F \rangle}$. We put the first cylinder centered at z_h in each spectrum. We place additional cylinders that lie next to each other around the first cylinder.

To obtain statistically reliable results, we make use of the cylinders whose $S/N_{\langle F \rangle}$ is the highest in each sightline. However, there is a possibility that this procedure for the cylinder placement would bias the results. Here, we change the central redshift of the first cylinder from z_h to the following two redshifts, (I) and (II). The two redshifts provide cylinders that cover (I) the lowest and (II) the highest wavelength ranges of the Lyα forest. Taking these different central redshifts of the cylinders, we find that our statistical results of Section 4 change only by 4%–5%.

Each quasar sightline has two to four cylinders usable for our analysis. In total, there are 26 cylinders. The number of cylinders is determined by the wavelength range where the following two ranges of (i) and (ii) overlap. The two ranges are (i) the Lyα forest wavelength range, $1041(1+z_{\rm q})$Å $\leqslant \lambda \leqslant$ 1185

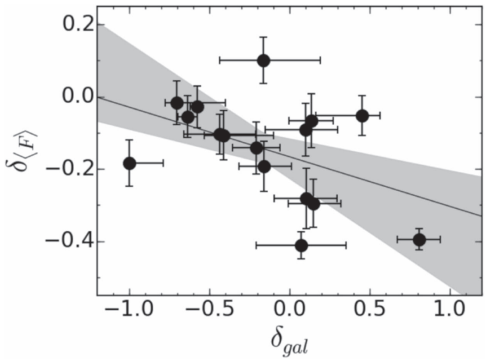

Figure 2. Lyα forest fluctuation $\delta_{\langle F\rangle}$ as a function of the galaxy overdensity $\delta_{\rm gal}$. The circles with the error bars represent the galaxy-H I properties of the cylinders. The best-fit linear model of Equation (6) is represented by a solid line, with the shaded region indicating the 1σ uncertainty range that is calculated by the perturbation method.

$(1+z_{\rm q})$Å and (ii) the $z_{{\rm Ly}\alpha}=2$–3 Lyα forest absorption line range, $3647\,{\rm Å}\leqslant\lambda\leqslant 4862\,{\rm Å}$, in the observed frame (Section 2.2). For example, a quasar at $z_{\rm q}=3.36$ has $4539\,{\rm Å}\leqslant\lambda\leqslant 5166\,{\rm Å}$ for (i). The wavelength range of the (i) and (ii) overlap is $4539\,{\rm Å}\leqslant\lambda\leqslant 4862\,{\rm Å}$. This wavelength range corresponds to $2.73\leqslant z_{{\rm Ly}\alpha}\leqslant 3.0$. Because the length of a cylinder is $0.025(1+z)\sim 0.1$ at this redshift range, we obtain two cylinders from this sightline of the $z_{\rm q}=3.36$ quasar.

We use the data of the cylinders with ${\rm S/N}_{\langle F\rangle}\geqslant 4$ and estimate both $\delta_{\rm gal}$ and $\langle F\rangle_{dz}$ in the cylinders. Here we test whether this cut of ${\rm S/N}_{\langle F\rangle}\geqslant 4$ gives impacts on our results. We change the ${\rm S/N}_{\langle F\rangle}$ cut from 4 to 3 and 5, and carry out the same analysis to evaluate how many different results can be obtained by the different S/N cuts. We find that results of ${\rm S/N}_{\langle F\rangle}\geqslant 3$ and $\geqslant 5$ cuts are very similar to those of ${\rm S/N}_{\langle F\rangle}\geqslant 4$. Thus the different S/N cut within this range has a minimal impact on our results. Note that the ${\rm S/N}_{\langle F\rangle}$ cut below 3 raises the noise level and that the correlation signals are diminished. Moreover, the ${\rm S/N}_{\langle F\rangle}$ cut beyond 5 gives a number of spectra that is too small to investigate the correlations.

By these definitions and selections of the cylinders, we have a total of 16 cylinders for the $\delta_{\rm gal}-\langle F\rangle_{dz}$ measurements. In each cylinder, we calculate the Lyα forest fluctuation for which the negative values correspond to a strong Lyα absorption:

$$\delta_{\langle F\rangle}\equiv\frac{\langle F\rangle_{dz}}{F_{\rm cos}(z)}-1, \tag{4}$$

where $F_{\rm cos}(z)$ is the cosmic Lyα forest mean transmission. We adopt $F_{\rm cos}(z)$ estimated by Faucher-Giguère et al. (2008),

$$F_{\rm cos}(z)=\exp[-0.001845(1+z)^{3.924}]. \tag{5}$$

The error of $\delta_{\langle F\rangle}$ is estimated with the Lyα forest transmission errors $\langle\sigma_F\rangle_{dz}$.

4. Galaxy-IGM H I Correlation

Figure 2 presents $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values in the cylinders. We calculate a Spearman's rank correlation coefficient $\rho_{\rm obs}$ of a nonparametric measure to investigate the existence of a correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$. We obtain $\rho_{\rm obs}=-0.39$, which corresponds to the ∼90% confidence level[11] (Wall & Jenkins 2012).

[11] We find $\rho_{\rm obs}=-0.57$, which corresponds to the ∼96% confidence level, if we remove outliers (Section 6.2).

We estimate the errors of $\rho_{\rm obs}$ by the perturbation method (Curran 2015). We generate 1000 data sets of 16 cylinders for which data values include random perturbations following the Gaussian distribution whose sigma is defined by the observational errors of the 16 cylinders. We obtain 1000 Spearman's ρ values for the 1000 data sets. We define the error of $\rho_{\rm obs}$ as the range of 68% distribution for the 1000 Spearman's ρ values. The estimated error of $\rho_{\rm obs}$ is 0.2 at the 1σ level. The 1σ-error range of $\rho_{\rm obs}$ is thus $\rho_{\rm obs}=-0.6-(-0.2)$. This range of $\rho_{\rm obs}$ corresponds to the confidence level range of ∼80%–100%. In other words, the $\rho_{\rm obs}$ error changes the results of the ∼90% confidence level only by ±∼10%.

We then apply chi-square fitting to the relation of $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$, and obtain the best-fit linear model,

$$\delta_{\langle F\rangle}=-0.17^{+0.06}_{-0.06}-0.14^{+0.06}_{-0.16}\times\delta_{\rm gal} \tag{6}$$

which is shown with the solid line in Figure 2. Figure 2 and Equation (6) suggest weak evidence of an anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$. The suggestive anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ indicates that high-z galaxies exist around an excess of H I gas in the Lyα forest.

There is a possibility that the spatial correlation in the same sightlines would bias the results. To evaluate this possible bias, we choose one cylinder for each sightline that does not have any spatial correlations with the other cylinders, and conduct the same analysis. We find an anti-correlation at the ∼80% confidence level that falls in the 1σ error range of the results with all of the cylinders. We thus conclude that the results do not change by the bias of the correlation that is not as large as the one of statistical uncertainties.

