

## FOREGROUND CONTAMINATION IN LY $\alpha$ INTENSITY MAPPING DURING THE EPOCH OF REIONIZATION

YAN GONG<sup>1</sup>, MARTA SILVA<sup>2,3</sup>, ASANTHA COORAY<sup>1</sup>, AND MARIO G. SANTOS<sup>2,3,4</sup>

<sup>1</sup>Department of Physics & Astronomy, University of California, Irvine, CA 92697

<sup>2</sup>CENTRA, Instituto Superior Tecnico, Technical University of Lisbon, Lisboa 1049-001, Portugal

<sup>3</sup>Department of Physics, University of Western Cape, Cape Town 7535, South Africa and

<sup>4</sup>SKA SA, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa

*Draft version March 3, 2014*

### ABSTRACT

The intensity mapping of Ly $\alpha$  emission during the epoch of reionization (EoR) will be contaminated by foreground emission lines from lower redshifts. We calculate the mean intensity and the power spectrum of Ly $\alpha$  emission at  $z \sim 7$ , and estimate the uncertainties according to the relevant astrophysical processes. We find that the low-redshift emission lines from 6563 Å H $\alpha$ , 5007 Å [OIII] and 3727 Å [OII] will be strong contaminants on the observed Ly $\alpha$  power spectrum. We make use of both the star formation rate (SFR) and luminosity functions (LF) to estimate the mean intensity and power spectra of the three foreground lines at  $z \sim 0.5$  for H $\alpha$ ,  $z \sim 0.9$  for [OIII] and  $z \sim 1.6$  for [OII], as they will contaminate the Ly $\alpha$  emission at  $z \sim 7$ . The [OII] line is found to be the strongest. We analyze the masking of the bright survey pixels with a foreground line above some line intensity threshold as a way to reduce the contamination in an intensity mapping survey. We find that the foreground contamination can be neglected if we remove pixels with fluxes above  $1.4 \times 10^{-20}$  W/m<sup>2</sup>.

*Subject headings:* cosmology: theory - diffuse radiation - intergalactic medium - large-scale structure of universe

### 1. INTRODUCTION

The epoch of reionization (EoR) is an important and largely unconstrained stage in the evolution of our universe (Barkana & Loeb 2001). The first stars and galaxies can start ionization of the neutral intergalactic medium (IGM) as early as  $z \sim 20$  or even before while completion of the ionization process should end much later, at a redshift as low as  $z \sim 6$ . This is suggested by the observations of the absorption of the Ly $\alpha$  emission from quasars by the neutral hydrogen in the IGM (Fan et al. 2006), i.e. the Gunn-Peterson effect (Gunn & Peterson 1965). Most of the fundamental stones of this era are still poorly known, such as the history, the formation and evolution of the stars and galaxies, and the formation of large scale structure (LSS).

One method to study the EoR is through 21-cm emission of neutral hydrogen in the IGM (Furlanetto et al. 2006), which provides a direct method to study the ionization history of the IGM. However, the 21-cm measurement can not directly trace the galaxy distribution or reflect the star formation. It also suffers from foreground contamination at several orders of magnitude higher than the cosmological 21-cm signal. These disadvantages restrict the 21-cm measurements to explore the LSS and the processes of the formation and evolution for stars and galaxies during the EoR.

A complementary measurement to 21-cm observations is the intensity mapping of the emission lines. This approach is more closely connected to the galaxy properties. The emission lines generated from the stars and molecular or atomic gas in galaxies are more sensible probes of the metallicity and star formation of galaxies. They trace the galaxy distribution at large scales, providing information about the LSS. Several emission lines from galaxies have been proposed to study the EoR (Visbal & Loeb

2010), e.g. CO rotational lines (Righi et al. 2008; Gong et al. 2011; Carilli 2011; Lidz et al. 2011), [CII] fine structure line (Gong et al. 2012), and the Ly $\alpha$  line (Silva et al. 2013; Pullen et al. 2013). Even molecular hydrogen H<sub>2</sub> in the pre-reionization era can be used to study the formation of first stars and galaxies (Gong et al. 2013).

