

Joint analysis of galaxy redshift surveys and Ly α forests I: the large-scale gasenous environment of Ly α galaxies with redshift-space distortion

Koki Kakiichi et al.^{1*}

¹Max Planck Institute for Astrophysics, Karl-Schwarzschild Straße 1, 85741 Garching, Germany

Released 2002 XXXXX XX

ABSTRACT

Redshift surveys of Ly α forests and Ly α -emitting/absorbing galaxies.

Key words: atomic processes – cosmology: theory – line: formation – radiative transfer – infrared: general – scattering

1 INTRODUCTION

We propose and the explore the potentials of deep redshift surveys of Ly α galaxies in the foreground of QSOs covering redshifts $2 < z < 7$. Joint analysis of Ly α forests and Ly α -emitting and absorbing galaxies over redshift $2 < z < 7$ enables us to investigate the Epoch of Reionization and the transition to post-reionization era, allowing to break the multiple degeneracies arising from analysis of high-redshift Ly α galaxies alone. For both Lyman-break and narrow-band selection technique, Ly α line can appear in galaxies both as emission or absorption.

The gasenous environment of galaxies is a natural consequence arising from the concordance hierarchical structure formation in ΛCDM cosmology, forming cosmic web. One effect of gasenous environment on galaxies is the impact on their formation and evolution (Dekel+), for example, the popular cold stream picture in the formation of galaxies to name a few. Another important impact of the environment is to induce the selection bias on the observation of galaxies. Since we see galaxies through their gasenous environments, any absorption or reprocessing of the light along a line of sight introduces the environmentally-induced selection bias upon observations of galaxies. One notable example of such selection bias is on the population of Ly α -emitting galaxies, which has been used to constrain the neutral fraction of the universe during the EoR. Furthermore, even at the post-reionized universe, the possible impact on LAEs clustering due to the Ly α RT through the gasenous environment as CGM and IGM (Zheng+; Wyithe & Dijkstra; Breg+) calls for the close examination and testing this hypothesis based on observations. The degree of the environmetal impact on the LAE selection influences the interpretation of redshift BAO survey such as HETDEX to do cosmology (Wyithe&Dijkstra, Breg+).

The galaxy redshift survey in the foreground of QSOs is not a new idea, which has been conducted by Aldelberger+, Cooke+, Keck Baryonic Structure Survey (KBSS, Steidel+) for $z \sim 2 - 3$. Yet, the connection to EoR and RT studies of galaxies and the IGM are not or only weakly made. The idea to perform the joint analysis of Ly α galaxies and QSO spectra occasionally has appeared in the literature (Baek, Ferrara & Semelin 2012, SDSS J1335+3533 good candidate?). Yet, the full potential of such survey strategy and the requirement still must be worked out. An idea is to extend such survey strategy to higher redshifts, which enable us the joint analysis of LLS/DLA, Ly α forests and galaxies for both Ly α in emission and absorption.

2 THEORY: JOINT MODEL OF REDSHIFT SPACE DISTORTION AND Ly α RT

Our primary motive is to study the intergalactic environment of Ly α -galaxies and its impact on Ly α RT. The Ly α RT effect on the Ly α -galaxy depends on the clustering and pairwise velocity statistics between galaxies and small-scale absorbers. We consider how to constrain the pairwise velocity staitistics from redshift-space distortion. To build the intuition, we first employ the analytic theory to connect the statistical formulations of the redshift-space anisotropy of galaxy-absorber systems and the cosmological Ly α radiative transfer. The formalism for the redshift-space distortion largely follows that by Fisher et al. (1994) and Reid & White (2011). The joint model will be used as an estimator to infer the pairwise velocity distribution and it impact on the Ly α visibility.

2.1 Dynamics and statistics: galaxy-absorber pairs

The dynamics of galaxy-absorber system and the real-space clustering form the basis of joint statistical formulation of

* E-mail: kakiichi@mpa-garching.mpg.de

redshift-space anisotropy and Ly α RT. The dynamical description of the galaxy-absorber pairs is provided by the phase-space information. The phase space distribution function of a pair is $f(v_{12}, r_{12})dr_{12}dv_{12}$, i.e. the number of galaxy-absorber pair in the interval from r_{12} to $r_{12}+dr_{12}$ and v_{12} to $v_{12}+dv_{12}$, where v_{12} is the pairwise relative velocity between galaxy and absorber and r_{12} is the galaxy-absorber comoving distance. Because of statistical isotropy and homogeneity, it only depends on the magnitude of pairwise velocity and comoving separation.

The real-space correlation function $\bar{n}^2\xi(r) = \int f(v, r)dv$ (check).

2.1.1 real-space clustering

The global abundance of absorber is characterized by the column density distribution function $\frac{\partial N_{\text{HI}}}{\partial N_{\text{HI}}\partial z}$. In the following treatment, it is more convenient to consider the CDDF per comoving distance $|\frac{dz}{dr}| \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z}$ where $|\frac{dz}{dr}| = c/H_0[\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}$. Recalling the definition of correlation function, the probability find an absorber along a one-dimensional skewer from galaxy at comoving distance r is given by

$$\begin{aligned} P(r)dr &= \int dN_{\text{HI}} \left| \frac{dz}{dr} \right| \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} [1 + \xi(r, N_{\text{HI}})] dr, \\ &= \lambda_a [1 + \bar{\xi}(r)] dr, \end{aligned} \quad (1)$$

where the prefactor $\lambda_a(z) = \int dN_{\text{HI}} \left| \frac{dz}{dr} \right| \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z}$ is the line density of absorbers along a skewer. For simplicity we consider the CDDF-integrated correlation function

$$\bar{\xi}(r) = \int dN_{\text{HI}} \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} \xi(r, N_{\text{HI}}) / \int dN_{\text{HI}} \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z}. \quad (2)$$

In the linear theory, $\bar{\xi}(r) = b_a b_g \int \Delta_L^2(k) j_0(kr) d\ln k$ where $j_0(kr)$ is the zeroth spherical Bessel function of the first kind. Another useful form is the power-law,

$$\bar{\xi}(r) = (r/r_0)^{-\gamma} \quad (3)$$

2.1.2 What is the effect of local ionizing source?

The local ionizing source from galaxy itself contribute to the photoionization rate Γ

$$\Gamma_{\text{local}} = \int \frac{4\pi J_\nu \sigma_{\text{HI}}}{h\nu} d\nu \approx \frac{\alpha}{3+\alpha} \frac{\sigma_{912} f_{\text{esc}}^{\text{LyC}} \dot{N}_{\text{ion}}}{4\pi r^2} \quad (4)$$

where $\sigma_{\text{HI}} = \sigma_{912}(\nu/\nu_{912})^{-3}$, $\sigma_{912} = 6.304 \times 10^{-18} \text{ cm}^2$, ν_{912} is the Lyman limit frequency, $f_{\text{esc}}^{\text{LyC}}$ is the escape fraction of Lyman continuum photons, and $\dot{N}_{\text{ion}} = \int_{\nu_{912}}^{\infty} \frac{L_\nu d\nu}{h\nu}$ is the ionizing photon emissivity. We have assumed the power-law spectrum for galaxy $L_\nu \propto \nu^{-\alpha}$.

Using the stellar population synthesis result from Schaefer (2003), the ionizing photon emissivity of the high-redshift star-forming galaxies is given by $\dot{N}_{\text{ion}}/(\text{ph s}^{-1}) = 10^{53.81 - 0.0029(\log_{10} Z + 9)^{2.5}} \text{ SFR}/(\text{M}_\odot \text{ yr}^{-1})$ for $10^{-9} \leq Z \leq 0.04$ with Salpeter IMF. The variation of metallicity in this range produces the difference within factor of 3. Since the metallicity is poorly constrained parameter with SED fitting analysis (de Barros et al. 2014), we take $Z = 0.02$ as our fiducial value. However note that extremely metal poor

galaxy $Z = 10^{-9}$ can exhibit the factor of 2.6 larger ionizing photon emissivity than the fiducial value. The Lyman-break selected and Ly α -selected galaxies show the diverse range of star formation rate ranging from $\text{SFR} \sim 1\text{M}_\odot \text{yr}^{-1}$ to $\text{SFR} \sim 100\text{M}_\odot \text{yr}^{-1}$ or as high as $\text{SFR} \sim 1000\text{M}_\odot \text{yr}^{-1}$ (Scheerer; Ouchi; Rhood; Hagen etc.).

The local contribution enhances the total photoionization rate by $\Gamma = \Gamma_{\text{local}} + \Gamma_{\text{global}}$. We can define the characteristic equality radius r_{eq} such that the local contribution to the photoionization rate is equal to the global contribution $\Gamma_{\text{local}}(r_{eq}) = \Gamma_{\text{global}}$,

$$\begin{aligned} r_{eq} &= \sqrt{\frac{\alpha}{3+\alpha} \frac{\sigma_{912} f_{\text{esc}}^{\text{LyC}} \dot{N}_{\text{ion}}}{4\pi \Gamma_{\text{global}}}} \\ &\approx 320 h_7^{-1} \text{ ckpc} \left(\frac{1+z}{4} \right) \sqrt{\frac{\alpha}{3+\alpha}} \left(\frac{f_{\text{esc}}^{\text{LyC}} \text{SFR}}{\text{M}_\odot \text{yr}^{-1}} \right)^{1/2} \end{aligned} \quad (5)$$

Then, $\Gamma_{\text{local}} = \Gamma_{\text{global}}(r/r_{eq})^{-2}$. The combination of $f_{\text{esc}}^{\text{LyC}} \text{SFR}$ enters, for example, the fiducial values of $f_{\text{esc}}^{\text{LyC}} = 0.1$ and $\text{SFR} = 10\text{M}_\odot \text{yr}^{-1}$.

By taking the Jeans argument by Schaye (2001), H I column density of an absorber scales as $N_{\text{HI}} \propto \Gamma^{-1}$,

$$\begin{aligned} N_{\text{HI}} &\sim 2.7 \times 10^{13} \text{ cm}^2 (1+\delta)^{3/2} T_4^{-0.26} \Gamma_{12}^{-1} \\ &\times \left(\frac{1+z}{4} \right)^{9/2} \left(\frac{\Omega_b h^2}{0.02} \right)^{3/2} \left(\frac{f_g}{0.16} \right)^{1/2}. \end{aligned} \quad (6)$$

Following Hui et al. (1997), rescaling the H I column density from a reference value $N_{\text{HI}}^{\text{ref}}$ corresponding to Γ^{ref} into a new value N_{HI} with Γ , $N_{\text{HI}} = (\Gamma^{\text{ref}}/\Gamma)N_{\text{HI}}^{\text{ref}}$, the CDDF scales as

$$\frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} \Big|_{N_{\text{HI}}} = \frac{\Gamma}{\Gamma^{\text{ref}}} \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}^{\text{ref}}\partial z} \Big|_{N_{\text{HI}}^{\text{ref}} = (\Gamma/\Gamma^{\text{ref}})N_{\text{HI}}} \quad (7)$$

In the case of the CDDF with the power-law form ¹

$$\frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} = A(z) (N_{\text{HI}}^{\text{ref}})^{-\beta_N} \quad (8)$$

The reference value Γ^{ref} is taken as the measured photoionization rate consistent with the derived CDDF at redshift of the measurement. We take $\Gamma^{\text{ref}} = 10^{-12} \text{ s}^{-1}$. Then, the dependence of the CDDF on the photoionization rate is given by

$$\frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} = A(z) \left(\frac{\Gamma}{\Gamma^{\text{ref}}} \right)^{-\beta_N+1} N_{\text{HI}}^{-\beta_N} \quad (9)$$

For $\beta_N > 1$, the normalization of the CDDF is suppressed by the increasing photoionization rate. The suppression factor $S(r)$ by the change in the photoionization rate is

$$S(r) = \left(\frac{\Gamma}{\Gamma^{\text{ref}}} \right)^{-\beta_N+1} = \left[\frac{\Gamma_{\text{global}}}{\Gamma^{\text{ref}}} \left(\left(\frac{r}{r_{eq}} \right)^{-2} + 1 \right) \right]^{-\beta_N+1} \quad (10)$$

Thus by substituting the suppression factor,

$$\begin{aligned} P(r) &= \int dN_{\text{HI}} \left| \frac{dz}{dr} \right| \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}}\partial z} S(r) [1 + \xi(r, N_{\text{HI}})] dr, \\ &= \lambda_a S(r) [1 + \bar{\xi}(r)], \end{aligned} \quad (11)$$

¹ In the literature the prefactor is often parameterised as $A(z) = A(1+z)^{\beta_z}$ where A is a constant.

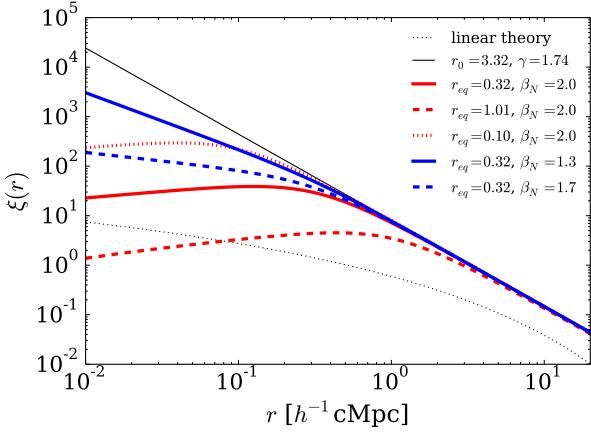


Figure 1. The galaxy-absorber correlation function. The black line corresponds to the best-fit power-law LBG-DLA correlation function at $z \sim 3$ by Cooke+ (2006). The red and blue lines modify the Cooke+ (2006) best-fit by inclusion of local photoionization effect. The red lines shows the variation over the equality radius r_{eq} and the blue lines shows the variation over the CDDF slope β_N . The turn-over of correlation function at small-scale only occurs when $\beta_N > \gamma/2 + 1$ is satisfied. The equality radius is in the unit of $h^{-1} \text{cMpc}$.