Strong Lyα absorption lines can be made by the CGM of galaxies that lie near the quasar sightlines. Rudie et al. (2012) have studied velocities and spatial locations of H I gas surrounding star-forming galaxies at $z\sim$ 2–3, and found that the H I column density rapidly increases with decreasing an impact parameter within 200 pkpc. We investigate the Lyα absorption lines associated with the CGM. In each cylinder used in Section 3.2, we calculate revised $\delta_{\rm gal}$ values for which galaxy numbers are estimated in hollow cylinders whose inner radius is 0.′4, corresponding to an impact parameter of 200 pkpc. We find that there are only 0–1 galaxies in a 0.′4 radius cylinder. The white circles in Figure 3 represent the hollow cylinder results that are very similar to the black circles in Figure 2. Figure 3 indicates no significant differences in the $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$ distributions and the $\rho_{\rm obs}$ value, and suggests that the CGM of galaxies is not the major source of the small $\delta_{\langle F\rangle}$ values.

5. Simulations

We perform cosmological hydrodynamical simulations with the RAMSES code (Teyssier 2002) to investigate the spatial correlations of galaxies and IGM H I of our observational results (Section 4). The initial conditions are generated with the COSMIC package (Bertschinger 1995) and are evolved using Zel'dovich approximation. We include both dark matter and baryon using N-body plus Eulerian hydrodynamics on a uniform grid. The simulations are performed in a box size of $80h^{-1}$ cMpc length with 512^3 cells and a spatial resolution of $156h^{-1}$ ckpc. We use 512^3 dark matter particles with a mass resolution of $3.16\times10^8\,M_\odot$. The mean gas mass per cell is $5.4\times10^7\,M_\odot$.

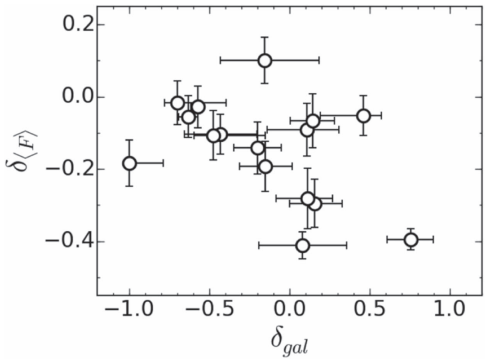

Figure 3. Same as Figure 2, but for revised $\delta_{\rm gal}$ values with hollow cylinders. The galaxy numbers of the revised $\delta_{\rm gal}$ are estimated in hollow cylinders whose inner and outer radii are $0\farcm4$ and $5'$, which corresponds to 200 pkpc and 2.5 pMpc at $z \sim 2.5$, respectively.

We include the ultraviolet background model of Haardt & Madau (1996) at the reionization redshift $z_{\rm reion} = 8.5$. We investigate the gas temperature value at the mean gas density in the simulations and find $T = 1.4 \times 10^4$ K, which is consistent with observational measurements at $z \sim 2$–3 (Becker et al. 2011). We assume the photoionization equilibrium. We apply the optically thin limit, and do not produce any DLAs. Note that our simulations do not include feedback effects on the Lyα forest. Because the feedback mostly affects high-density absorbers with $N_{\rm H\,I} > 10^{16}$ cm^{-2} (Theuns et al. 2001), the lack of the feedback effects does not significantly change the large-scale correlation of galaxy overdensities and Lyα forest absorption lines.

Dark matter halos in the simulations are identified by the HOP algorithm (Eisenstein & Hut 1998). We use dark matter halos containing more than 1000 dark matter particles. We have compared the halo mass functions at $z \sim 2.5$ in our simulations with the halo mass function of the high-resolution N-body simulations (Reed et al. 2007), and found a good agreement within ∼30% in abundance. Our simulations resolve dark matter halos with a mass of $\log_{10} M_h/M_\odot > 11$.

5.1. Mock Galaxy Catalog

We create mock galaxy catalogs from the simulations using the abundance matching technique (e.g., Peacock & Smith 2000; Vale & Ostriker 2004; Moster et al. 2010; Behroozi et al. 2013), which explains observational results of stellar mass functions. We make simulated galaxies, populating each halo with one galaxy. We assume the stellar-to-halo mass ratio (SHMR) with a functional form of

$$M_*/M_h = f_0 \frac{(M_h/M_1)^\alpha}{1 + (M_h/M_1)^{\alpha-\beta}}, \quad (7)$$

where f_0, M_1, α, and β are free parameters. We produce SMFs at $z \sim 2.5$ with many sets of these parameters. We compare these SMFs with observed SMFs (Muzzin et al. 2013b; Tomczak et al. 2014), and find the best-fit parameter set reproducing the observed SMFs. The best-fit parameter set is $(M_1, \alpha, \beta, f_0) = (1.2 \times 10^{12} M_\odot, 1.0, -0.3, 0.04)$. The SHMR with these best-fit parameters is consistent with the one estimated by Behroozi et al. (2013) within the 1σ error levels. We use the simulated galaxies whose stellar mass is $\log_{10}(M_*/M_\odot) > 9.7$ that is the same stellar mass limit in the COSMOS/UltraVISTA catalog (Section 2.1). The stellar mass limit corresponds to the minimum halo mass of $M_h > 4.4 \times 10^{11} M_\odot$. The simulated galaxies consist of 2221 galaxies at $z = 2.4$–2.5, which agree with the number of galaxies at $z = 2.4$–2.5 in our COSMOS/UltraVISTA photometric samples.

5.2. Lyα Forest Catalog

In the simulation boxes, we make mock spectra along the random sightlines parallel to a principal axis defined as the redshift direction. The Lyα transmitted flux is computed with the fluctuating Gunn-Peterson approximation (FGPA; e.g., Weinberg et al. 1998, 2003; Meiksin 2009; Becker et al. 2015), because the FGPA method is simple and fast in computing. We ignore the gas velocities and the effect of redshift-space distortion on the Lyα forest, testing whether this method gives reliable results. The FGPA is a good approximation for absorbers with densities around and below the cosmic mean (Rakic et al. 2012).

We choose two typical sightlines, and conduct full optical depth calculations with gas velocities and redshift-space distortions (e.g., Lukić et al. 2015; Meiksin et al. 2015). We then compare the results of these two sightlines with those given by our original method. We find that the difference is only <6%, which is not as large as the one of statistical uncertainties.

The FGPA gives the Lyα optical depth as

$$\tau = \frac{c\sigma_\alpha \bar{n}_{\rm H}(z)}{\nu_\alpha H(z)} x_{\rm H\,I} \Delta_b \propto \Delta_b^{2-0.72(\gamma-1)}, \quad (8)$$

where σ_α is the Lyα cross-section, $\bar{n}_{\rm H}(z)$ is the cosmic mean density of hydrogen atoms at redshift z, ν_α is the Lyα resonance frequency, $H(z)$ denotes the Hubble constant at redshift z, $x_{\rm H\,I}$ is the fraction of neutral hydrogen, Δ_b is the baryonic density in units of the mean density, and γ is the power-law slope of the temperature–density relation in the IGM (Hui & Gnedin 1997). The γ value is set as $\gamma = 1.5$, which is consistent with observations of the Lyα forest transmissions (Becker et al. 2011; Boera et al. 2014). We investigate the fidelity of our simulated Lyα forest model. We compare the one-dimensional power spectrum of the transmitted flux in our simulations with the SDSS and the BOSS measurements (McDonald et al. 2006; Palanque-Delabrouille et al. 2013), and we find a good agreement. We scale the mean Lyα transmitted flux to the cosmic Lyα forest mean transmission of Faucher-Giguère et al. (2008; Equation (5) of this paper). This scaling method is widely used in the literature (e.g., White et al. 2010; Lukić et al. 2015). We rebin the simulated spectra, and produce the SDSS and the BOSS pixel width of 69 km s^{-1}. We add Gaussian noises to the simulated spectra, accomplishing the $S/N_{\rm Ly\alpha} \sim 5$, which corresponds to the typical $S/N_{\rm Ly\alpha}$ in our quasar samples (Section 2.2).