The intensity mapping technique is a powerful tool to explore large areas at poor spatial or angular resolution. It does not attempt to resolve individual sources, but measures the cumulative emission from all sources. Thus, intensity mapping of emission lines from galaxies provides a suitable way to study the statistical properties of galaxies during the EoR in an acceptable survey timescale. However, intensity maps are easily contaminated by other emission lines from lower redshifts in the same observed frequency range. Since intensity mapping cannot resolve individual sources, it is important to independently identify and eliminate the contamination from the foreground lines.

In this work, we study the foreground contamination on Ly $\alpha$  intensity maps of the EoR. We first compute the Ly $\alpha$  emission from both galaxies and the IGM due to recombination and collision processes. We calculate the Ly $\alpha$  mean intensity and power spectrum at  $z \sim 7$ . We find the 6563 Å H $\alpha$ , 5007 Å [OIII] and 3727 Å [OII] are the strongest contamination on the Ly $\alpha$  observations. We estimate the mean intensities of these three lines using the star formation rate (SFR) derived from observations and simulations. Also, for comparison, we make use of luminosity functions (LFs) from different observations to calculate the mean intensities and derive their anisotropy power spectra. We then discuss ways to remove contamination by masking the bright pixels of the intensity maps.

The paper is organized as follows: in the next Section, we estimate the Ly $\alpha$  emission from both galaxies and the IGM, and calculate the mean intensity and power spectrum during the EoR. In Section 3, we explore the intensity and power spectra of the foreground emission lines of the H $\alpha$ , [OII] and [OIII] around  $z = 0.5, 0.9$  and  $1.6$  that can contaminate the Ly $\alpha$  emission at  $z \sim 7$ , and derive their mean intensity and power spectra. In Section 4, we discuss the projection effect of the foreground power spectra and investigate the methods of removing the foreground contamination from low-redshift emission lines. We summarize our results in Section 5. We assume the flat  $\Lambda$ CDM with  $\Omega_M = 0.27$ ,  $\Omega_b = 0.046$ , and  $h = 0.71$  for the calculation throughout the paper (Komatsu et al. 2011).

## 2. THE LY $\alpha$ INTENSITY AND ANISOTROPY POWER SPECTRUM DURING THE EOR

In this Section, we estimate the mean intensity and power spectrum of the Ly $\alpha$  emission during the EoR. The processes that originate Ly $\alpha$  in galaxies and in the IGM are mainly recombinations and collisions. The other processes, e.g. gas cooling by the falling of the IGM gas into the potential wells of the dark matter, and continuum emission by stellar, free-free, free-bound and two-photon emission, can be safely neglected according to previous studies (Cooray et al. 2012; Silva et al. 2013).

### 2.1. The Ly $\alpha$ mean intensity

Following Silva et al. (2013), we estimate the luminosity of the Ly $\alpha$  emission from galaxies by the recombination and collision processes as

$$L_{\text{rec}}^{\text{gal}}(M, z) = 1.55 \times 10^{42} f_{\text{Ly}\alpha}(z) [1 - f_{\text{esc}}^{\text{ion}}(M, z)] \times \frac{\text{SFR}(M, z)}{M_{\odot} \text{ yr}^{-1}} \text{ (erg s}^{-1}\text{)}, \quad (1)$$

$$L_{\text{coll}}^{\text{gal}}(M, z) = 4.03 \times 10^{41} f_{\text{Ly}\alpha}(z) [1 - f_{\text{esc}}^{\text{ion}}(M, z)] \times \frac{\text{SFR}(M, z)}{M_{\odot} \text{ yr}^{-1}} \text{ (erg s}^{-1}\text{)}, \quad (2)$$

respectively<sup>1</sup>. Here the  $f_{\text{Ly}\alpha} = 10^{-3} \times C_{\text{dust}}(1+z)^{\xi}$  is the fraction of the Ly $\alpha$  photons which are not absorbed by the dust in the galaxy, where  $C_{\text{dust}} = 3.34$  and  $\xi = 2.57$  (Hayes et al. 2011). Note that this fitting formula is mainly derived from low-redshift sources and extrapolated to high redshifts ( $z > 6$ ). The  $f_{\text{Ly}\alpha}$  can be much lower at  $z \sim 7$  (e.g. Stark et al. 2010, 2011). The  $f_{\text{esc}}^{\text{ion}} = \exp[-\alpha(z)M^{\beta(z)}]$  is the escape fraction of the ionizing photons, where  $M$  is the halo mass, and  $\alpha = 5.18 \times 10^{-3}$  and  $\beta = 0.244$  (Razoumov & Sommer-Larsen 2010; Silva et al. 2013). The  $\text{SFR}(M, z)$  is the star formation rate, which is parameterized as

$$\text{SFR}(M, z = 7) = 1.6 \times 10^{-26} M^a \left(1 + \frac{M}{c_1}\right)^b \times \left(1 + \frac{M}{c_2}\right)^d \left(1 + \frac{M}{c_3}\right)^e, \quad (3)$$