It is interesting to see the slope of $P(r)$ in the limit of $r \rightarrow 0$ ($r < r_{eq}$ and $r < r_0$) for the power-law correlation function, which approaches to $P(r) \propto r^{2(\beta_N - 1) - \gamma}$. In order that $P(r)$ to turn around and produce the positive slope, we need $\beta_N > \gamma/2 + 1$.

2.1.3 The effect of reionization (or UV background fluctuation)

Suppose a Ly α forest-galaxy survey deeper towards the EoR. Toward the EoR, we expect some diffuse H $_1$ patches before the IGM engulfed by the I-front. These H $_1$ patches in the QSO absorption spectra appears as DLA-like feature (or LLS depending on H $_1$ column density). However, in contrast to the ordinary LLS/DLAs around galaxies, the spatial position is expected to be greater than the typical size of H $_{\text{II}}$ bubbles $r > R_b$. Therefore, in addition to post-reionized universe galaxy-absorber distribution based on cosmic web, the correlation function on the scale $r > R_b$ will have extra contribution by the EoR-induced absorbers. (Just to illustrate I plotted Fig. 2. simplified version of McQuinn+ 2005 model).

While in practice we have only one $z \sim 7$ QSO, the availability of handful of $z \sim 6$ QSOs might allow us to measure this possible effect towards the end of reionization. The existence of H $_1$ patch with increasing correlation with galaxy on the scale $r > R_b$ will be a necessary diagnostic criterion that we are approaching to the EoR at $z > 6$. However, it is worthwhile to note the this may not be a sufficient condition because the UV background fluctuation can produce additional correlation on the large-scale. Pontzen+ (2014a,b) and Gontcho+ (2014) have considered the effect of UV background fluctuation on Ly α forest correlation function and showed the increase of correlation at large scale. Since the

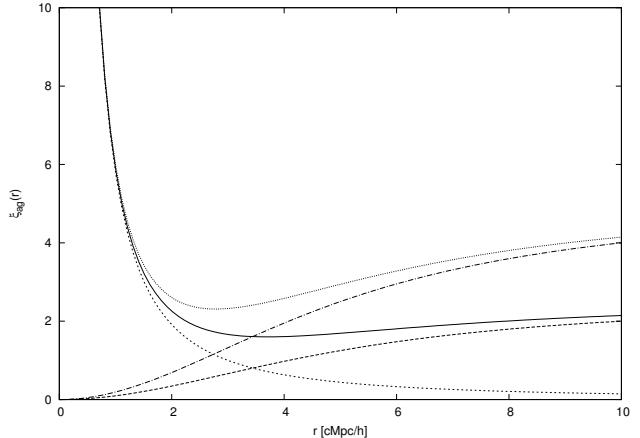


Figure 2. galaxy-absorber correlation function. Illustrating example effect of EoR. quick calc. just not to forget. hyper-simplified model based on McQuinn+ 2005 & Furlanetto+ 2004. **do this using simulation, then take a quick functional form here as a fitting function.**

mean free path decrease to higher redshift, the power can be added even on the scale compatible to the H $_{\text{II}}$ bubble size.

The existence of spikes on $z \sim 6$ QSO absorption spectra provides the interesting opportunity to apply galaxy survey in the $z \sim 6$ QSO fields.

2.1.4 Pairwise velocity statistics

The conditional pairwise velocity distribution function $f(v_{12}|r_{12})$ at the fixed comoving separation r_{12} is followed from the rule of probability $f(v_{12}, r_{12}) = f(v_{12}|r_{12})f(r_{12})$. The mean pairwise velocity is given by the first moment, $\langle v_{12}(r) \rangle = \int v_{12}f(v_{12}|r)dv_{12}$. The pairwise velocity statistics is the central quantity in our formulation.

The streaming models require the modelling of the pairwise velocity statistics; mean pairwise velocity $v_{12}(r)$ and pairwise velocity dispersion $\sigma_{12}^2(r, \mu)$. The BBGKY hierarchy provides a framework to relate the pairwise velocity statistics to the real-space clustering (Davis&Peebles 1977; Peebles 1980; Fry+; Hamilton 1988). Formally, the BBGKY hierarchy provides the relation between velocity statistics such as pairwise streaming velocity and velocity dispersion and N-point correlation function. We cast the traditional BBGKY hierarchy expressed in terms of dark matter or galaxies in the application to galaxy-absorber systems. The absorbers are the result of the IGM fluctuation in the cosmic web environment. We approximate the absorbers as a cloud. However in reality, Ly α forest or self-shielded system, LLS/DLA, should be viewed in terms of cosmic web. We mitigate this by considering the filaments or walls of cosmic web as connection of clouds in a appropriate geometry.

Suppose the phase-space distribution of galaxy and absorbers. The first velocity moment of the BBGKY hierarchy expresses the the conservation of pairs. This can more easily derived from the conservation argument (Peebles 1976?; White's book). Suppose particles (galaxy-absorber pair) within a comoving radius $< r$ centred at a particle, $N(< r, t) = \bar{n}(0) \int_0^r 4\pi r'^2 dr' (1 + \xi(r, t))$ where $\bar{n}(0)$ is the

mean comoving number density (constant over redshift) and $\xi(r, t)$ is the particle-particle (galaxy-absorber) correlation function. The change in the total particle number $N(< r, t)$ is determined by the flux through the surface of sphere of radius r , which is $4\pi r^2 \bar{n}(0)(1 + \xi(r, t))\langle v_{12} \rangle$. Note that $\langle v_{12} \rangle$ is the mean pairwise *comoving* velocity.² We adopt the convention that the pairwise velocity is negative for inflow and positive for outflow.

Thus, when the total number of particles are conserved,

$$\frac{\partial N}{\partial t} + 4\pi r^2 \bar{n}(0)(1 + \xi(r, t))\langle v_{12} \rangle = 0 \quad (14)$$

By rearranging for $\langle v_{12} \rangle$, this follows the mean pairwise velocity field as (Davis&Peebles 1977; Cooray & Sheth 2002; White)

$$\langle v_{12}(r, t) \rangle = -\frac{1}{1 + \xi(r, t)} \frac{1}{r^2} \frac{\partial}{\partial t} \int_0^r \xi(r', t) r'^2 dr' \quad (15)$$

For linear theory, the correlation function scales as $\propto D(t)^2$, the pairwise velocity field for unbiased tracer from the linear theory.

$$v_{12}(r) = -\frac{Hf(\Omega_m)}{\pi^2} \int P_L(k) j_1(kr) k dk \quad (16)$$

where $f(\Omega) = d \ln D / d \ln a$.

For self-similar clustering,

$$v_{12}(r) = -\frac{v_{in}}{1 + (r/r_0)^\gamma} \quad (17)$$

2.1.5 Pairwise velocity dispersion

The pairwise velocity dispersion is given by the second moment of BBGKY hierarchy, which depends on three-point correlation function. The pairwise velocity dispersion has the contribution from (1) the anisotropy of the absorber's velocity onto each central galaxies (2) The dispersion among different galaxies' environment. The former is, for example, the absorber accretion from filaments is expected to have different velocity from the absorbers in the direction of voids. The latter is caused by the fact that, even at the same radial pair separation, the absorbers at different central galaxies has the different pairwise velocity as they reside in the different part of cosmic web.

The magnitude of this pairwise velocity dispersion is shown in Fig.3. The pairwise velocity dispersion spans from about 100 to 250km/s. The velocity dispersion of galaxy-absorber pairs increase towards lower redshift. This is a natural expectation from hierarchical structure formation.

While the simulation indicates the scale-dependence of pairwise velocity dispersion, for simplicity, we consider a constant pairwise velocity dispersion to construct the joint model of redshift-space distortion and Ly α RT. Such assumption may introduce a bias on the inferred properties of

² To avoid a confusion in various form appear in the literature, proper velocity is $\langle v_{12}^{proper} \rangle = a \langle v_{12}^{comoving} \rangle$.

$$\frac{\langle v_{12}^{proper}(r, a) \rangle}{Har} = -\frac{1}{1 + \xi(r, a)} \frac{a}{r^3} \frac{\partial}{\partial a} \int_0^r \xi(r', a) r'^2 dr' \quad (12)$$

in the proper coordinates??,

$$\frac{\partial \xi}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} [r^2(1 + \xi) \langle v_{12}(r, a) \rangle] = 0 \quad (13)$$

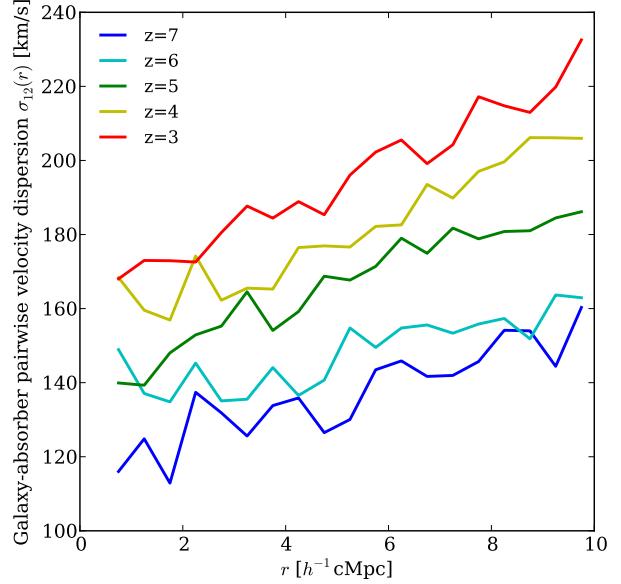


Figure 3. Radial pairwise velocity dispersion from simulation.

pairwise velocity field once the model is fitted to observations. These complications should be addressed by the mock observation of simulations with expected observational errors. In this paper, we would like to convey the framework that how the RSD of galaxy-absorber pairs helps to understand Ly α RT on the galaxy's visibility. This would aid to break the degeneracy in the H α measurement in the EoR and Ly α RT effect in cosmology with Ly α -emitting galaxies.

2.1.6 Effect of feedback and outflow

The feedback from galaxies is expected to modify the pairwise velocity statistics. So far we have only considered the inflow which is purely cosmological origin. What is the region of influence of feedbacks and the outflow velocity?

From the observations of Steidel+ and Ouchi+, the velocity shift of metal lines range from $v_{out} = 100 - 600$ km/s, which is believed to be launched by some feedback mechanism. For the back-of-envelop estimate of the region of influence of feedback, we consider a simple mechanical kick of gas with velocity v_{out} from a galaxy. To get a rough upper estimate of the region of influence of feedback, ignore the gravitational force onto central galaxy. Then, during the age of a galaxy, say $t_{age} = 1$ Gyr, the inertial motion of gas can reach to $v_{out} t_{age} \sim 100 - 610$ pkpc ($r_{infl} \sim 70 - 430(1+z)h_{70}^{-1}$ ckpc), which is about the conventional definition of CGM (<300pkpc). The hydrodynamical simulation indicates the IGM properties beyond the turnaround radius are well converged among different hydro scheme (SPH or grid) and implemented feedback mechanisms (is there papers by Hopkins?, Schaye?, Fumagalli?; Meiksin+2014).

Motivated by this argument, we consider a phenomenological model to include the outflow,

$$v_{12}(r) = 2v_{out} e^{-r/r_{infl}} - \frac{v_{in}}{1 + (r/r_0)^\gamma} \quad (18)$$

As a remark, any effect beyond the region of influence by ‘mechanical’ feedback would only be possible by radiation field from galaxies. To influence the pairwise velocity of galaxy-absorber pairs we can postulate either by energy-driven (radiative heating) or momentum-driven (radiation pressure). Note that the effect can be realized by the change of absorber abundance by the increased photoionization rate by the local source as discussed in previous section. Possible candidates for radiation pressure exerted on absorbers is by Ly α line, metal lines (for metal-enriched absorber) and dust (for dusty absorber). If these mechanism is in action, it is possible to have a larger region of influence by radiation feedback.

2.2 Redshift-space distortion

The redshift-space distortion is the result of the line-of-sight peculiar velocity between the tracers. For our interest, the tracers are galaxies (LAEs and LBGs) and absorbers (Ly α forests and LLS/DLAs). The pairwise velocity between galaxies and absorbers imprint the redshift space distortion signature. Turner? Rudie+? have measured the redshift space distortion consistent with coherent cosmological inflow for LBG samples.