5.3. Simulated Galaxy-IGM H I Correlation

In Figure 4, the red dots represent a cylinder of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ in our simulations. We make 1000 sets of the 16 cylinders selected from the simulations, mocking our observed 16 cylinders. The red circles in Figure 4 denote one example set of the 16 mock cylinders. We fit a linear model to each of these

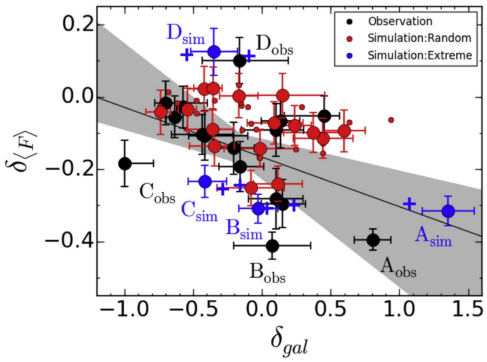

Figure 4. Same as Figure 2, but we overplot our simulation results. The black circles represent the observation data. The red dots are the simulation results. The red circles show an example of 16 cylinders that are randomly chosen from our simulation results (red dots). The blue circles present the simulated results for the four cylinders that have extreme values of δ_{gal} and/or $\delta_{\langle F\rangle}$ similar to A_{obs}–D_{obs} (Section 6.2). The blue crosses are other counterparts found in our simulations (Appendix C).

1000 sets of the 16 mock cylinders, and obtain Spearman's rank correlation coefficient ρ_{sim} values for the 1000 sets. We find that a 1σ distribution of ρ_{sim} corresponds to the range of -0.35–(-0.60), which indicates the existence of the anti-correlation between δ_{gal} and $\delta_{\langle F\rangle}$ in the simulations. This ρ_{sim} distribution corresponds to the $\sim$90% $\pm$ 10% confidence level that is the same as our conclusions for the observational data (Section 4).

To test the convergence of our simulations, we perform simulations that have a box size of 80 h^{-1} cMpc length with 256^3 cells. We detail these results in Appendix A.

We test whether the noise distribution makes significant changes from our conclusions. We use the same $S/N_{Ly\alpha}$ probability distribution as that of our observational spectra of the 16 cylinders. We add noise to our 16 simulation spectra, following the $S/N_{Ly\alpha}$ probability distribution. We calculate Spearman's rank correlation coefficient ρ_{sim} values for 16 simulation spectra with the noise. Conducting this test for $\sim$10 times, we find that the ρ_{sim} values are not different from our original result beyond the statistical errors.

We estimate how the correlation changes when the full redshift range of the observations is considered. From $z = 2.5$ (3) to $z = 2$ (2.5), the structure growth increases only by $\sim$14%. The galaxy clustering and Lyα forest clustering are expected to grow accordingly by $\sim$10%. In Figure 4, this redshift evolution shifts δ_{gal} ($\delta_{\langle F\rangle}$) values of the simulation data points rightward (downward) only by $\sim$10%, which is not as large as that of the statistical uncertainties.

6. Discussion

6.1. Comparison between the Observation and the Simulation Results

Sections 4 and 5.3 present the observation and the simulation results. Both observation and simulation results indicate weak evidence of the anti-correlation between δ_{gal} and $\delta_{\langle F\rangle}$ (Figure 4). Moreover, the Spearman's rank correlation coefficient $\rho_{obs} = -0.39$ from the observations (Section 4) falls in the ρ_{sim} range of -0.35–(-0.60), indicating that simulations well reproduce observational results.

6.2. Four Cylinders with an Extreme Value

In Figure 4, we find four cylinders with the labels of A_{obs}, B_{obs}, C_{obs}, and D_{obs} that have the largest (smallest) values of δ_{gal} or $\delta_{\langle F\rangle}$ among the observational data points. Note that, in Figure 4, A_{obs} falls on the best-fit linear model within the errors, and that the other three cylinders, B_{obs}, C_{obs}, and D_{obs}, show significant departures from the best-fit linear model by 0.1–0.2 in $\delta_{\langle F\rangle}$ at the 2–3σ significance levels. These three cylinders weaken the anti-correlation signal found in Section 4. If we exclude these three cylinders, we obtain $\rho_{obs} = -0.57$, which corresponds to the $\sim$96% confidence level.[12]

We present background quasar spectra of the four cylinders in Figure 5. The full width of the abscissa axis in the four panels of Figure 5 corresponds to the full redshift range of the cylinder. The black points in Figure 5 present the positions of galaxies with the best estimate of photometric redshifts and the impact parameter b in reference to quasar sightlines. We note that the photometric redshift uncertainty is comparable to the full redshift range of the cylinder.

We investigate the physical origin of the extreme δ_{gal} and/or $\delta_{\langle F\rangle}$values in the four cylinders from our observations. We use the simulations performed in Section 5, and search for cylinders in the simulations whose δ_{gal}–$\delta_{\langle F\rangle}$ values are similar to the four cylinders from our observations. With the simulation data, we create the sky map (Figure 6), and identify four mock cylinders that show δ_{gal}–$\delta_{\langle F\rangle}$ values most similar to A_{obs}–D_{obs}. We refer to these mock cylinders as the A_{sim}, B_{sim}, C_{sim}, and D_{sim}. The maps of $\delta_{\langle F\rangle}$ around the sightlines of A_{sim}–D_{sim} are shown in Figures 7 and 8, respectively. In Figure 4, we overplot the δ_{gal}–$\delta_{\langle F\rangle}$ values of A_{sim}–D_{sim}. Below, we describe properties and comparisons of A_{obs}–D_{obs} and A_{sim}–D_{sim}.

Cylinder A: A_{obs} has the largest δ_{gal} and one of the smallest $\delta_{\langle F\rangle}$ values. The top panel of Figure 5 indicates that A_{obs} has the largest number of galaxies and one of the strongest Lyα forest absorption lines among the four cylinders of A_{obs}–D_{obs}. Interestingly, A_{obs} coincides with one of the proto-cluster candidates reported by Chiang et al. (2014). The large δ_{gal} and the small $\delta_{\langle F\rangle}$ values of A_{obs} suggest that a large galaxy overdensity is associated with the large amount of H I gas (Cucciati et al. 2014; Chiang et al. 2015; Cai et al. 2016; Lee et al. 2016). Our simulation results in Figure 7 show that the sightline of A_{sim} penetrates gas filaments of LSSs and a galaxy overdensity like a proto-cluster at $z \sim 2.46$ (label "a"). Moreover, the top panel of Figure 8 indicates that A_{sim} would include a Coherently Strong Lyα Absorption system (CoSLA) that traces massive overdensities on the scale of $\sim$15 h^{-1} cMpc (Cai et al. 2016).