<sup>1</sup> Note that the ratio  $L_{\text{coll}}^{\text{gal}}/L_{\text{rec}}^{\text{gal}} = 0.26$  can be larger at high redshifts due to strong galaxy forming process (Laursen et al. 2013). This ratio is still observationally unconstrained, and also has large uncertainty suggesting by simulations.

where  $a = 2.59$ ,  $b = -0.62$ ,  $c_1 = 8 \times 10^8 M_{\odot}$ ,  $c_2 = 7 \times 10^9 M_{\odot}$ ,  $c_3 = 10^{11} M_{\odot}$ ,  $d = 0.4$  and  $e = -2.25$ . This SFR parameterization was derived to fit the properties of the simulated galaxies catalogs from De Lucia et al. (2007) and Guo et al. (2011), and is available at  $z = 7$  in (Silva et al. 2013). Then the integrate Ly $\alpha$  mean intensity from individual galaxies is given by

$$\bar{I}_{\text{gal}}(z) = \int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn}{dM} \frac{L_{\text{gal}}(M, z)}{4\pi D_L^2} y(z) D_A^2, \quad (4)$$

where  $L^{\text{gal}} = L_{\text{rec}}^{\text{gal}} + L_{\text{coll}}^{\text{gal}}$  is the total luminosity of Ly $\alpha$  emission from galaxies,  $dn/dM$  is the halo mass function (Sheth & Tormen 1999), and  $D_L$  and  $D_A$  are the luminosity and comoving angular diameter distance respectively. The factor  $y(z) = dr/d\nu = \lambda_{\text{Ly}\alpha}(1+z)^2/H(z)$ , where  $r(z)$  is the comoving distance at  $z$ ,  $\lambda_{\text{Ly}\alpha} = 1216 \text{ \AA}$  is the Ly $\alpha$  wavelength in the rest frame, and  $H(z)$  is the Hubble parameter. We take  $M_{\text{min}} = 10^8 h^{-1} M_{\odot}$  and  $M_{\text{max}} = 10^{13} h^{-1} M_{\odot}$ , which denote the mass range of the halos that host the galaxies with the Ly $\alpha$  emission in this work.

For the Ly $\alpha$  emission from the IGM, we use the luminosity density to estimate the mean intensity

$$\bar{I}_{\text{IGM}} = \frac{l^{\text{IGM}}(z)}{4\pi D_L^2} y(z) D_A^2, \quad (5)$$

where  $l^{\text{IGM}} = l_{\text{rec}}^{\text{IGM}} + l_{\text{coll}}^{\text{IGM}}$  is the Ly $\alpha$  luminosity density for the IGM in  $\text{erg s}^{-1} \text{cm}^{-3}$ . The  $l_{\text{rec}}^{\text{IGM}}$  and  $l_{\text{coll}}^{\text{IGM}}$  are the luminosity densities from recombination and collision processes, respectively. The  $l_{\text{rec}}^{\text{IGM}}$  can be estimated by  $l_{\text{rec}}^{\text{IGM}} = \epsilon_{\text{Ly}\alpha}^{\text{rec}} h \nu_{\text{Ly}\alpha}$ , where  $\epsilon_{\text{Ly}\alpha}^{\text{rec}}$  is the Ly $\alpha$  recombination emission rate per  $\text{cm}^3$ , which is given by

$$\epsilon_{\text{Ly}\alpha}^{\text{rec}} = f_{\text{Ly}\alpha}^{\text{rec}} n_e n_{\text{HII}} \alpha_{\text{B}}^{\text{rec}}. \quad (6)$$