The mapping between redshift space and real space is

$$s_{\perp} = r_{\perp}, \quad s_{\parallel} = r_{\parallel} + v_{12,\parallel}/H, \quad (19)$$

The redshift-space correlation function $\xi_s(s_{\parallel}, s_{\perp})$ can be modelled by the integration of real-space correlation function $\xi(r)$ with the pairwise velocity distribution function $f(v_{12,\parallel}|r)$ along a line-of-sight (Peeble 1980; Reid&White (2006?); Fisher 1994-6?),

$$1 + \xi_s(s_{\parallel}, s_{\perp}) = \int dv_{12,\parallel} f(v_{12,\parallel}|r) [1 + \xi(r)] \quad (20)$$

where $r = \sqrt{s_{\perp}^2 + (s_{\parallel} - v_{12,\parallel}/H)^2}$. This is often referred to as streaming model (Peebles 1980; Jing+).

The central ingredient of the streaming model is the pairwise velocity distribution $f(v_{12,\parallel}|r)$. Fisher (1995) has shown that the streaming model with linear theory for $f(v_{12,\parallel}|r)$ is equivalent to the familiar Kaiser effect. Scoccimaro (2004) has shown that if pairwise velocity PDF is correct the streaming is valid for all scale (check the argument).

2.3 Redshift-space correlation function

Our aim is to develop a theory to interpret the observation of the galaxy-absorber correlation function in the redshift space. The redshift-space correlation function $\xi(s_{\parallel}, s_{\perp})$ shows the anisotropy due to the redshift-space distortion imprinted by the peculiar velocity field. Although we do not intend to use such observations to measure $f = d \ln D / d \ln a$ as in modern LLS survey (e.g. BOSS, WiggleZ), we start with the linear theory to build up a solid understanding.

The two-dimensional correlation function is usually expanded in Legendre Polynomials $L_{\ell}(\mu_s)$ where $\mu_s = s_{\parallel}/s$ (Reid&White, Reid+2012; Hamilton 1992),

$$\xi(s, \mu_s) = \sum_{\ell=0}^{\infty} \xi_{\ell}(s) L_{\ell}(\mu_s), \quad (21)$$

where

$$\xi_{\ell}(s) = \frac{2\ell+1}{2} \int d\mu_s \xi(s, \mu_s) L_{\ell}(\mu_s). \quad (22)$$

For large-scale structure studies, the first three even Legendre moments $\xi_0(s)$, $\xi_2(s)$ and $\xi_4(s)$ are usually considered. The quadrupole-monopole ratio ξ_2/ξ_0 is positive for stretching in radial direction and negative for flattening. Coherent inflow produce the negative quadrupole-monopole ratio, while random velocity dispersion and coherent outflow give positive value.

In linear theory, Kaiser formula translates to (Hamilton 1992; Reid&White 2011)

$$\xi_{\ell}(s) = i^{\ell} \int \frac{dk}{k} \Delta_{\ell}^2(k) j_{\ell}(ks), \quad (23)$$

where

$$\Delta_0^2(k) = \Delta_L^2(k)(b^2 + \frac{2}{3}bf + \frac{1}{5}f^2), \quad (24)$$

$$\Delta_2^2(k) = \Delta_L^2(k)(\frac{4}{3}bf + \frac{4}{7}f^2), \quad (25)$$

$$\Delta_4^2(k) = \Delta_L^2(k)(\frac{8}{35}f^2), \quad (26)$$

2.3.1 Pairwise velocity PDF

For the perfectly coherent velocity field,

$$1 + \xi^s(s_{\parallel}, s_{\perp}) = \int dr_{\parallel} \left[1 + \xi(\sqrt{s_{\perp}^2 + r_{\parallel}^2}) \right] \delta_D [s_{\parallel} - r_{\parallel} - \mu v_{12}(r)] \quad (27)$$

where $r = \sqrt{r_{\perp}^2 + r_{\parallel}^2}$ and $\mu = r_{\parallel}/r$. The redshift-space correlation function is then given by $\xi^s(s_{\parallel}, s_{\perp}) = \xi(\sqrt{s_{\perp}^2 + r_{\parallel}^2(s_{\parallel}))$ where $r_{\parallel}(s_{\parallel})$ is the solution of implicit equation $s_{\parallel} = r_{\parallel} \left(1 - \frac{v_{12}(r)}{r} \right)$.

By assuming the velocity PDF is gaussian, we arrive at the Gaussian streaming model (Reid&White 2011)

$$1 + \xi^s(s_{\parallel}, s_{\perp}) = \int \frac{dy}{\sqrt{2\pi\sigma_{12}^2(r, \mu)}} \times \left[1 + \xi(\sqrt{s_{\parallel}^2 + y^2}) \right] \exp \left[-\frac{(s_{\parallel} - y - \mu v_{12}(r))^2}{2\sigma_{12}^2(r, \mu)} \right] \quad (28)$$

3 THEORY: THE ENVIRONMENTAL DEPENDENCE OF Ly α VISIBILITY

3.1 Cosmological Ly α radiative transfer

The statistical formulation of the Ly α radiative transfer can start with Paresce, McKee Bowyer 1980. The critical difference between continuum RT and line RT is that the former is sensitive to the clustering in real space whereas the latter is to in *velocity space*. For Ly α RT what matters is where the absorber occupies in the velocity space (marginalised over real space position). The probability of an absorbering cloud’s velocity within v and $v+dv$ is $p(v)dv$. Then the probability that a line-of-sight of Ly α galaxy to the observer does not coincide is $[1-p(v)]dv$. The fraction of photons absorbed is $p(v)e^{-\tau(v)}dv$

$$\langle I(v + \Delta v) \rangle = \langle I(v) \rangle \left\{ (1 - p(v))\Delta v + p(v)e^{-\tau(v)}\Delta v \right\} \quad (29)$$

Forming the differential equation by $d\langle I \rangle / dv \approx \langle I \rangle [p(v)(1 - e^{-\tau(v)})]$, whose solution is $\langle I \rangle = \langle I \rangle_0 e^{-\tau_{\text{eff}}}$

$$\tau_{\text{eff}} = \int_{-\infty}^{\infty} p(v_{12}) [1 - e^{-\tau(v_{12})}] dv_{12} \quad (30)$$

where v_{12} is the pairwise velocity of absorber and galaxy.

The pairwise velocity PDF $p(v_{12})$. Maybe an approach is cumulant expansion with respect to N-point correlation function derived from BBGKY hierarchy. The truncate with 2PCF (see Scherrer & Bertschinger 1991)? From the Bayes' rule

$$p(v_{12}) = \int p(v_{12}|r)p(r)dr \quad (31)$$

where $p(r)dr = \frac{4\pi r^2}{V}(1 + \xi(r))dr$ is the PDF from the real-space clustering of absorber around galaxy.

The end result for Gaussian pairwise velocity PDF.

$$\tau_{\text{eff}} = \int dN_{\text{HI}} \frac{d\mathcal{N}}{dN_{\text{HI}}} \int dv p(v_{12}) [1 - e^{-\tau(v_{12}, N_{\text{HI}})}] \quad (32)$$

$$p(v_{12}) = \int \frac{d^3r}{\sqrt{2\pi\sigma_{12}^2(\mathbf{r})}} (1 + \xi(\mathbf{r})) \exp \left[\frac{(v_{12} - \langle v_{12}(\mathbf{r}) \rangle)^2}{2\sigma_{12}^2(\mathbf{r})} \right] \quad (33)$$

Pairwise velocity at \mathbf{r} is $v_{12}(\mathbf{r})$. $\xi_v(\mathbf{r}) = \langle v_{12}(\mathbf{x})v_{12}(\mathbf{x} + \mathbf{r}) \rangle$

We can imagine two models, dirac delta PDF $p(v_{12}|r) = \delta_D(v_{12} - aHr)$

3.2 The region of influence

For both EoR $z > 6$ (McQuinn+; Dijkstra+) and the post-reionized universe $2 < z < 6$ (Zheng+; Dijkstra&Wyithe), the IGM environment influence the visibility of Ly α -emitting galaxies. The degree of the impact depends on the amount of clustering of gas, velocity field, and photoionization background in the vicinity of the galaxies. The Ly α radiative transfer is sensitive to the velocity structure. The gas velocity field is nonlinear in and out of galactic haloes, but at large enough distance it would converge to Hubble flow. The region of influence of Ly α RT is most directly characterized in the velocity space. The wing approximation to the Lorentz profile $\varphi_\nu = \frac{\Lambda/4\pi^2}{(\nu - \nu_\alpha)^2 + (\Lambda/4\pi)^2} \approx \frac{\Lambda}{4\pi^2(\nu - \nu_\alpha)^2}$ where $\Lambda = 6.25 \times 10^8 s^{-1}$ is the damping coefficient. The optical depth of an absorber of column density N_{HI} and total (hubble flow plus peculiar) proper velocity v_c is $\tau_\alpha(\nu_e) = \sigma_\alpha N_{\text{HI}} \varphi_\nu [\nu_e(1 - v_c/c)]$ The absorber velocity that gives the optical depth τ_α against Ly α line emission from galaxies is

$$v_c(\tau_\alpha) = c \sqrt{\frac{\sigma_\alpha N_{\text{HI}} \Lambda}{4\pi^2 \nu_\alpha^2 \tau_\alpha}} = 507.3 \tau_\alpha^{-1/2} \left(\frac{N_{\text{HI}}}{10^{20} \text{cm}^2} \right)^{1/2} \text{km/s} \quad (34)$$

For a strong absorbing cloud to give the optical depth of unity (37% transmission) against the Ly α line emission, the absorbing cloud cannot be outflowing relative to the central galaxy more than $\sim 500 - 2800 \text{km/s}$. This defines the region of influence of Ly α RT in the velocity space. To translate the region of influence to the real space we need to know the velocity field around a galaxy. For pure Hubble flow the comoving region of influence is

$$D_{\text{infl}} = \frac{v_c(\tau_\alpha)}{H_0} \frac{1+z}{[\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}} \quad (35)$$

3.3 Statistical formulation

The statistical formulation of the Ly α radiative transfer can start with Paresce, McKee Bowyer 1980. An alternative derivation, which clarifies the implicit assumptions and allow generalization, is shown in Appendix. The critical difference between continuum RT and line RT is that the former is sensitive to the clustering in real space whereas the latter is to in *velocity space*. The probability of an absorbing cloud's velocity within v and $v + dv$ is $p(v)dv$. Then the probability that a line-of-sight of Ly α galaxy to the observer does not coincide is $1 - p(v)dv$. The fraction of photons absorbed is $p(v)e^{-\tau(v)}dv$

$$\langle I(v + \Delta v) \rangle = \langle I(v) \rangle \left\{ (1 - p(v)\Delta v) + p(v)e^{-\tau(v)}\Delta v \right\} \quad (36)$$

Forming the differential equation by $d\langle I \rangle / dv \approx \langle I \rangle [p(v)(1 - e^{-\tau(v)})]$, whose solution is $\langle I \rangle = \langle I \rangle_0 e^{-\tau_{\text{eff}}}$

$$\tau_{\text{eff}} = \int_{-\infty}^{\infty} p(v_{12}) [1 - e^{-\tau(v_{12})}] dv_{12} \quad (37)$$

where $v_{12}^{tot} = aHr_{12} + v_{12}$ is the *total, proper pairwise velocity* of absorber and galaxy. By transforming the variables r_{12}, v_{12} into v_{12}^{tot} ,

$$\begin{aligned} f(v_{12}^{tot}) &= \iint \delta_D [v_{12}^{tot} - (Hr_{12} + v_{12})] f(v_{12}, r_{12}) dv_{12} dr_{12} \\ &= \int f(v_{12}^{tot} - Hr_{12}|r_{12}) f(r_{12}) dr_{12} \end{aligned} \quad (38)$$

where $f(r)dr = \frac{4\pi r^2}{V}(1 + \xi(r))dr$ (NO!) CORRECT ONE is $f(r)dr = \frac{1}{L}(1 + \xi(r))dr$ (pre-factor dependence is geometric factor for 1D along a skewer it does not enter!) is the PDF from the real-space clustering of absorber around galaxy.

total velocity $v_{\parallel}^{tot} = Hs_{\parallel} = H(r_{\parallel} + v_{\parallel}/H)$

$$I(s+ds) = I(s) \left\{ 1 - p(s_{\parallel}, N_{\text{HI}})ds_{\parallel} + p(s_{\parallel}, N_{\text{HI}})e^{-\tau(s_{\parallel}, N_{\text{HI}})}ds \right\} \quad (39)$$

The end result for Gaussian pairwise velocity PDF.

$$\tau_{\text{eff}} = \int dN_{\text{HI}} \frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}} \partial z} \left| \frac{dz}{dr} \right| \int dv f(v_{12}) [1 - e^{-\tau(v_{12}, N_{\text{HI}})}] \quad (40)$$

$$f(v_{12}) = \int \frac{dr}{\sqrt{2\pi\sigma_{12}^2(r)}} (1 + \xi(r)) \exp \left[\frac{(v_{12}^{tot} - Hr - \langle v_{12}(r) \rangle)^2}{2\sigma_{12}^2(r)} \right] \quad (41)$$

The end result for Gaussian pairwise velocity PDF.

$$\tau_{\text{eff}} = \int dN_{\text{HI}} \frac{d\mathcal{N}}{dN_{\text{HI}}} \int dv f(v_{12}) [1 - e^{-\tau(v_{12}, N_{\text{HI}})}] \quad (42)$$

$$f(v_{12}) = \int \frac{4\pi r^2 dr}{\sqrt{2\pi\sigma_{12}^2(r)V}} (1 + \xi(r)) \exp \left[\frac{(v_{12}^{tot} - Hr - \langle v_{12}(r) \rangle)^2}{2\sigma_{12}^2(r)} \right] \quad (43)$$