Because there exist high-quality spectra of A_{obs} taken with VLT/X-shooter (Vernet et al. 2011). we show the characteristic of the A_{obs} sightline based on the X-shooter spectrum in Appendix B.

Cylinder B: B_{obs} has the moderate δ_{gal} and the smallest $\delta_{\langle F\rangle}$ values. The second top panel of Figure 5 shows that B_{obs} has the strong Lyα forest absorption lines over the entire redshift range of the cylinder. In our simulations, Figure 7 presents that the sightline of B_{sim} goes through gas filaments at $z \sim 2.43$ (label "b"). Our simulations indicate that the sightline of B_{obs}

[12] In our simulations, there exist sets of cylinders that show significant (>95%) confidence levels after we remove extreme cylinders, which agrees with our observational results.

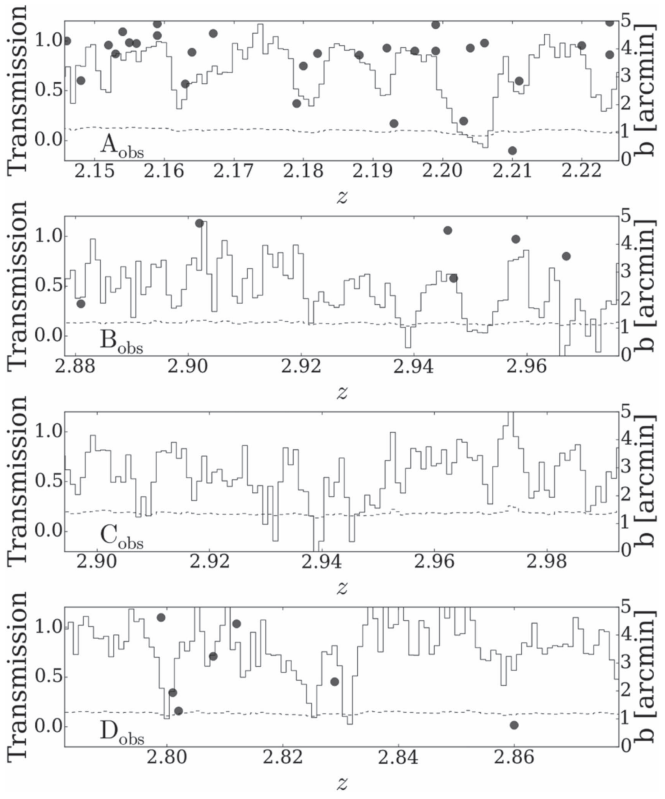

Figure 5. Observed background quasar spectra of the four cylinders that have extreme values of δ_{gal} and/or $\delta_{\langle F\rangle}$ (Section 6.2). The solid lines depict transmission per pixel as a function of redshift for the Lyα forest absorption lines (left-hand axis). The dashed lines represent the noise per pixel. The black points denote redshifts and impact parameters (right-hand axis) of our photo-z galaxies in the cylinders. Note that the photometric redshift uncertainty is comparable to the full redshift range of the cylinder.

would penetrate gas filaments with the moderate number of galaxies.

Cylinder C: C_{obs} has the smallest δ_{gal}. Note that C_{obs} has $\delta_{gal} = -1$ corresponding to no galaxy in the cylinder. The second bottom panel of Figure 5 indicates that C_{obs} does not have galaxies but it has moderately strong Lyα forest absorption lines. In our simulations, we identify C_{sim}, which has δ_{gal}–$\delta_{\langle F\rangle}$ values most similar to those of C_{obs}. Our simulation results in Figure 7 show that the sightline of C_{sim} penetrates a large void of LSSs at $z \sim 2.43$ (label "c1") and goes across gas filaments at $z \sim 2.45$ (label "c2"). Our simulations suggest that C_{obs} would penetrate a large void, and go across gas filaments. Note that C_{sim} has $\delta_{gal} \sim -0.5$, which is larger than the δ_{gal} value of C_{obs}. Our simulations find no cylinders with $\delta_{gal} = -1$ and $\delta_{\langle F\rangle} \sim -0.2$ that C_{obs} has. This is the difference between C_{obs} and C_{sim}. This difference is probably made, because (1) photometric redshifts of galaxies in C_{obs} have catastrophically large errors, (2) there exist faint galaxies whose luminosities are just below the observational limit of $K_s = 23.4$ mag, or (3) a void of galaxies similar to C_{obs} is missing in the limited box size of the simulations.

Cylinder D: The cylinder with the largest $\delta_{\langle F\rangle}$ is D_{obs}. The bottom panel of Figure 5 presents that D_{obs} has the moderately weak Lyα forest absorption lines. In our simulations, Figure 7 shows that the sightline of D_{sim} crosses the low-density

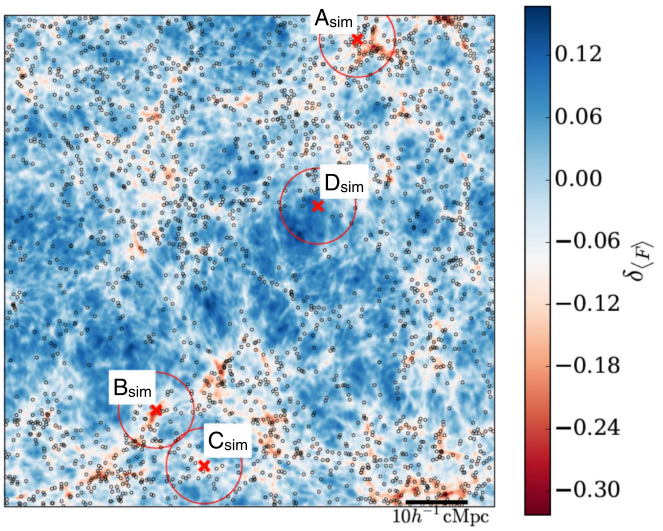

Figure 6. Projected sky map of the distribution of the mock galaxies and the Lyα forest absorption lines. The black circles represent mock galaxies whose stellar masses fall in $\log_{10}M_*/M_\odot > 9.7$. Note that the background color scale represents $\delta_{\langle F\rangle}$ defined in the full redshift range. The positions of sightlines are indicated by the red crosses enclosed by the red circles corresponding to the circumference of cylinders. In the simulations, 10 h^{-1} cMpc corresponds to 8.′5 at $z \sim 2.5$.

filaments. Our simulation results suggest that D_{obs} would go through the orthogonal low-density filaments.

With the results of Figures 5 and 8, we count the numbers of galaxies with photo-z errors in each cylinder. We find (29, 6, 7) galaxies in (A_{obs}, B_{obs}, D_{obs}), while there are (32, 15, 10) galaxies in (A_{sim}, B_{sim}, D_{sim}). These numbers agree within the $\sim$1–2σ levels. However, the number of galaxies in C_{obs} is 0, which is significantly smaller than that of C_{sim} (see above for the difference of C_{obs} and C_{sim}).

Note that there exist other counterparts of A_{obs}–D_{obs} in our simulations. We detail these counterparts in Appendix C.

We investigate properties of other sightlines that do not have extreme values of δ_{gal} and $\delta_{\langle F\rangle}$. We find that these sightlines penetrate neither dense structures, filaments in parallel, nor large voids.