The  $n_e$  and  $n_{\text{HII}}$  are the number density of the electron and HII. Here we have  $n_e n_{\text{HII}} = C_{\text{IGM}}(z) \bar{n}_e(z) \bar{n}_{\text{HII}}(z)$ , where  $\bar{n}_e(z)$  and  $\bar{n}_{\text{HII}}(z)$  are the mean number density of the electron and HII at  $z$ , which are dependent on the reionization fraction  $x_i$  during the EoR. The  $C_{\text{IGM}}(z) = \langle n^2 \rangle / \langle n \rangle^2$  is the clumping factor of the IGM at  $z$  which can be set by  $C_{\text{IGM}} = 6$  at  $z=7$  (Pawlik et al. 2009; Silva et al. 2013). The  $\alpha_{\text{B}}^{\text{rec}}$  is the hydrogen case B recombination coefficient which can be fitted by (Hummer 1994; Seager et al. 1999)

$$\alpha_{\text{B}}^{\text{rec}}(T) = 10^{-13} \frac{aT_4^b}{1 + cT_4^d} \text{ (cm}^3/\text{s)}, \quad (7)$$

where  $a = 4.309$ ,  $b = -0.6166$ ,  $c = 0.6703$ ,  $d = 0.5300$ , and  $T_4 = T/10^4 \text{ K}$ . The  $f_{\text{Ly}\alpha}^{\text{rec}}$  is the the fraction of the Ly $\alpha$  photons produced in the case B recombination, and we use the fitting formula from Cantalupo et al. (2008) to evaluate it as

$$f_{\text{Ly}\alpha}^{\text{rec}}(T) = 0.686 - 0.106 \log_{10}(T_4) - 0.009 T_4^{-0.44}. \quad (8)$$

This formula is accurate to 0.1% for  $100 < T < 10^5 \text{ K}$ . We assume the mean gas temperature of the IGM to be  $1.5 \times 10^4 \text{ K}$  in this work, which is in a good agreement with the results from simulations and the current measurements from quasars (Theuns et al. 2002;

Tittley et al. 2007; Trac et al. 2008; Bolton et al. 2010, 2012).

The Ly $\alpha$  collisional emission in the IGM involves free electrons that collide with the neutral hydrogen atoms (HI) and transfer their kinetic energy by exciting the HI to high energy levels. The hydrogen atoms then decay by emitting photons, including Ly $\alpha$  photons. During the EoR, this process mainly occurs in the ionizing fronts of the reionization bubbles, since this emission will be the strongest when  $n_e \sim n_{\text{HI}}$ . Similar to the recombination emission, the luminosity density of the collisional emission in the IGM can be calculated by  $l_{\text{coll}}^{\text{IGM}} = \epsilon_{\text{Ly}\alpha}^{\text{coll}} h\nu_{\text{Ly}\alpha}$ . The  $\epsilon_{\text{Ly}\alpha}^{\text{coll}}$  is the collisional emission rate per  $\text{cm}^3$

$$\epsilon_{\text{Ly}\alpha}^{\text{coll}} = C_{\text{Ly}\alpha}^{\text{eff}} n_e n_{\text{HI}}, \quad (9)$$

where  $n_{\text{HI}}$  is the number density of the neutral hydrogen atoms. We assume  $n_e n_{\text{HI}} = f_{\text{front}}^{\text{ion}} C_{\text{IGM}}(z) \bar{n}_e(z) \bar{n}_{\text{HI}}(z)$ , where  $\bar{n}_{\text{HI}}(z)$  is the mean number density of neutral hydrogen at  $z$ . The  $f_{\text{front}}^{\text{ion}}$  is the volume fraction of the bubble ionizing fronts. Because the ionizing front just takes a very small part of the whole ionizing bubble, as a good approximation, the volume fraction of the ionizing front can be estimated as  $f_{\text{front}}^{\text{ion}} = \frac{4}{3}\pi(r_2^3 - r_1^3)$ , where  $r_1 = [(3/4\pi)x_i]^{1/3}$ , and  $r_2 = r_1 + d$  and  $d = \frac{1}{6}[n_{\text{H}}\sigma_{\text{H}}(\langle\nu\rangle)]^{-1}$  is the thickness of the bubble ionizing front. The factor  $1/6$  is derived from the integration of  $x_{\text{HI}}(1 - x_{\text{HI}})$  over the ionizing front where  $x_{\text{HI}}(r)$  is the fraction of the HI in the front at radius  $r$  (Cantalupo et al. 2008). Here  $\sigma_{\text{H}}$  denotes the cross section of the hydrogen ionization, where  $\langle\nu\rangle$  is the mean frequency of the ionizing photons, and it is given by (Osterbrock 1989)

$$\sigma_{\text{H}}(\nu) = \frac{\sigma_0}{1 - \exp(-2\pi/\epsilon)} \left[ \frac{\nu_0}{\nu} \exp\left(1 - \frac{\tan^{-1}\epsilon}{\epsilon}\right) \right]^4. \quad (10)$$