The end result for Gaussian pairwise velocity PDF.

$$\tau_{\text{eff}} = \int dN_{\text{HI}} \frac{d\mathcal{N}}{dN_{\text{HI}}} \int dv f(v_{12}) [1 - e^{-\tau(v_{12}, N_{\text{HI}})}] \quad (44)$$

$$f(v_{12}) = \int \frac{4\pi r^2 dr}{\sqrt{2\pi\sigma_{12}^2(r)V}} (1 + \xi(r)) \exp \left[\frac{(v_{12}^{tot} - Hr - \langle v_{12}(r) \rangle)^2}{2\sigma_{12}^2(r)} \right] \quad (45)$$

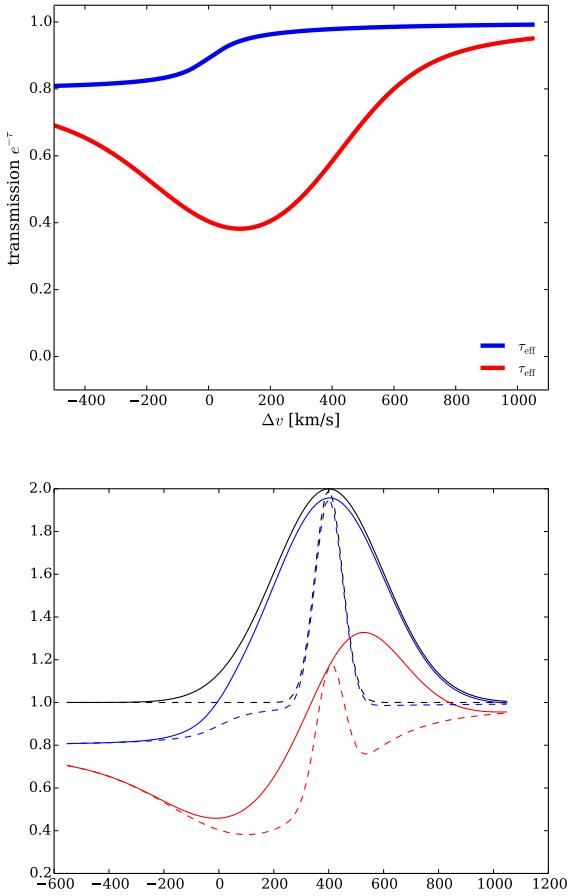


Figure 5. Example calc. of effective optical depth with statistical formalism

For generalization to the anisotropic real-space clustering, e.g. biconical outflow + two cold streaming model

$$f(v_{12}) = \int \frac{d^3 r}{\sqrt{2\pi\sigma_{12}^2(\mathbf{r})}} (1 + \xi(\mathbf{r})) \exp \left[\frac{(v_{12} - \langle v_{12}(\mathbf{r}) \rangle)^2}{2\sigma_{12}^2(\mathbf{r})} \right] \quad (46)$$

Pairwise velocity at \mathbf{r} is $v_{12}(\mathbf{r})$. $\xi_v(\mathbf{r}) = \langle v_{12}(\mathbf{x})v_{12}(\mathbf{x} + \mathbf{r}) \rangle$

We can imagine two models, dirac delta PDF $p(v_{12}|r) = \delta_D(v_{12} - aHr)$

3.4 Caveats

3.4.1 Pairwise velocity PDF

How good is the gaussian approximation to pairwise velocity PDF? The assumption that the pairwise velocity PDF may introduce systematic bias in analytic model. Scoccimaro 2004 has studied that the pairwise velocity PDF is in fact highly non-gaussian with skewness and kurtosis. Tinker 2005 modelled the non-gaussian pairwise velocity PDF in the HOD framework. Generalisation can also be done possibly by the cumulant expansion or Edgeworth expansion (e.g. Matsubara2008+; Sherrer&Bertschinger 1991).

Scoccimaro 2004 argued that the direct reconstruction of the PDF from redshift-space clustering measurement is

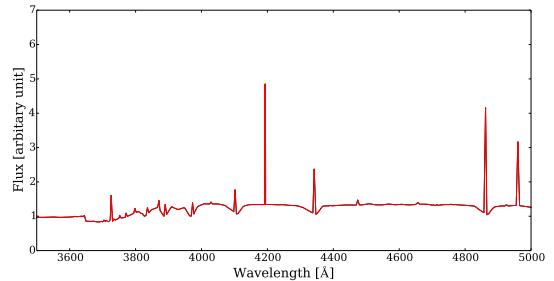


Figure 7. Optical part of synthetic spectrum.

in general impossible, which must rely on the assumption of scale-independent PDF or Gaussianity.

The pairwise velocity PDF $p(v_{12})$. Maybe an approach is cumulant expansion with respect to N-point correlation function derived from BBGKY hierarchy. The truncate with 2PCF (see Scherrer & Bertschinger 1991)?

4 APPLICATION OF THEORY

4.1 Spectral synthesis with the intergalactic environment

We show an application of the joint model of RSD and Ly α RT to the stellar population synthesis. We have been considering the deep spectroscopic galaxy survey in the foreground of QSOs. The availability of spectroscopy of individual galaxies were supposed to perform the RSD measurement.

The statistical formalism presented in the section 3 allows us to perform the spectral synthesis including the RT effect of large-scale environment. Since the independent constraint on Ly α effective optical depth can be provided from the RSD measurement.

The distribution of foreground absorbers of the galaxies attenuates the restframe wavelength shorter than 1216 \AA by Lyman series line blanketing and photonionization absorption (< 912 \AA) (Madau 95, Meiksin 06, Inoue 12).

5 SIMULATION

5.1 Cosmological hydrodynamical simulation

We carried out the cosmological hydrodynamical simulation of the IGM using AMR hydro/N-body code RAMSES (Teyssier 2002). We performed a series of adiabatic simulation with varying box size and resolution elements for convergence test.

We adopt the quasi-Lagrangian refinement strategy with refinement criterion of 10 particles by which the cell is refined when the number of dark matter particle or equivalent gas mass exceeds 10.

The initial condition is generated with COSMICS package (Bertschinger 1996?). The COSMICS package is initialized by setting the initial rms fluctuation $\sigma_8(z = z_{ini}) = 0.03$. Because the resolution of simulations are varying the initial redshift is elevated to lower redshift for lower resolution simulation. The initial temperature is set as $T = 650\text{K}$. This is compromise because the adiabatic simulation

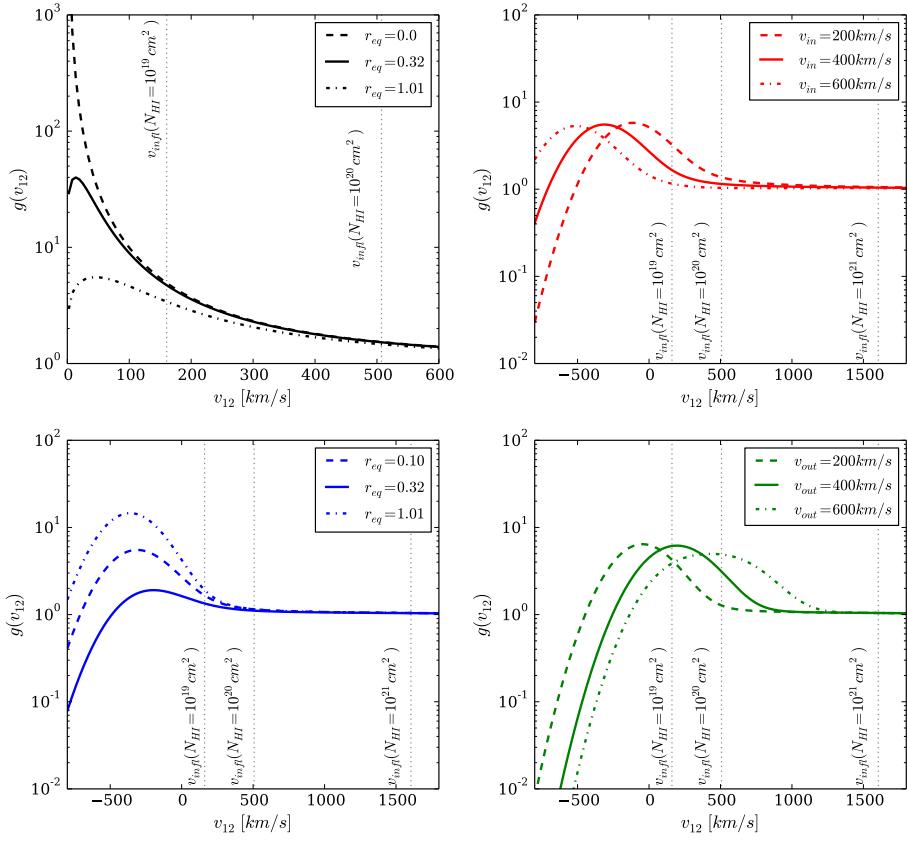
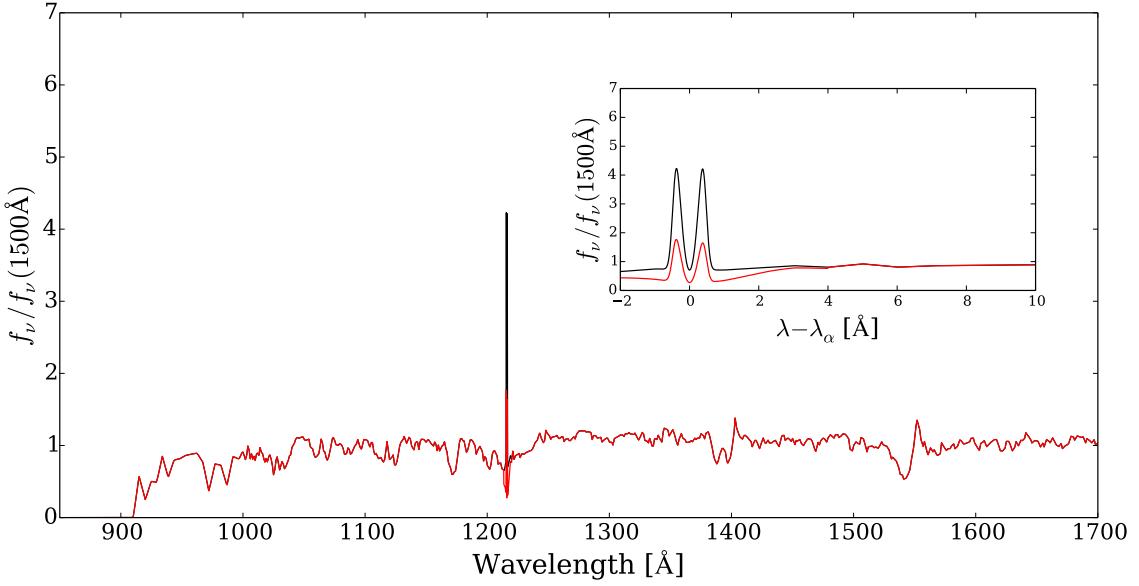
**Figure 4.** pairwise velocity distribution. Enhancement due to velocity space clustering**Figure 6.** example spectral synthesis.

Table 1. Simulation set up

name	Box size L [h^{-1} cMpc]	N_{DM}	$N_{\text{grid}}(\text{base})$	$N_{\text{grid}}(\text{fine})$	m_{DM} [$h^{-1}\text{M}_{\odot}$]	$\Delta L(\text{base})$ [h^{-1} ckpc]	$\Delta L(\text{fine})$ [h^{-1} ckpc]	z_{ini}	$J_{21}, \alpha, z_{\text{reion}}$
<u>static grid + adiabatic</u>									
L10P256G256R0	10	256 ³	256 ³		3.57×10^6	39.1		276	
L20P256G256R0	20	256 ³	256 ³		2.86×10^7	78.1		237	
L40P256G256R0	40	256 ³	256 ³		2.29×10^8	156.2		199	
L60P256G256R0	60	256 ³	256 ³		7.72×10^8	234.4		178	
<u>AMR grid + adiabatic</u>									
L40P256G256R2	40	256 ³	256 ³	1024 ³	2.29×10^8	156.2	39.1	199	
<u>static grid run + UV background</u>									
L40P256G256R0	40	256 ³	256 ³		2.29×10^8	156.2		199	1.0, 1.0, 8.5

does not include Compton heating from CMB and cooling. The IGM only be cooled or heated by adiabatic process.

The haloes are identified by HOP algorithm (Eisenstein+; Aubert).

5.2 Radiative transfer

We introduce the self-shielded gas by Rhamati+ criterion. The method is calibrated based on radiative transfer simulation.

5.3 Synthetic QSO spectra

5.4 Synthetic galaxy catalogue

6 SURVEY: MOCK

7 RESULTS

7.1 Dynamics of galaxy-absorber pairs

To build the understanding of the dynamics of large-scale environment of galaxies, i.e. galaxy-absorber pairs, we compare the analytic models (linear theory and self-similar ansatz) with numerical simulation.