6.3. Summary of the Simulation Comparisons

In Section 6.2, we discuss the physical origins of four cylinders (A_{obs}, B_{obs}, C_{obs}, D_{obs}) that have extremely large (small) values of δ_{gal} and $\delta_{\langle F\rangle}$. We use the simulations performed in Section 5, and identify four cylinders (A_{sim}, B_{sim}, C_{sim}, D_{sim}), whose δ_{gal}–$\delta_{\langle F\rangle}$ values are close to (A_{obs}, B_{obs}, C_{obs}, D_{obs}). The comparisons between A_{obs}–D_{obs} and A_{sim}–D_{sim} suggest that sightlines in the observation would penetrate (1) a galaxy overdensity like a proto-cluster in A_{obs}, (2) gas filaments in B_{obs}, (3) a large void in C_{obs}, and (4) orthogonal low-density filaments in D_{obs}. In this way, our simulations provide the possible physical pictures of these four cylinders based on the structure formation models.

The similarity between our observation and simulation results (Figure 4) supports the standard picture of the galaxy formation scenario in the filamentary LSSs (Mo et al. 2010) on which our simulations are based.

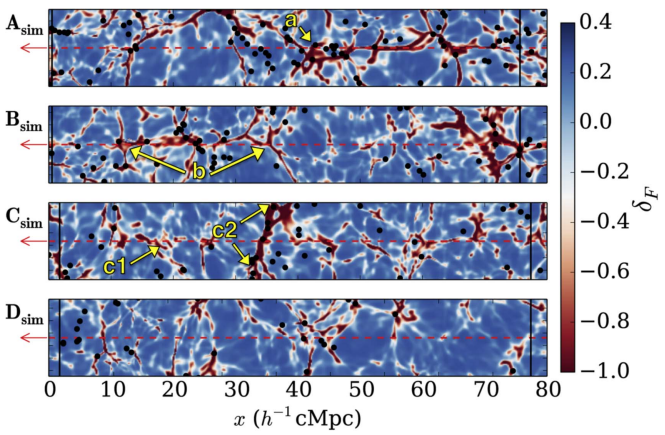

Figure 7. Maps of $\delta_{\langle F\rangle}$ around the sightlines of $A_{\rm sim}$, $B_{\rm sim}$, $C_{\rm sim}$, and $D_{\rm sim}$ (Section 6.2). The black points are projected positions of the mock galaxies in the cylinders. The background color scale represents a $\delta_{\langle F\rangle}$ value per pixel. The black vertical lines represent the lower and upper edges of each cylinder. The red dashed lines represent the sightlines of $A_{\rm sim}$, $B_{\rm sim}$, $C_{\rm sim}$, and $D_{\rm sim}$. The arrows indicate a galaxy overdensity in $A_{\rm sim}$(label "a"), two edges of gas filaments in $B_{\rm sim}$ and $C_{\rm sim}$ (labels "b" and "c2"), and a galaxy void with gas filaments in $C_{\rm sim}$(label "c1").

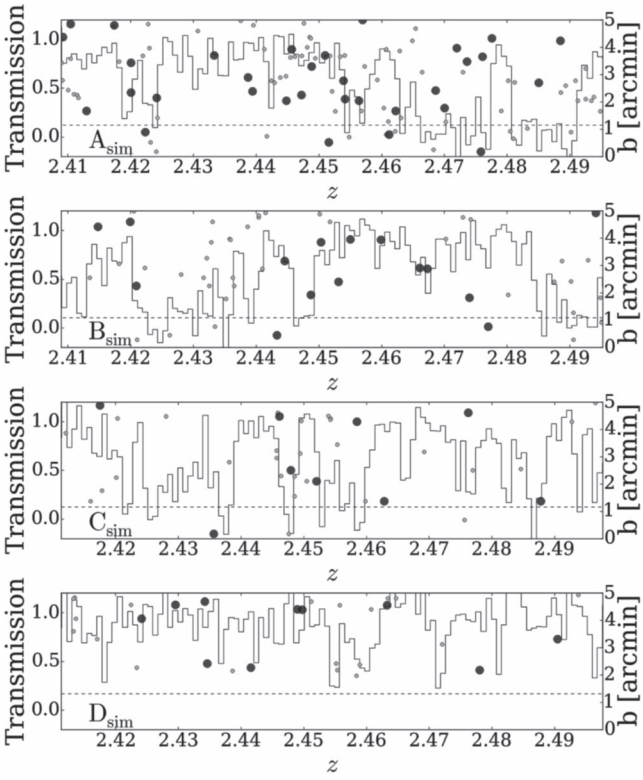

Figure 8. Same as Figure 5, but for the background quasar spectra in the simulations. Each panel corresponds to a cylinder that has extreme values of $\delta_{\rm gal}$ and/or $\delta_{\langle F\rangle}$ in the simulations (Section 6.2). The black points denote redshifts and impact parameters (right-hand axis) of our photo-z galaxies in the cylinders. The gray points present photo-z galaxies without photometric redshift uncertainties.

As noted in Section 6.2, the three cylinders, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$ depart from the anti-correlation of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ in Figure 4. Because the simulation counterparts of these three cylinders penetrate gas filaments, a large void, and orthogonal low-density filaments by chance, the comparisons with our simulations suggest that the significant departures from the anti-correlation are produced by the filamentary LSSs and the observation sightlines. These chance alignment effects reduce the anti-correlation signal.

7. Summary

We investigate spatial correlations of galaxies and IGM H I with the 13,415 photo-z galaxies and the Lyα forest absorption lines of the background quasars with no signature of damped Lyα system contamination in the 1.62 deg^2 COSMOS/UltraVISTA field. The results of our study are summarized below.

1. We estimate the Lyα forest fluctuation $\delta_{\langle F\rangle}$ and the galaxy overdensity $\delta_{\rm gal}$ within the impact parameter of 2.5 pMpc from the quasar sightlines at $z\sim$ 2–3. We identify an indication of anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values (Figure 2). The Spearman's ρ value of -0.39 indicates that there is weak evidence of an anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ at a $\sim$90% confidence level. This anti-correlation suggests that high-z galaxies are found in the excess regions of H I gas in the Lyα forest.
2. We perform cosmological hydrodynamical simulations with the RAMSES code, and identify an anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values in our simulation model that is similar to the one found in our observational data. We estimate the Spearman's ρ for the $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values in our simulation results that suggest the anti-correlation agrees with the observational results.
3. In our observational data, we identify four cosmic volumes that have very large or small values of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ that are dubbed $A_{\rm obs}$, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$ (Figure 4). Three out of these four cylinders, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$, present significant departures from the anti-correlation of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$, and weaken the signal of the anti-correlation.
4. In our simulations, we identify model counterparts of $A_{\rm obs}$, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$ in the $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ plane (Figure 4), which are referred to as $A_{\rm sim}$, $B_{\rm sim}$, $C_{\rm sim}$, and $D_{\rm sim}$, respectively. The comparisons of $A_{\rm obs}$–$D_{\rm obs}$ with $A_{\rm sim}$–$D_{\rm sim}$ indicate that the observations pinpoint (1) a galaxy overdensity like a proto-cluster in $A_{\rm obs}$, (2) gas filaments lying on the quasar sightline by chance in $B_{\rm obs}$, (3) a large void in $C_{\rm obs}$, and (4) orthogonal low-density filaments in $D_{\rm obs}$. Our simulations suggest that the three cylinders, $B_{\rm obs}$, $C_{\rm obs}$, and $D_{\rm obs}$, significantly departing from the anti-correlation, are produced by the filamentary LSSs and the observation sightlines. The chance alignment effects reduce the anti-correlation signal of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$.