Here  $\nu_0$  is the Lyman limit frequency,  $\sigma_0 = 6.3 \times 10^{-18} \text{ cm}^2$  and  $\epsilon \equiv (\nu/\nu_0 - 1)^{1/2}$ . Here we take  $\langle\nu\rangle = 2.4\nu_0$  in our calculation (Gould & Weinberg 1996). Note that we set  $f_{\text{front}}^{\text{ion}} = 0$  when  $x_i \geq 1$ , which means there is no collisional emission in the IGM when the universe is totally ionized.

In Eq. (9), the  $C_{\text{Ly}\alpha}^{\text{eff}}$  is the effective collisional excitation coefficient, which can be estimated by (Cantalupo et al. 2008)

$$C_{\text{Ly}\alpha}^{\text{eff}} = C_{1,2p} + C_{1,3s} + C_{1,3d}. \quad (11)$$

Here we take into account the excitation up to  $n = 3$  energy level to decay and produce Ly $\alpha$  photons. The contribution of emission from higher energy levels can be neglected at  $z \approx 7$  given that the gas temperature and density of the IGM are on average too low to excite these levels. The collisional excitation rate  $C_{1,u}$ , in  $\text{cm}^3/\text{s}$  per second, is given by

$$C_{1,u} = \frac{8.629 \times 10^{-6}}{g_l \sqrt{T}} \gamma_{1,u}(T) \exp\left(\frac{-\Delta E_{1,u}}{kT}\right) \quad (\text{cm}^3/\text{s}), \quad (12)$$

where  $T$  is the gas temperature that we assume  $T = 1.5 \times 10^4 \text{ K}$ ,  $\Delta E_{1,u}$  is the energy difference between lower level and higher level,  $g_l$  is the statistic weight for lower level, and  $\gamma_{1,u}(T)$  is the effective collision strength calculated

by the fitting formulae from Giovanardi et al. (1987).

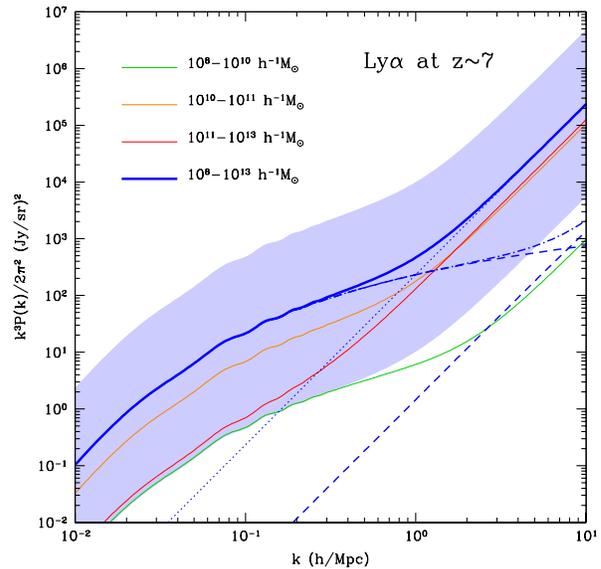


FIG. 1.— The power spectrum of the Ly $\alpha$  emission at  $z = 7$ . The dash-dotted line is the clustering power spectrum which consists of 1-halo and 2-halo components in dashed lines, the dotted line denotes the shot-noise power spectrum, and the total power spectrum is shown in solid line. The uncertainty of the power spectrum is shown in shaded region which is estimated by the uncertainties of the  $f_{\text{esc}}^{\text{ion}}$ ,  $f_{\text{Ly}\alpha}$ , SFR and IGM clumping factor. The contributions of different halo mass scales to the total power spectrum are also shown. We find  $10^{10}\text{-}10^{11} h^{-1} M_{\odot}$  dominates the clustering power spectrum.