Firstly, we assess the reliability of the simulation by comparing the 2PCF of galaxy and absorber, which is shown in Fig.8. The 2PCF was measured using pair-count estimator $\xi = DD/RR - 1$. Given the simplicity of the IGM simulation, the simulated 2PCF is a reasonable agreement with the best-fit power-law correlation function from observation by Cooke+06. While the detailed convergence tests are necessary to assess the robustness of the simulated 2PCF, we take our simulation as a first fiducial reference point to consider the pairwise velocity statistics (and the redshift evolution of 2PCF).

The radial component of mean pairwise velocity between galaxy and absorber is shown in Fig.9. We compare the linear theory, self-similar theory, and simulations at $z = 3$. As expected, the linear theory severely underestimate the magnitude of cosmological inflow velocity. While we allowed the nonlinear evolution in the self-similar ansatz, which we left the amplitude v_{in} as a free parameter, the radial dependence has different form when we apply the best-fit correlation function of Cooke+06. However, the functional form derived in the self-similar ansatz provide a convenient functional form that can be well fitted to simulation. The best fit parameters are $v_{in} = 143\text{km/s}$, $r_{in} = 8.4h^{-1}\text{cMpc}$, and $\gamma = 2.26$.

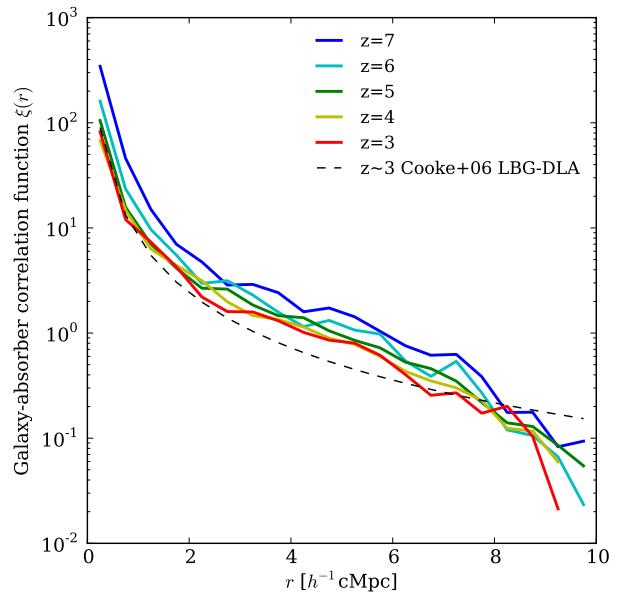


Figure 8. The comparision between the simulated 2PCF for different redshift and the best-fit 2PCF from observation by Cooke+06.

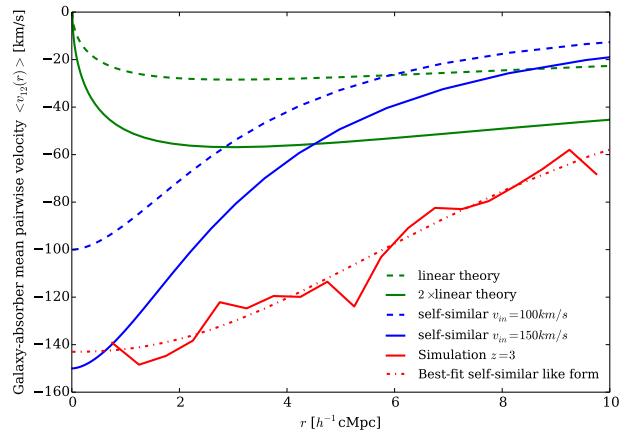
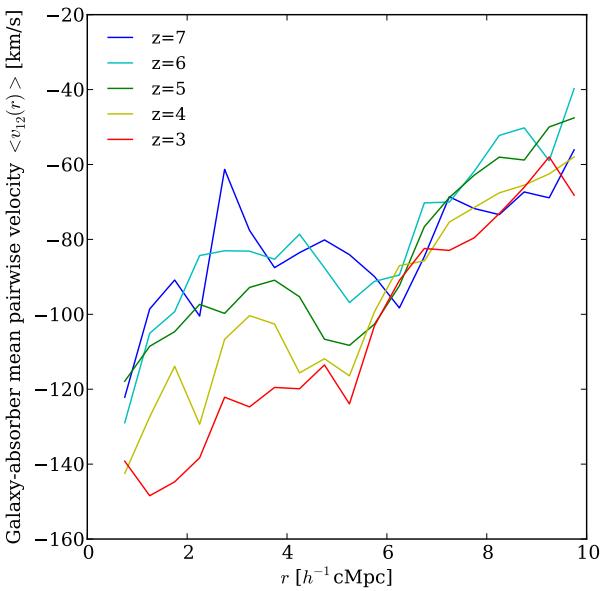


Figure 9. The comparison of mean pairwise velocity between analytic theory and simulation at $z = 3$.

Table 2. Model grids

		Available source of constraints
CDDF: $\frac{\partial^2 \mathcal{N}}{\partial N_{\text{HI}} \partial z}$	Haardt&Madau (2012) fit (fid.)	Ly α forest, LLS/DLA survey
real-space clustering: ξ	$r_0 = 3.32 h^{-1} \text{cMpc}$ (fid.)	$\gamma = 1.76$ (fid.) angular correlation function
local photoionization effect: S	$r_{eq}(f_{esc}^{LyC} \text{SFR}) = 0.1, 0.32(\text{fid.}), 1.01 h^{-1} \text{cMpc}$	$\beta_N^{\text{eff}} = 2$ (fid.) SFR from UV magnitude or SED/spectral fitting
velocity statistics:		primary interest of constraining by RSD
v_{12}	$v_{in} = 100, 143(\text{fid.}), 200 \text{km/s}$	$\gamma_{in} = 2.26(\text{fid.})$
v_{out}	$r_{out} = 1(\text{fid.})$	possibly helped with metal line
σ_{12}	$180, 200(\text{fid.}), 220 \text{km/s}$	

**Figure 10.** The redshift evolution of the mean pairwise velocity from simulation.

7.2 Redshift-space distortion

We consider how the RSD measurement between galaxies and absorbers helps to constrain the pairwise velocity statistics, which is a critical ingredient to predict the Ly α visibility affected by the intergalactic environment. While it may be possible to let all the model parameters free and attempt to constrain by RSD only, we take rather a moderate approach that some model quantities can be constrained prior to RSD measurement.

CDDF can be constrained by Ly α forest and LLS/DLA survey. The real-space correlation function is from the angular correlation function $\omega(\theta)$. This reduces the parameter space to be local photoionization effect and pairwise velocity statistics. Analysis of SED for wavelength $> 1216 \text{\AA}$ can give a helpful information to characterize the local photoionization effect such as SFR. If the metal lines are present, the outflow velocity may be guessed.

We explore the model grids as shown in Table 2, which has the total number of 13 models composed of the 3 variants of local photoionization, and 3 variants of inflow velocity, 4

variants of outflow velocity, and 3 variants of pairwise velocity dispersion. For each variant, the rest of parameter is taken as the fiducial values.

7.2.1 The variation of local photoionization

In Fig. 11, we show that the 2D plot of the redshift-space correlation function. The fiducial (left) model with $f_{esc}^{LyC} \text{SFR} = 1 M_{\odot} \text{yr}^{-1}$ is varied for low (middle) and high (right) local photoionization effect. The low and high photoionization models correspond to $f_{esc}^{LyC} \text{SFR} = 0.1, 10 M_{\odot} \text{yr}^{-1}$ respectively. The low photoionization model produces the higher value of $\xi_s(\sigma, \pi)$ for small separation because the real-space correlation function truncate much smaller scale. Consequently, the figure-of-God effect on small scale become more apparent for lower photoionization effect. More quantitative effect is seen in the Legendre moments of redshift-space correlation function as shown in Fig. 12. The monopole as affected by local photoionization effect with more small scale suppression for higher photoionization effect. In the quadrupole moment, this propagates as larger finger-of-God effect resulting in the larger value of quadrupole moment for smaller scale. The monopole-quadrupole moment is an indicator of stretching (positive) and squashing (negative) of RSD. The suppression of the small-scale power by local photoionization may produce negative monopole-quadrupole ratio because the squashing of RSD is weighted more in the quadrupole moment.

7.2.2 The variation of inflow velocity

Figs. 13 and 14 show that the RSD and Legendre moment varying the inflow velocity. As well-known (Kaiser 96?), the larger inflow velocity produces more squashing of the RSD. We explore how the variation of cosmological inflow velocity affect the RSD signature. As seen in Fig. 14, the monopole is not affected by the change of inflow velocity. This is understood because the change in velocity field is washed out in the monopole moment. The quadrupole and monopole-quadrupole ratio shows the clear sign of increasing squashing of RSD on larger scale, producing more negative value for larger inflow velocity.

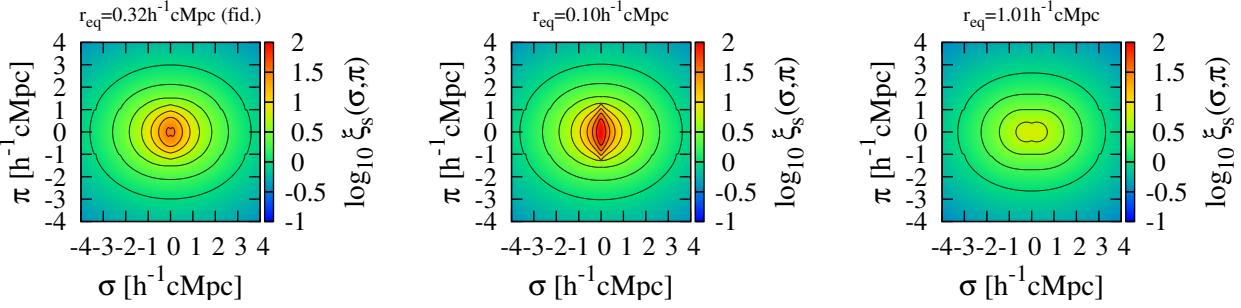


Figure 11. RSD varying the local photoionization effect.

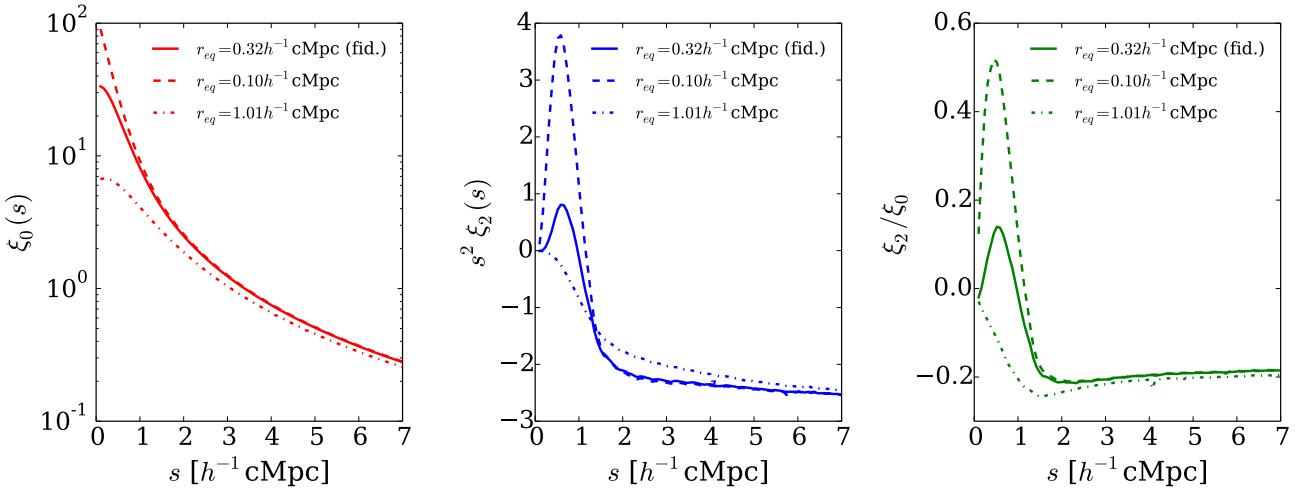


Figure 12. The monopole, quadrupole and monopole-quadrupole ratio of the Legendre moments of redshift-space correlation function.