The large-scale correlation of $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$ found in Section 4 is relatively weak. This is because the correlation is based on the relatively large cylinder, whose redshift range is limited in photo-z accuracy (see Figure 1 of Cai et al. 2016). The small-scale galaxy-H I relations can be studied with galaxies with spectroscopic redshifts. Here, the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX) survey (Hill & HETDEX Consortium 2016) will carry out the wide-field observations, and provide 0.8 million galaxies in $\sim$400 deg^2. The HETDEX survey will reveal the galaxy-H I relations in the large cosmological volumes including a number of proto-cluster cadidates, gas filaments, and voids of LSSs. The galaxy-IGM relation study with HETDEX will be complementary to the

programs of the MAMMOTH (Cai et al. 2016) and the CLAMATO (Lee et al. 2014, 2016).

We are grateful to Toru Misawa, Suzuka Koyamada, Toru Yamada, Ken Mawatari, Takashi Kojima, Andreas Schulze, Henry Joy McCracken, Rich Bielby, and Shun Arai for their useful comments and constructive discussions. We are also grateful to Khee-Gan Lee for kindly providing the MF-PCA code.

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium.

The COSMOS/UltraVISTA K_s-band selected galaxy catalog used in this work is compiled by Muzzin et al. (2013a). The catalog contains PSF-matched photometry in 30 photometric bands covering the wavelength range of 0.15 μm $\rightarrow$ 24 μm and includes the available *GALEX* (Martin et al. 2005), CFHT/Subaru (Capak et al. 2007), UltraVISTA (McCracken et al. 2012), S-COSMOS (Sanders et al. 2007), and zCOSMOS (Lilly et al. 2009) data sets.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III Web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

This work has made use of the services of the ESO Science Archive Facility. Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 086.A-0974.

This work is supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan, and KAKENHI (15H02064) Grant-in-Aid for Scientific Research (A) through Japan Society for the Promotion of Science (JSPS).

Appendix A
Convergence Test

We perform simulations that have a box size of 80 h^{-1} cMpc length with 256^3 cells. Figure 9 presents $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values in the cylinders. The green (red) dots represent a cylinder of $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ in our simulations with 256^3(512^3) cells. In the same manner as our simulations with 512^3 cells (Section 5.3), we calculate a Spearman's rank correlation coefficient of the 256^3-cell simulations ρ_{256}, using 16 cylinders that are randomly chosen from the 256^3-cell simulation results. We obtain

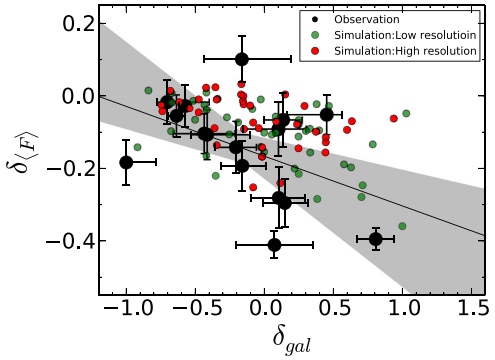

Figure 9. Same as Figure 2, but for the different resolution results with 256^3 (green dots) and 512^3 (red dots) cells.

$\rho_{256} \sim -0.4$, which corresponds to the $\sim$90% confidence level. We find that results are very similar to those of our simulations with 512^3 cells.

Appendix B
Detailed Properties of the $A_{\rm obs}$ Sightline

In this appendix, we show the supplementary VLT/X-shooter observations for the background quasar of $A_{\rm obs}$. The X-shooter observations were carried out in the service mode on 2010 December 31 (Program ID: 086.A-0974, PI: S. Lilly). We use the reduced X-shooter spectra that are publicly available on the European Southern Observatory (ESO) Science Archive Facility.[13] We select the spectra of the UVB and VIS arms, which cover the wavelength ranges of 3000–5595 Å and 5595–10240 Å, respectively. The spectral resolutions of the UVB and VIS arms are medium high, $R \sim 4000$ and $R \sim 7400$, respectively. The observational details are summarized in Table 1.

Figure 10 shows the X-shooter spectrum of the background quasar of $A_{\rm obs}$ in the same wavelength range as Figure 5. Although the spectral resolution of the X-shooter spectrum is significantly higher than that of the BOSS spectrum (Figure 5), we confirm that the sightline of $A_{\rm obs}$ has Lyα forest absorption lines with no signature of DLAs. The only exception is the absorbers at $z \sim 2.203$, where saturated Lyα absorption lines are detected. The top left panel of Figure 11 presents a zoom-in X-shooter spectrum around the saturated Lyα absorption lines. Note that their Lyβ absorption lines are also covered by the X-shooter spectrum. However, as shown in the bottom left panel of Figure 11, the Lyβ absorption lines are also saturated. Moreover, the Lyβ absorption lines are contaminated by foreground Lyα absorbers at $z \sim 1.70$, which makes it difficult to precisely measure HI column densities of the HI absorption lines.

As shown in the right panel of Figure 11, we identify metal absorption lines of C II, Si IV, C IV and Mg II at $z \sim 2.203$ in the X-shooter spectrum. There are three metal absorptions dubbed Systems 1, 2 and 3 that are labeled in Figure 11. We fit Voigt profiles to these metal lines with VPFIT[14] to measure column densities. The best-fit profiles for the metal lines are presented in the right panel of Figure 11. Table 2 summarizes the measured column densities.

[13] http://archive.eso.org
[14] http://www.ast.cam.ac.uk/~rfc/vpfit.html

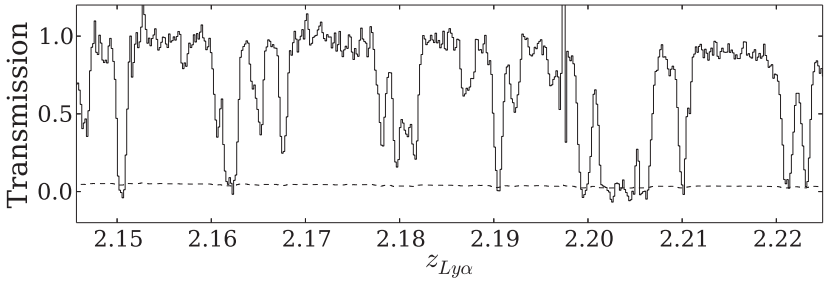

Figure 10. VLT/X-shooter spectrum of the background quasar of $A_{\rm obs}$. The solid lines represent transmission per pixel as a function of redshift for the Lyα forest absorption lines. The dashed curve denotes the noise per pixel.