7.2.3 The variation of pairwise velocity dispersion

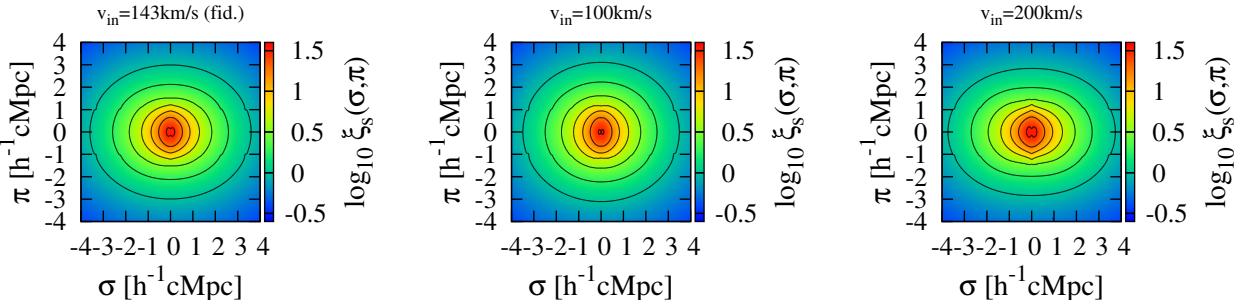
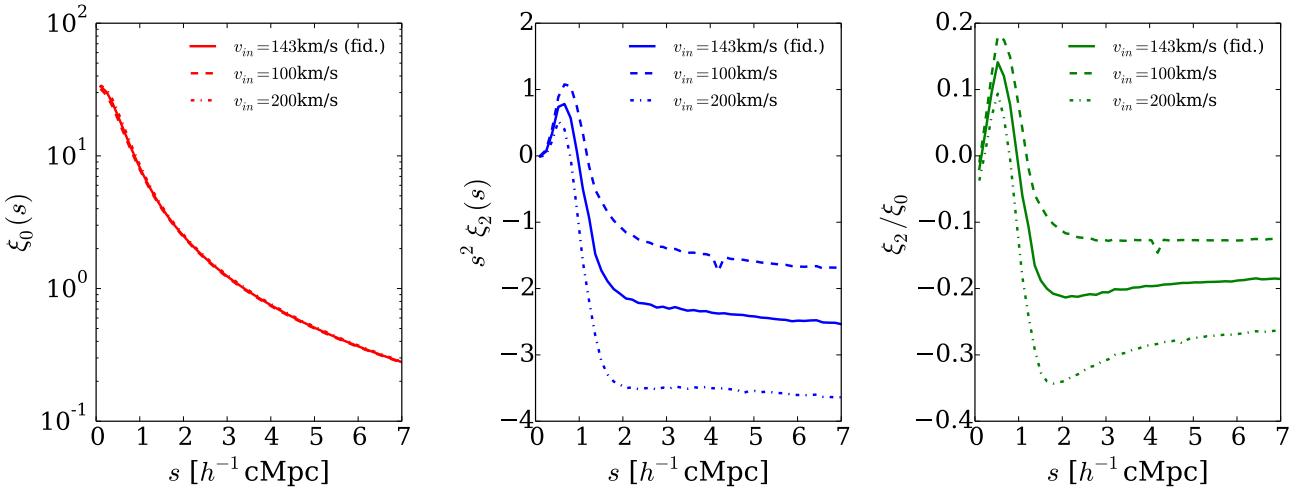
The variation of RSD and Legendre moments for different pairwise velocity dispersion is shown in Figs.15 and 16. Increasing Finger-of-God effect is seen for larger pairwise velocity dispersion. Similarly to the change of inflow velocity, the monopole moment has only a minor impact from the difference in pairwise velocity dispersion. The amplitude of the small-scale part of the quadrupole moment varies as a result of different size of Finger-of-God effect, while the large scale is less affected by the pairwise velocity dispersion.

7.2.4 The variation of outflow velocity

The outflow as considered in the phenomenological model show a peculiar signature in RSD as shown in Fig.17. The outflow stretches the RSD in the line-of-sight direction, creating two island-like feature in the RSD. This can be interpreted as a result of the combination of outflow and the local photoionization. Since the local photoionization creates the suppression of correlation function in small-scale separation, the line-of-sight stretching produces the islands in the RSD. The higher outflow velocity creates the larger line-of-sight separation of the two islands.

In Fig.18, the Legendre moments also shows the sign

of outflow effect. Unlike the inflow and pairwise velocity dispersion, the monopole is also affected by outflow. The apparent suppression on the small-scale monopole moment is the result of the line-of-sight stretch of local photoionization effect of RSD. Since the small-scale suppression in real-space is systematically displaced in the line-of-sight direction, the stronger outflow produce more suppression of the monopole. Note that since inflow displaces the local photoionization effect into smaller scale, such effect is not visible in the monopole. Also, as the pairwise velocity dispersion gives random displacement the effect on monopole is very minor (but slightly larger than the impact by inflow). The quadrupole clearly indicates the larger stretching and redshift-space displacement by outflow by pushing and amplifying the positive peak of quadrupole moment. Both for quadrupole and monopole-quadrupole ratio shows that for larger scale the impact of cosmological inflow still produces the squashing of RSD. Finally, note that the negative monopole-quadrupole ratio (yet the RSD is stretching) is a result of local photoionization. Since we take the power-law slope of CDDF in the DLA regime to be $\beta_N^{\text{eff}} = 2$, the real-space correlation function bend to have positive slope by strong local photoionization suppression on smaller scale. Hence, as the quadrupole moment positively weights

**Figure 13.** RSD varying the inflow velocity.**Figure 14.** The monopole, quadrupole and monopole-quadrupole ratio of the Legendre moments of redshift-space correlation function.

in the line-of-sight direction and negatively weights in the perpendicular direction, the net result is negative value of quadrupole, although the RSD is being stretched in the line-of-sight direction.

7.2.5 Can we handle the pairwise velocity statistics from RSD measurement?

The consideration of how RSD and Legendre moments are affected by the change in our parameters of interest to be constrained, i.e. pairwise velocity statistics of galaxies and absorbers, helps how well can we constrain these parameters by the RSD measurement. As illustrated above, the local photoionization effect and the pairwise velocity statistics (inflow velocity, pairwise velocity dispersion, outflow velocity) influence the RSD and Legendre moment in different way. Fitting of RSD model provides insightful constraints on the pairwise velocity statistics (**Can I do Fisher matrix forecast around a fiducial model?**). One may hope that even in the presence of redshift error on the galaxy and absorber position, which produces the degenerate effect to scale-independent pairwise velocity dispersion, other interested parameter such as inflow velocity, may be able to constrain.

7.3 Ly α RT through the intergalactic environment

8 APPLICATIONS

We present a couple of interesting applications of the RSD measurement enabled by joint analysis of galaxy redshift survey and Ly α forests over $2 < z < 7$.

8.1 Breaking H i fraction-topology degeneracy of EoR

8.1.1 Strategy to break global H i fraction-topology degeneracy

The expectation that the reionization effect on galaxy-absorber correlation function imprinted on large-scale has very important consequence on breaking the degeneracy between small-scale absorbers and reionization. By using small scale RSD $r \lesssim 10 h^{-1} \text{cMpc}$ we can infer the dynamics and distribution of small-scale absorbers. This then inform us, as we show later, that the effective optical Ly α depth due to the small-scale absorbers for Ly α -emitting galaxies. This is then compared with the observed Ly α visibility $\tau = \tau_{\text{web}} + \tau_{\text{bub}}$. The required additional contribution from H II bubble distribution inform us the global H i fraction of the IGM. The reionization model should in turn consistently reproduce the

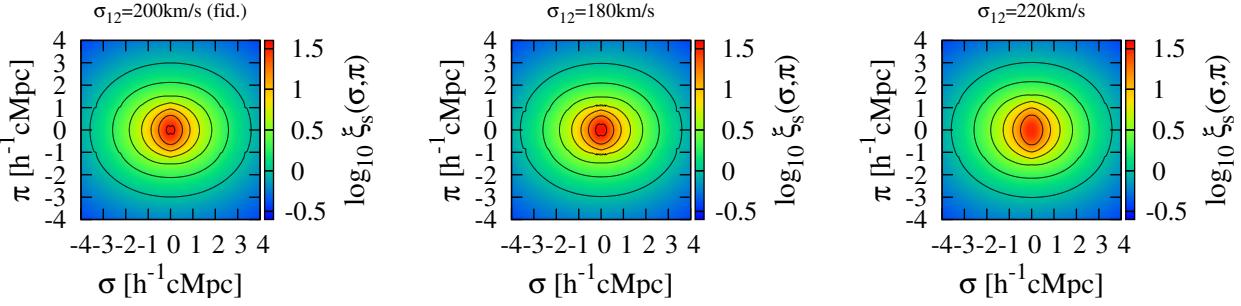


Figure 15. RSD varying the pairwise velocity dispersion.

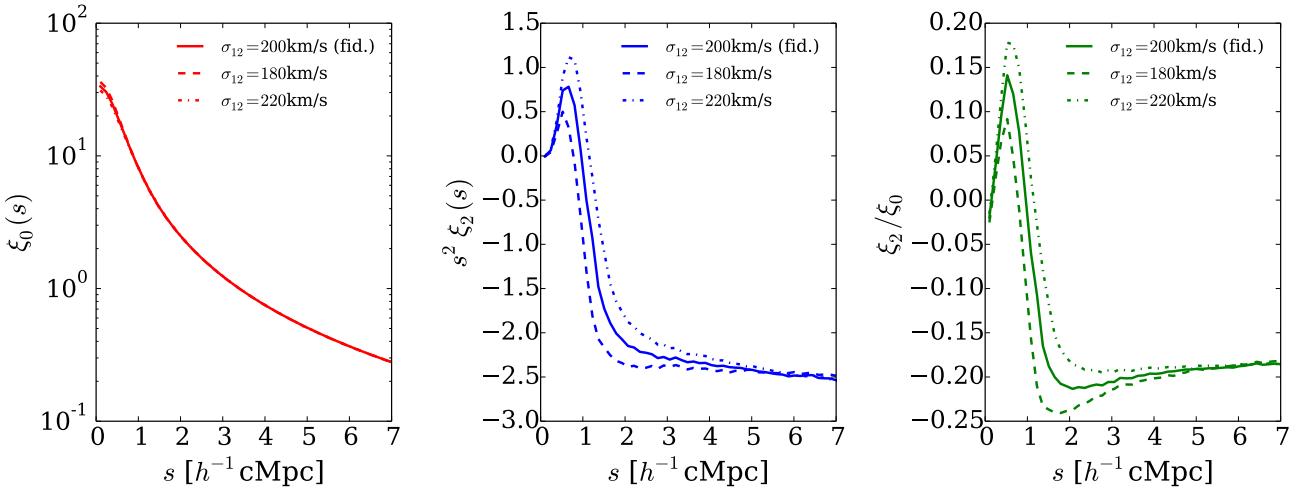


Figure 16. The monopole, quadrupole and monopole-quadrupole ratio of the Legendre moments of redshift-space correlation function.

observed galaxy-absorber correlation function on large-scale $r > R_b$ as a result of H I patches outside the H II bubbles. This ability to handle the small-scale absorbers and reionization effect allows us to make a diagnostic to break the global H I fraction-topology degeneracy, hence simultaneously measuring the two essential quantities of the EoR astrophysics.

8.2 Testing Ly α RT effect for HETDEX cosmology survey

The measurement of small-scale RSD between galaxies and absorbers using the QSO spectra allows us to test the Ly α RT effect on Ly α -selected galaxy cosmology survey. By masking the Ly α line, we select by using OIII line or Lyman break galaxies. Then, we measure the RSD of UV-selected or OIII-selected (or any other ISM nebular lines) galaxies with absorbers (this can be small sub-field of HETDEX). Based on the statistical formalism presented above, we can predict the possible average Ly α transmission effect. If the impact of Ly α optical depth is large, we need to take into account the Ly α RT effect in clustering measurement, whereas if the Ly α optical depth is negligible it provides an negative evidence for Ly α RT effect. However, note that RSD-Ly α visibility diagnostic does not perfectly discard the Ly α RT effect since

it only takes into account the average, i.e. stacked, Ly α visibility. If the substantial spatial variation of IGM environment of Ly α -emitting galaxies is possible, only characterizing the mean visibility is not sufficient to argue that Ly α RT effect is negligible in HETDEX. This case the repeating RSD diagnostic for different sub-fields would give an insight on the variation of IGM environment of Ly α -emitting galaxies.

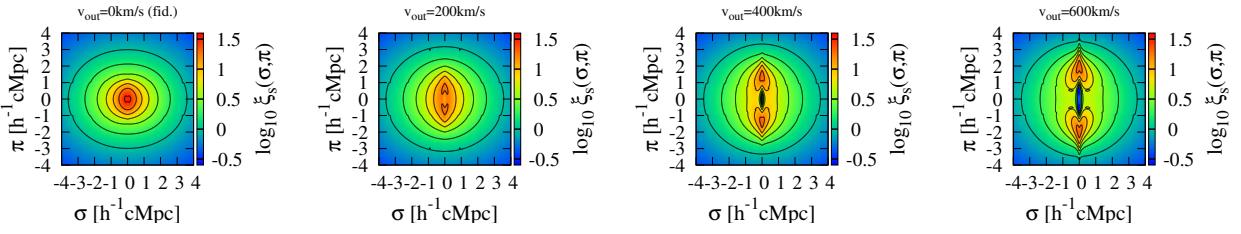
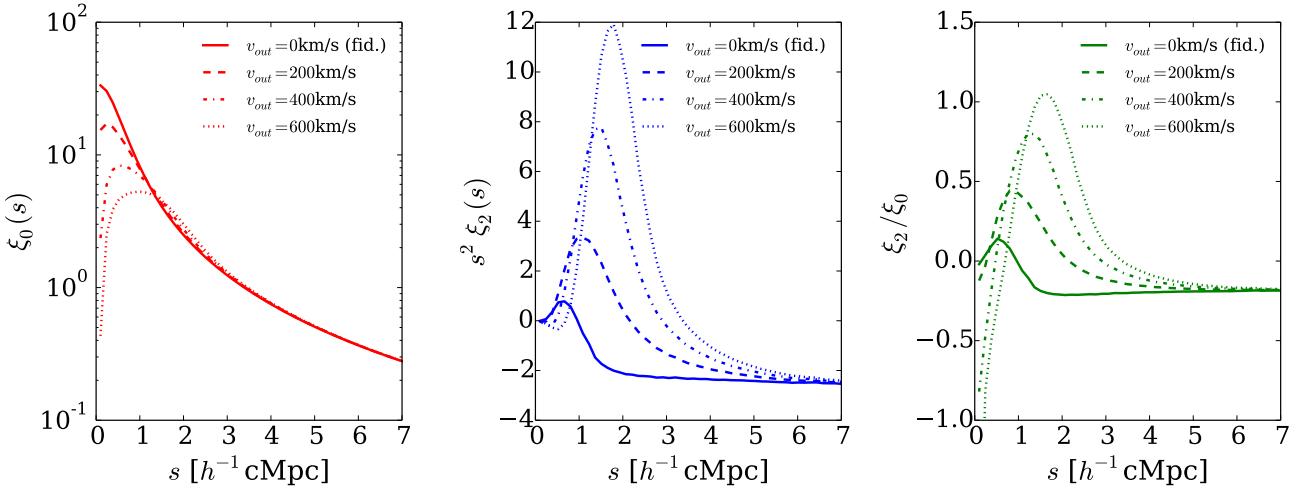
9 SURVEY REQUIREMENTS

We propose the sketch of survey requirement to characterise the IGM/CGM environment of Ly α -galaxies for $2 < z < 7$.

9.1 QSO fields

Firstly, the selection of the potential target fields is listed in Table 3. The redshift range covered by the Ly α forest between Ly α and Ly β lines in the restframe of a background QSO at redshift z_Q is $\Delta z/(1+z_Q) = \Delta\lambda/\lambda_\alpha$ where $\Delta\lambda = \lambda_\alpha - \lambda_\beta$ is the wavelength interval between Ly α and Ly β lines.

To perform the Voigt profile decomposition of Ly α forests, both high signal-to-noise ratio per pixel and high resolution ($\text{FWHM} \lesssim 25$ km/s) spectroscopy of a QSO should

**Figure 17.** RSD varying the outflow velocity.**Figure 18.** The monopole, quadrupole and monopole-quadrupole ratio of the Legendre moments of redshift-space correlation function.

be obtained (?). For example, KBSS survey (?) has observed 15 QSOs with Keck/HIRES, obtaining $R \simeq 45000$ (FWHM $\simeq 7$ km/s) and S/N $\sim 50 - 200$ pixel $^{-1}$. The resolved FWHM sets the lower limit of b -parameter can be measured from voigt profile decomposition.

While the high resolution and high S/N spectroscopy is always desirable whenever available, there are a couple of ways to characterise the absorbing systems with the equivalent width and redshift using intermediate resolution spectroscopy (?). We explore the pixel optical depth method, EW-redshift identification, and voigt profile decomposition to characterise the CGM/IGM environment using the synthetic spectra from simulations in the subsequent section.

9.2 Galaxy fields

9.2.1 Angular coverage

The angular coverage of a survey should be large enough that the region Ly α RT influence is well included. $\theta_{\text{infl}} = D_{\text{infl}}/D_A(z)$ where D_A is the angular diameter distance. This requirement is easily met for most of telescope with a single field-of-view (FoV), e.g. for Subaru/Supreme-Cam $34' \times 27'$ FoV.

The region of influence can extend as large as $\sim 50 h^{-1} \text{cMpc}$ for DLAs, whereas it is as small as $\sim 2 h^{-1} \text{cMpc}$ for LLS. Since the angular size of the region of influence (ca. 10-100 arcmin) stays approximately constant, the same mo-

saicing can be applied to survey the wide range of redshift. For the Subaru/Magellan-like ground-based telescope, the region of influence of LLSs can be covered by the single field of view. For DLA, the several mosaicing, say 2-7 tiles for one direction, is required. For the HST-like space-based telescope, also for JWST, order of 10 tiles for LLS and of 100 tiles for DLA are required.

Note that the estimate of the region of influence presented here only includes the Hubble flow. The systematic deviation due to peculiar velocity by the large-scale structure formation produces inflow. In the presence of large-scale inflow, the comoving region of influence will become larger.

9.2.2 Spectroscopic requirement

The redshift error of objects (Ly α forests and galaxies) should be below the inflow or outflow velocity scale that we want to probe. For $z = 2 - 3$ LBGs, ? observed using metal lines that the large-scale galactic outflow spans over 100km/s-600km/s. To gain the velocity space error below 100km/s, the required absolute redshift error should be below $\Delta z = \delta v/c = 0.0003$. In terms of spectroscopy, the spectral resolution should exceed $R = \lambda/\Delta\lambda = (1+z)/\Delta z \approx 3000(1+z)$.

Table 4. Telescopes and instruments for redshift surveys with Ly α forests

Telescope/Instrument	Field of View [arcmin \times arcmin]	Pixel resolution [arcsec/pixel]	Ref.
<u>Imaging</u>			
Subaru/Supreme-Cam	34' \times 27'	0.202	Ouchi+2008 ++
HST/WFC3(ACS)	$\sim 2.1'()$	0.128	Ellis+2012 (HUDF12)
Magellan/IMACS	27.2' \times 27.2'	0.2	Dressler+2014
<u>Spectroscopy</u>			
Keck/HIRES			
Keck/DEIMOS			
VLT/X-Shooter			
Gemini/GMOS			

Table 5. Future telescopes and instruments for redshift surveys with Ly α forests

Telescope/Instrument	Field of View [arcmin \times arcmin]	Pixel resolution [arcsec/pixel]	Ref.
<u>Imaging</u>			
Subaru/HSC			
ELT			
JWST			
<u>Spectroscopy</u>			
Subaru/PFS			

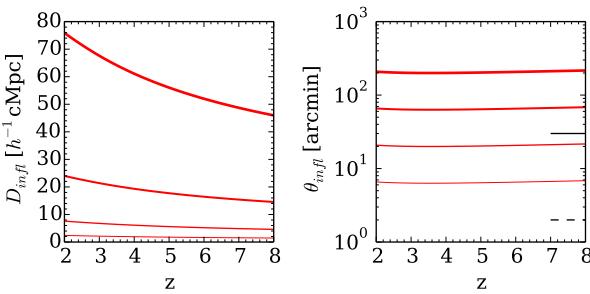


Figure 19. The comoving region of influence (left pannel) and the angular size on the sky (right pannel) for the strong absorbers with column density $N_{\text{HI}} = 10^{19}, 10^{20}, 10^{21}, 10^{22} \text{ cm}^{-2}$ (from bottom to top). The field-of-views of the Subaru/Magellan-like ground-based telescope (30arcmin, solid) and HST-like space-based telescope (2arcmin, dashed) are shown as horizontal black lines.

9.2.3 Number counts of LAE/LBG selections

We estimate the expected number counts of Ly α /UV-selected galaxies in the QSO field within the field-of-view to cover the region of influence. The required flux limit provide the feasibility of such survey strategy.

The expected number of galaxies within the survey volume is

$$N(> F_{\text{lim}}) = \Omega_{\text{survey}} \int_{z_1}^{z_2} dz \frac{d^2 V}{dz d\Omega} \int_{4\pi D_L^2(z) F_{\text{lim}}}^{\infty} \frac{dn(> L, z)}{dL} dL \quad (47)$$

where $\frac{d^2 V}{dz d\Omega} = |\frac{dr}{dz}| D_A^2(z)$ and we have assume plane-parallel approximation and \bar{z} is the mean redshift depth of the survey. $z_1 = z_Q - \Delta z$ and $z_2 = z_Q$. Since the obseravtion showed that for $z = 3 - 6$ the LAE luminosity function evolves very

slowly, for Schechter fit $\phi(L)dL = \phi_*(L/L_*)^\alpha e^{-L/L_*} dL/L_*$,

$$N(> F_{\text{lim}}) \approx \Omega_{\text{survey}} \int_{z_1}^{z_2} \frac{c D_A^2(z) dz}{H(z)} \phi_* \Gamma \left(1 + \alpha, \frac{4\pi D_L^2(z) F_{\text{lim}}}{L_*} \right) \quad (48)$$

where $\Gamma(a, x)$ is the upper incomplete Gamma function.

9.2.4 Galaxy-absorber pair counts

To measure the galaxy-absorber correlation function, we need to sample the sufficient number of the galaxy-absorber pairs to lower the random sampling error. For the simple case of Poisson error estimate, the random uncertainty in the correlation function is $\frac{\Delta \xi(r)}{1 + \xi(r)} = 1/\sqrt{N_{\text{pair}}(r)}$, which is the lower bound of more realistic error (Peacock 1998). To measure the correlation function below 10% error, more than 100 pairs per bin are required. The expected number of galaxies centred at each absorber ($N_{a,\text{pair}}(r) = N_{\text{pair}}(r)/N$) is $N_{a,\text{pair}}(r) = \frac{4\pi R^3}{3} \int_{L_{\text{lim}}}^{\infty} \frac{dn}{dL} \int_{r-\frac{1}{2}\Delta r}^{r+\frac{1}{2}\Delta r} p(r) dr$ where R is the radius of survey field-of-view and $p(r)dr$ is the probability that a galaxy is found in the interval r and $r + dr$ around an absorber. For the Poisson distribution of galaxies, $p(r)dr = 4\pi r^2 dr / ((4\pi/3)R^3)$. If one expects some degree of clustering around an absorber, the probability is modified as $p(r)dr = 4\pi r^2 [1 + \xi_{\text{ag}}^{\text{expt}}(r)] dr / ((4\pi/3)R^3)$ where $\xi_{\text{ag}}^{\text{expt}}(r)$ is the expected absorber-galaxy correlation function. Thus, the total number of galaxy-absorber pairs for each bin Δr for each QSO field is given by $N_{\text{pair}}(r) = \mathcal{N} N_{a,\text{pair}}$,

$$N_{\text{pair}}(r) = \mathcal{N} \int_{L_{\text{lim}}}^{\infty} \frac{dn}{dL} dL \int_{r-\frac{1}{2}\Delta r}^{r+\frac{1}{2}\Delta r} 4\pi r^2 (1 + \xi_{\text{ag}}^{\text{expt}}(r)) dr. \quad (49)$$

The linear theory expectation of galaxy-absorber correlation function is $\xi_{\text{ag}}^{\text{lin}}(r) = b_a b_g \int \Delta^2(k) \frac{\sin kr}{kr} dk$ (check!) where b_a and b_g are the linear bias of absorbers and galaxies respectively. For $2 < z < 3$ BOSS, $b_{\text{LyF}}, b_{\text{DLA}}, b_{\text{LAE}}, b_{\text{LBG}}$. The

Table 3. List of potential targets of QSO fields. QSOs used in Fan+2006 and Becker+2014 are listed. Also three QSOs from Pan-STARRS1 survey (Venemans+15).

QSO name	QSO redshift	
<i>z > 6</i>		
PSO J338.2298+29.5089	6.658	Venemans+15
PSO J036.5078+03.0498	6.527	Venemans+15
PSO J167.6415-13.4960	6.508	Venemans+15
SDSS J1148+5251	6.4189	Fan+06
SDSS J1030+0524	6.3110	Fan+06
CFHQS J0050+3445	6.25	
SDSS J1623+3112	6.2470	Fan+06
SDSS J1048+4637	6.2284	Fan+06
SDSS J125051.93+313021.9	6.1300	Fan+06
ULAS J1319+0950	6.13	
SDSS J23150023	6.12	
SDSS J1602+4228	6.0700	Fan+06
SDSS J1630+4012	6.0650	Fan+06
SDSS J20540005	6.06	
SDSS J0353+0104	6.05	
SDSS J0818+1722	6.02	Fan+06
SDSS J1306+0356	6.0160	Fan+06
SDSS J113717.73+354956.9	6.0100	Fan+06
<i>5 < z < 6</i>		
ULAS J0148+0600	5.98	
SDSS J1411+1217	5.9270	Fan+06
SDSS J133550.80+353315.8	5.9012	Fan+06
SDSS J0005-0006	5.85	Fan+06
SDSS J0840+5624	5.8441	Fan+06
SDSS J143611.74+500706.9	5.8300	Fan+06
SDSS J104433.04-012502.2	5.7824	Fan+06
SDSS J0836+0054	5.774	Fan+06
SDSS J092721.82+200123.7	5.7722	Fan+06
SDSS J0203+0012	5.72	
SDSS J02310728	5.42	
SDSS J1659+2709	5.32	
SDSS J1208+0010	5.27	
SDSS J0915+4244	5.20	
SDSS J12040021	5.09	
<i>4 < z < 5</i>		
SDSS J00400915	4.98	
SDSS J0011+1446	4.95	
SDSS J22250014	4.89	
SDSS J1616+0501	4.88	
BR 12020725	4.70	
SDSS J21470838	4.60	
BR 03533820	4.59	
BR 10330327	4.52	
BR 00066208	4.52	
BR 07146449	4.49	
BR 04185723	4.48	
<i>z < 3 – 4</i>		
see BOSS Ly α survey		Kee-Gee?+

expected number of absorbers for each QSO spectrum is

$$\mathcal{N}(N_{\text{HI}}) = \int_{z_1}^{z_2} dz \int_{N_{\text{HI}}^{\min}}^{N_{\text{HI}}^{\max}} dN_{\text{HI}} \frac{d^2 \mathcal{N}}{dN_{\text{HI}} dz} \quad (50)$$

Fig.22 shows that expected galaxy-absorber pair counts. We have assumed the fixed power-law correlation function $\xi(r) = (r/r_0)^{-1.74}$ (Cooke+2003) and the CDDF (O'Meara+2012).