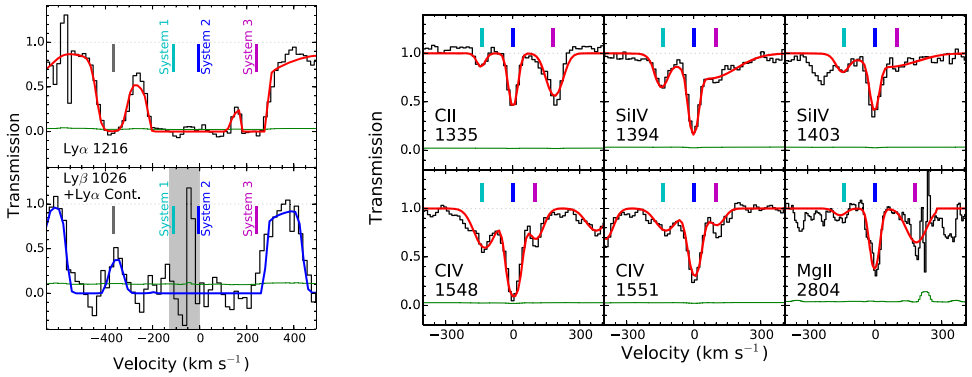

Figure 11. Left panel: zoom-in X-shooter spectrum around the $z \sim 2.203$ H I Lyα absorption systems (top) and Lyβ absorption systems with $z \sim 1.70$ Lyα absorption contaminations (bottom). The black (green) histogram presents the transmission (1σ noise) per pixel. The Voigt-profile decompositions (red curve) are made of the three Lyα absorption systems, System 1 (cyan ticks), System 2 (blue ticks), System 3 (magenta ticks), and a Lyα absorption system that has no metal lines (gray ticks). The blue curve shows the fitting results for the Lyβ absorption systems and $z \sim 1.70$ Lyα absorption contaminations. Note that we simultaneously fit $z \sim 2.203$ Lyα and Lyβ absorbers and $z \sim 1.70$ Lyα absorption contaminations. The decompositions indicate that System 2 is the LLS at $z = 2.203$. The gray shaded region in the bottom panels presents the wavelength range where we find a large sky subtraction systematic errors. Right panel: same as the left panel, but for metal absorption lines associated with the Lyα absorption systems at $z \sim 2.203$.

Table 1
X-shooter Observations

Source	R.A. (J2000)	Decl. (J2000)	Integration Time (s)	Dates of Observations	S/N[a] [pix^{-1}]
COSMOS-QSO-199[b]	09 58 58.72	+02 01 38.6	UVB: 2700[c]	31 Dec 2010	27
COSMOS-QSO-199[b]	09 58 58.72	+02 01 38.6	VIS: 2700[c]	31 Dec 2010	20

Notes.
[a] Median S/N per pixel for the combined spectra.
[b] The background quasar of $A_{\rm obs}$, [VV2006] J095858.7+020138 at $z = 2.448$.
[c] Three individual exposures of 900 s.

Table 2
Column Densities of the Ions in the $z \sim 2.203$ Absorption Systems

Ion	C II	Si IV	C IV	Mg II
System 1 log N (cm^{-2})	13.19 ± 0.70	13.14 ± 0.57	13.86 ± 0.19	12.31 ± 0.68
System 2 log N (cm^{-2})	13.92 ± 0.17	13.58 ± 0.24	14.29 ± 0.12	13.31 ± 0.21
System 3 log N (cm^{-2})	14.03 ± 0.16	13.54 ± 0.35	13.59 ± 0.35	13.29 ± 0.13

To estimate HI column densities of Systems 1–3, we first make photoionization models with the input observational measurements of C II, Si IV, C IV, and Mg II with those predicted from photoionization models. We perform multi-phase photoionization calculations with version 13.04 of the CLOUDY software (Ferland et al. 2013). We conduct the CLOUDY modeling for high-ionization phase clouds (CIV, SiIV) and low-ionization phase clouds (Mg II, C II; e.g.,

Misawa et al. 2008). We model these clouds in each phase, assuming a gas slab exposed by a uniform ultraviolet background (Haardt & Madau 2012) with a range of ionization parameters $U \equiv \Phi/cn_{\rm H}$ $(-4.2 < \log U < -0.6)$, where $n_{\rm H}$ and Φ are the hydrogen volume density and the ionizing photon flux incident on the gas cloud, respectively. The solar relative abundances of Asplund et al. (2009) are assumed. We then search for the best-fit model that minimizes χ^2 between the measured metal column densities and the photoionization model predictions. We find that the best-fit models for Systems 1, 2, and 3 have HI column densities of $\log N_{\rm H_I}$ (cm^{-2}) $\sim$ 16.0, 19.0, and 16.5, respectively. We then fit Voigt profiles to the spectrum in the wavelength ranges of the Lyα absorption lines (top left panel of Figure 11) and Lyβ absorption lines (bottom left panel of Figure 11) by using the column densities of the CLOUDY model results. Because the spectrum in the Lyβ wavelength range is contaminated by the foreground Lyα absorbers at $z \sim 1.70$, we conduct simultaneous fitting to the spectrum in these two wavelength ranges with the $z \sim 2.203$ Lyα and Lyβ absorbers, and $z \sim 1.70$ contaminations, together with the other Lyα absorbers. We obtain a self-consistent model that is shown with the red and blue curves in the left panels of Figure 11.

Based on the CLOUDY model results, we find that System 2 is classified as a Lyman limit system (LLS), which is an optically thick clouds with an HI column density of $17.2 \leqslant \log N_{\rm H_I}$ (cm^{-2}) $\leqslant 20.3$. Note that the presence of this LLS does not change our conclusions. We confirm that the weak anti-correlation between $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ is found at the $\sim$80% confidence level, even if the LLS is masked out in the BOSS spectrum. Our CLOUDY model indicates that System 2 has a metallicity $Z/Z_\odot \simeq 0.02$ and an ionization parameter $\log U \simeq -3.0$ that are comparable with those of typical LLSs at $z \sim 2$ (Fumagalli et al. 2016, 2013).

Because Systems 1 and 3 have HI column densities of $\log N_{\rm H_I}$ (cm^{-2}) $\sim$ 16.0 and 16.5, Systems 1 and 3 are classified as Lyα forest absorbers based on the moderately low column densities. However, the CLOUDY models imply that their metallicities are $Z/Z_\odot \simeq 0.3$ (System 1) and 1.0 (System 3), which are two orders of magnitude higher than the median IGM metallicity (Simcoe 2011). These results would suggest that Systems 1 and 3 are gas clumps in the CGM and/or the intra-cluster medium of the proto-cluster candidate discussed in Section 6.2.

Appendix C
Counterparts for Each Extreme Cylinder

In addition to $A_{\rm sim}$–$D_{\rm sim}$, there exist counterparts of $A_{\rm obs}$–$D_{\rm obs}$ in our simulations. We find additional two counterparts for each extreme cylinder, $B_{\rm obs}$, $C_{\rm obs}$, or $D_{\rm obs}$, in different volumes of the simulations, which are referred to as ($B1_{\rm sim}$, $B2_{\rm sim}$), ($C1_{\rm sim}$, $C2_{\rm sim}$), or ($D1_{\rm sim}$, $D2_{\rm sim}$). Moreover, we identify one additional counterpart of $A_{\rm obs}$, $A1_{\rm sim}$. We cannot find another counterpart of $A_{\rm obs}$, because the large $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values of $A_{\rm obs}$ are very rare in the simulation box. Figure 4 presents the $\delta_{\rm gal}$ and $\delta_{\langle F\rangle}$ values of these additional counterparts with the blue crosses. These $\delta_{\rm gal}$–$\delta_{\langle F\rangle}$ values are comparable with those of $A_{\rm sim}$–$D_{\rm sim}$ at the $\sim 1\sigma$ error levels. Figure 12 shows that these sightlines penetrate large overdensities, gas filaments parallel with (orthogonal to) the sightline, or large voids.