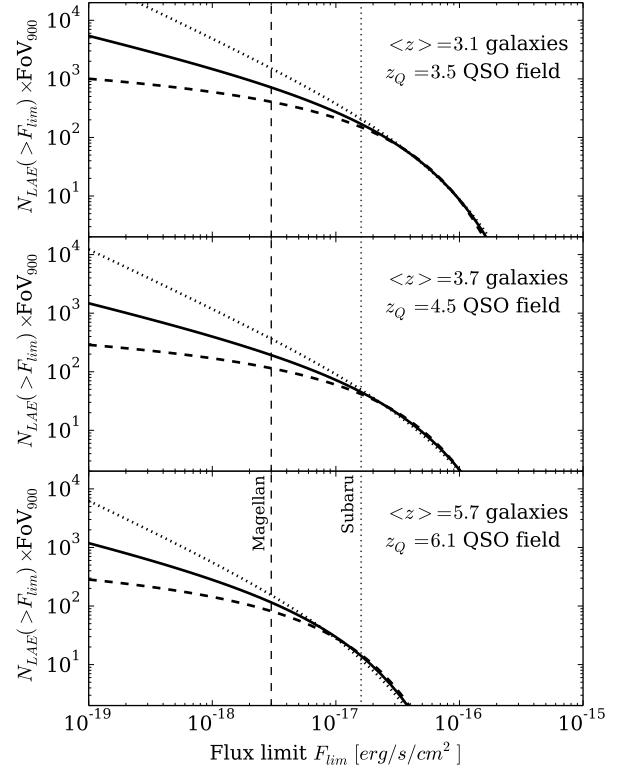


Figure 20. The expected number counts of LAEs in redshift galaxy survey in candidate QSO fields for $3 < z < 7$. y-axis shows the LAE number counts in a $30 \times 30 \text{arcmin}^2$ field-of-view $\text{FoV}_{900} = (\text{FoV}/900 \text{arcmin}^2)$. The vertical lines show the practical flux limit for current survey using Subaru/Supreme-Cam (dotted; Ouchi+2008) and Magellan/IMACS (dashed; Dressler+2014). The LAE luminosity function is taken from the best-fit Schechter function of Ouchi+(2008) with three fixed faint-end slope, $\alpha = -1.0, 1.5, 2.0$ (dashed, solid, dotted lines). We assumed the entire Ly α forest region between Ly α and Ly β lines is covered by the survey.

9.3 Survey strategy

Fig.22 guides us the first estimate of the survey strategy to measure the correlation function between galaxy and absorber. If we employ the present Magellan/IMACS like instrument, it is feasible to measure the correlation function with $\sim 10 - 30\%$ Poisson error by observing the multiple QSO fields (ca. ~ 10 QSO fields) as we expect to see $10 - 100$ pairs per radial bin. The single deep QSO field of futuristic flux limit $10^{19} \text{erg/s/cm}^2$ can give the similar Poisson error.

In the following treatment, we construct the mock survey to study the survey strategy and the feasibility in more detail.

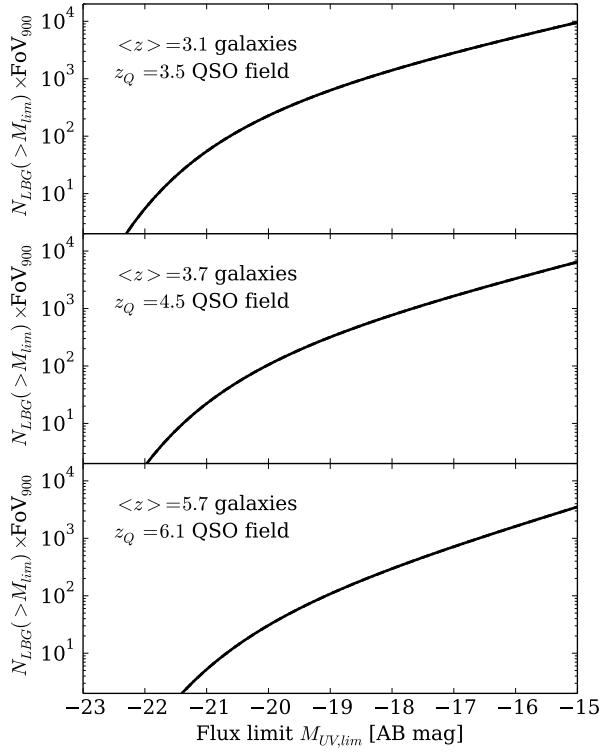


Figure 21. The expected number counts of LBGs in redshift galaxy survey in candidate QSO fields for $3 < z < 7$. y-axis shows the LBG number counts in a $30 \times 30\text{arcmin}^2$ field-of-view $\text{FoV}_{900} = (\text{FoV}/900\text{arcmin}^2)$. The LBG UV luminosity function is taken from the best-fit Schechter function of Bouwens+ (2014). We assumed the entire Ly α forest region between Ly α and Ly β lines is covered by the survey.

10 DISCUSSIONS

10.1 Case for saturating Ly α forests

use Ly-beta forest. or extrapolate low-z to high-z. note that from $z \sim 5$ to $z = 7$ only 400Myr time interval.

One approach is to employ the cosmological hydrodynamical simulations calibrated to match the RSD measurement of galaxy-absorber correlation function. Suppose the simulations are calibrated against $z = 2 - 4$ measurements. Then, the expected evolution of galaxy-absorber phase space dynamics to higher redshift $z > 6$ allow us to perform the simulation-aided measurement of global H $_1$ fraction with breaking the degeneracy with reionization topology.

10.2 Low-ionization metal (ISM) lines

The absorption lines of the low-ionization metals (CIV? etc) is detected for $z = 2 - 3$ galaxies, typically blueward of line centre as interpreted as metal-enriched outflow. ISM lines suppliments the modelling of the intrinsic Ly α line. Self-consistent modelling of both ISM lines and Ly α line would confine the parameter space for outflow of model galaxies. Hence it offers a possibility to break the intrinsic galaxy

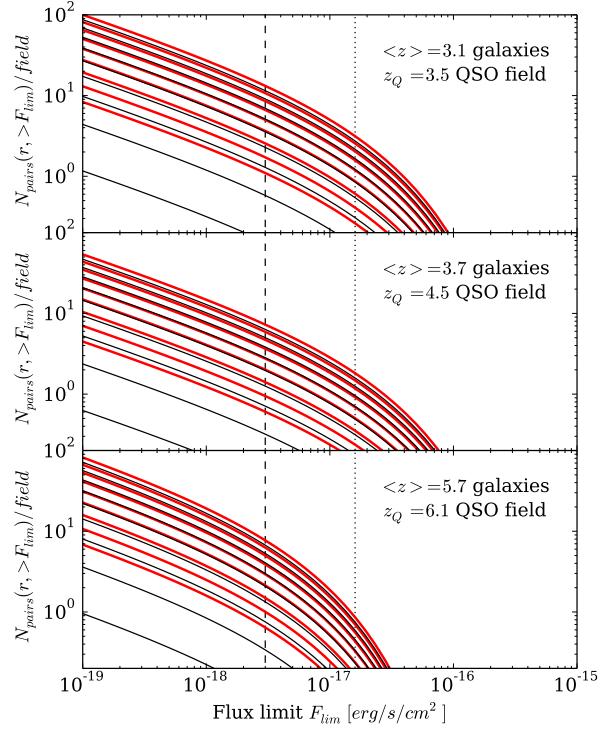


Figure 22. The expected number of LAE-absorber ($N_{\text{HI}} > 10^{19}\text{cm}^{-2}$) pairs per QSO field as a function of flux limit. The red lines are the case with galaxy-absorber clustering and black lines are for no clustering. The pair for each radial bins are $10, 9, 8, 7, 6, 5, 4, 3, 2, 1 h^{-1}\text{cMpc}$ with bin size of $1 h^{-1}\text{cMpc}$ from top to bottom. The vertical lines are Magellan/IMACS like instrument (dashed) and Subaru/Supreme-Cam like instrument (dotted).

property - IGM environment degeneracy on Ly α visibility. If correlation with UV magnitude with Ly α line profile (esp. line shift) the calibration of model with ISM line-detected galaxies would help to peeling into EoR era and more robust measurement of IGM neutral fraction.

10.3 Synergy with 21cm radio interferometry

While we have restricted ourselves in near-infrared observation with telescope such as HST, Subaru, Keck, VLT. When the radio interferometric observations become available, the combined analysis offer a way to explore deeper into the EoR. One possible way is to utilize 21cm forests, which unlike Ly α forests provides the unsaturated spectrum encoding the small-scale IGM structures. Hence the absorber-galaxy redshift-space distortion may be measured by galaxies and 21cm forests. Of course, the availability of number of 21cm forest sightlines is still questionable due to the lack of high-redshift radio-loud objects. Nonetheless, it is worthwhile to keep our mind open for this opportunity.

ACKNOWLEDGMENTS

K.K. thanks to ... Chi-Ting Chiang for discussion

E. Komatsu for Cosmology Routine Library and J. Carlson, B. Reid, M. White for CLPT and GSRSD code publicly available, from which the analytical computation routine for this paper is developed. R. Teyssier and RAMSES development team to make code publicly available and user friendly. Some of the calculations in this paper is computed using python mpmath library (?).

REFERENCES

- de Barros S., Schaefer D., Stark D. P., 2014, A&Ap, 563, A81
 Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1994, MNRAS, 267, 927
 Hui L., Gnedin N. Y., Zhang Y., 1997, ApJ, 486, 599
 Reid B. A., White M., 2011, MNRAS, 417, 1913
 Schaefer D., 2003, A&Ap, 397, 527
 Schaye J., 2001, ApJ, 559, 507

APPENDIX A: LAGRANGIAN FORMULATION

The linear theory and the self-similar clustering hypothesis provides the prototype of real-space 2PCF evolution. The pair conservation law is more general, which can be applied to outflow. We employ the Lagrangian picture to illustrate the evolution of 2PCF (Matsubara 2008, Carlson+; White 2014).

Lagrangian displacement field Ψ for a fluid element at \mathbf{q} at some initial time,

$$\mathbf{r} = \mathbf{q} + \Psi(\mathbf{q}, t) \quad (\text{A1})$$

The mass conservation implies

$$1 + \delta(\mathbf{r}) = \int d^3 q [1 + \delta(\mathbf{q})] \delta_D[\mathbf{r} - \mathbf{q} - \Psi(\mathbf{q}, t)] \quad (\text{A2})$$

APPENDIX B: STATISTICAL FORMULATION OF RADIATIVE TRANSFER

We present the statistical formulation of radiative transfer, which is the application and the substantial generalization of the ideas appeared in Parsce+ 1980, Haardt & Madau; Zuo; Meiksin& White; Kakiichi+2012.

The cosmological radiative transfer through the IGM, neglecting the effect of re-emission and scattering along a line-of-sight is described by

$$\frac{1}{c} \frac{\partial I_\nu}{\partial t} + \mathbf{n} \cdot \nabla I_\nu - \frac{H + \mathbf{n} \cdot \nabla \mathbf{v} \cdot \mathbf{n}}{c} \nu \frac{\partial I_\nu}{\partial \nu} + 3 \frac{H}{c} I_\nu = -\sigma_\alpha n_{\text{HI}} \varphi_\nu I_\nu \quad (\text{B1})$$

where I_ν is the specific intensity, \mathbf{v} is the peculiar velocity, \mathbf{n} is the unit direction vector of rays, $\sigma_\alpha = 0.011 \text{ cm}^2 \text{ Hz}$ is the Ly α cross section, φ_ν is the line profile of Ly α resonance line. $\mathbf{n} \cdot \nabla \mathbf{v} \cdot \mathbf{n}$ term is the Doppler shift effect. The solution to the cosmological radiative transfer equation is

$$I_\nu = \frac{I_\nu(z_s)}{(1+z_s)^3} e^{-\tau_\alpha(\nu_e, z_s)} \quad (\text{B2})$$

where the optical depth τ_α is given by

$$\tau_\alpha(\nu_e, z_s) = \int_0^{z_s} dz' \left| \frac{dl_p}{dz'} \right| n_{\text{HI}}(z') \varphi_\nu \left[T(z'), \nu_e \left(\frac{1+z'}{1+z_s} \right) \left(1 - \frac{v(z')}{c} \right) \right] \quad (\text{B3})$$

We are interested in the statistical properties of the radiation field I_ν . The probability distribution function $P(I_\nu) dI_\nu$ provides the probability that an radiating source is observed with the specific intensity I_ν . The average specific intensity for all the sources $\langle I_\nu \rangle = \int I_\nu P(I_\nu) dI_\nu$ corresponds to stacking the observed spectra of galaxies or QSOs.

Suppose a sphere of radius R centred at the observer that is sufficiently large to contain the source. In the picture that the IGM consists of absorbers of column density $N_{\text{HI},i}$ and redshift

Suppose a line-of-sight skewer of length R . The optical depth along a line-of-sight is

$$\tau(\nu_e) = \sigma_\alpha \int_0^R dr' n_{\text{HI}}(r') \varphi(r') = \sigma_\alpha \sum_{i=1}^N N_{\text{HI},i} \varphi(r_i) \Theta(r_i - R) \quad (\text{B4})$$