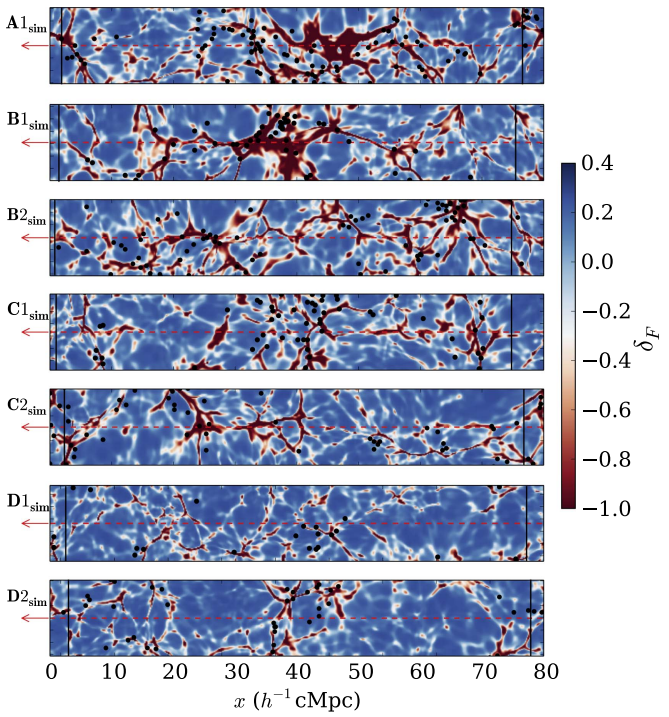

Figure 12. Same as Figure 7, but for $A1_{\rm sim}$, $B1_{\rm sim}$, $B2_{\rm sim}$, $C1_{\rm sim}$, $C2_{\rm sim}$, $D1_{\rm sim}$, and $D2_{\rm sim}$.

References

Adelberger, K. L., Shapley, A. E., Steidel, C. C., et al. 2005, ApJ, 629 636
Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, ApJ, 584, 45
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Becker, G. D., Bolton, J. S., Haehnelt, M. G., & Sargent, W. L. W. 2011, MNRAS, 410, 1096
Becker, G. D., Bolton, J. S., & Lidz, A. 2015, PASA, 32, e045
Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013, MNRAS, 430, 2067
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bertschinger, E. 1995, arXiv:astro-ph/9506070
Boera, E., Murphy, M. T., Becker, G. D., & Bolton, J. S. 2014, MNRAS, 441, 1916
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Cai, Z., Fan, X., Peirani, S., et al. 2016, ApJ, 833, 135
Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99
Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, ApJL, 782, L3
Chiang, Y.-K., Overzier, R. A., Gebhardt, K., et al. 2015, ApJ, 808, 37
Crighton, N. H. M., Bielby, R., Shanks, T., et al. 2011, MNRAS, 414, 28
Cucciati, O., Zamorani, G., Lemaux, B. C., et al. 2014, A&A, 570, A16
Curran, P. A. 2015, arXiv:1411.3816
Eisenstein, D. J., & Hut, P. 1998, ApJ, 498, 137
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Faucher-Giguère, C.-A., Prochaska, J. X., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, ApJ, 681, 831
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2016, MNRAS, 455, 4100
Fumagalli, M., O'Meara, J. M., Prochaska, J. X., & Worseck, G. 2013, ApJ, 775, 78
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Harikane, Y., Ouchi, M., Ono, Y., et al. 2016, ApJ, 821, 123
Hill, G. J. & HETDEX Consortium 2016, in ASP Conf. Ser. 507, Multi-Object Spectroscopy in the Next Decade: Big Questions, Large Surveys, and Wide Fields, ed. I. Skillen, M. Barcells, & S. Trager (San Francisco, CA: ASP), 393

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hui, L., & Gnedin, N. Y. 1997, MNRAS, 292, 27
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Lee, K.-G., Bailey, S., Bartsch, L. E., et al. 2013, AJ, 145, 69
Lee, K.-G., Hennawi, J. F., Stark, C., et al. 2014, ApJL, 795, L12
Lee, K.-G., Hennawi, J. F., White, M., et al. 2016, ApJ, 817, 160
Lee, K.-G., Suzuki, N., & Spergel, D. N. 2012, AJ, 143, 51
Lee, K.-S., Giavalisco, M., Conroy, C., et al. 2009, ApJ, 695, 368
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218
Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, MNRAS, 446, 3697
Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005, ApJL, 619, L1
Mawatari, K., Inoue, A. K., Kousai, K., et al. 2016, ApJ, 817, 161
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
McDonald, P., Seljak, U., Burles, S., et al. 2006, ApJS, 163, 80
Meiksin, A., Bolton, J. S., & Tittley, E. R. 2015, MNRAS, 453, 899
Meiksin, A. A. 2009, RvMP, 81, 1405
Misawa, T., Charlton, J. C., & Narayanan, A. 2008, ApJ, 679, 220
Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (Cambridge: Cambridge Univ. Press)
Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, ApJS, 206, 8
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, ApJ, 777, 18
Noterdaeme, P., Petitjean, P., Carithers, W. C., et al. 2012, A&A, 547, L1
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Palanque-Delabrouille, N., Yèche, C., Borde, A., et al. 2013, A&A, 559, A85
Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144
Prochaska, J. X., Hennawi, J. F., Lee, K.-G., et al. 2013, ApJ, 776, 136
Rakic, O., Schaye, J., Steidel, C. C., & Rudie, G. C. 2012, ApJ, 751, 94
Reed, D. S., Bower, R., Frenk, C. S., Jenkins, A., & Theuns, T. 2007, MNRAS, 374, 2
Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, ApJ, 750, 67
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Schechter, P. 1976, ApJ, 203, 297
Scoville, N., Arnouts, S., Aussel, H., et al. 2013, ApJS, 206, 3
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Simcoe, R. A. 2011, ApJ, 738, 159
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9
Tejos, N., Morris, S. L., Finn, C. W., et al. 2014, MNRAS, 437, 2017
Teyssier, R. 2002, A&A, 385, 337
Theuns, T., Mo, H. J., & Schaye, J. 2001, MNRAS, 321, 450
Thomas, R., Le Fèvre, O., Cassata, V. L. B. P., et al. 2014, arXiv:1411.5692
Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, ApJ, 783, 85
Tummuangpak, P., Bielby, R. M., Shanks, T., et al. 2014, MNRAS, 442, 2094
Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014, MNRAS, 445, 794
Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189
Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, A&A, 536, A105
Wall, J. V., & Jenkins, C. R. 2012, Practical Statistics for Astronomers (Cambridge: Cambridge Univ. Press)
Weinberg, D. H., Burles, S., Croft, R. A. C., et al. 1998, arXiv:astro-ph/9810142
Weinberg, D. H., Davé, R., Katz, N., & Kollmeier, J. A. 2003, in AIP Conf. Ser. 666, The Emergence of Cosmic Structure, ed. S. H. Holt & C. S. Reynolds (Melville, NY: AIP), 157
White, M., Pope, A., Carlson, J., et al. 2010, ApJ, 713, 383