Then we estimate the total mean intensity of the Ly $\alpha$  emission  $\bar{I}_{\text{Ly}\alpha} = \bar{I}_{\text{gal}} + \bar{I}_{\text{IGM}}$  using Eq. (4) and Eq. (5). At  $z \sim 7$ , we get  $\bar{I}_{\text{gal}} = 9.2 \text{ Jy/sr}$  and  $\bar{I}_{\text{IGM}} = 1.2 \text{ Jy/sr}$  if assuming  $\bar{x}_i = 0.85$ . Thus, according to the assumptions we have made, Ly $\alpha$  emission from galaxies is larger than that from the IGM around  $z = 7$ . Note that the collisional emission in the IGM is much smaller than the recombination emission with  $\bar{l}_{\text{coll}}^{\text{IGM}} \simeq 0.1 \text{ Jy/sr}$  and  $\bar{l}_{\text{rec}}^{\text{IGM}} \simeq 1.1 \text{ Jy/sr}$ , respectively, assuming  $T = 1.5 \times 10^4 \text{ K}$  for the gas temperature of the IGM.

There are large uncertainties for the parameters in the estimation of the Ly $\alpha$  intensity. According to the current measurements (Razoumov & Sommer-Larsen 2010; Hayes et al. 2011; Blanc et al. 2011), we find the  $f_{\text{esc}}^{\text{ion}}$  and  $f_{\text{Ly}\alpha}$  in Eq. (1) and (2) have an uncertainty of a factor of 2 each. Besides, we take into account of the uncertainties in the SFR and the clumping factor, which can be factors of 3 and 5 at  $z \sim 7$  respectively (Silva et al. 2013). Eventually, we find  $\bar{I}_{\text{Ly}\alpha} = 10.4_{-9.0}^{+39.3} \text{ Jy/sr}$  around

<sup>2</sup> This  $f_{\text{Ly}\alpha}$  is Ly $\alpha$  “effective” escape fraction instead of real escape fraction which should include the scattered Ly $\alpha$  emission from the diffuse halos surrounding galaxies (Dijkstra & Jeason-Daniel 2013). The Ly $\alpha$  emission from the diffuse halos has low surface brightness, but the total flux can exceed the direct emission from galaxies by a significant factor (e.g. Zheng et al. 2010; Steidel et al. 2010; Matsuda et al. 2012). Thus the real Ly $\alpha$  escape fraction can be much larger than the effective escape fraction. This diffused emission could contribute to the intensity mapping. We find our fiducial  $f_{\text{Ly}\alpha} \simeq 0.7$  at  $z = 7$  which is close to one, and we also add an uncertainty of a factor of 2 on it. So the effect of the diffused Ly $\alpha$  emission is safely covered in our calculation.

$z = 7$ . Our results are well consistent with other works when allowing for uncertainty (e.g. Silva et al. 2013; Pullen et al. 2013). Besides, we note that the estimated Ly $\alpha$  emission of high- $z$  metal-poor galaxies with standard case-B assumption can be underestimated by a factor of  $\sim 2 - 3$  compared to accurate prediction (Raiter et al. 2010). This possible departure can be covered by the intensity uncertainty we consider.

### 2.2. The Ly $\alpha$ power spectrum

According to the calculations above, the absolute value of the Ly $\alpha$  intensity background during the EoR is small and hard to measure directly in an absolute intensity experiment. Instead, we can try to measure the fluctuations of the Ly $\alpha$  intensity and estimate the anisotropy power spectrum. Since the Ly $\alpha$  emission from galaxies and IGM in the ionization bubbles surrounding the galaxies trace the underlying matter density field, we can calculate the Ly $\alpha$  intensity fluctuations by

$$\delta I_{\text{Ly}\alpha} = \bar{b}_{\text{Ly}\alpha} \bar{I}_{\text{Ly}\alpha} \delta(\mathbf{x}). \quad (13)$$

Here we set  $\bar{I}_{\text{Ly}\alpha} \simeq \bar{I}_{\text{gal}}$ , since the main Ly $\alpha$  emission at  $z \sim 7$  comes from galaxies. The  $\delta(\mathbf{x})$  is the matter over-density at the position  $\mathbf{x}$ , and  $\bar{b}_{\text{Ly}\alpha}(z)$  is the average galaxy bias weighted by the Ly $\alpha$  luminosity

$$\bar{b}_{\text{Ly}\alpha}(z) = \frac{\int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} L_{\text{gal}}^{\text{Ly}\alpha} b(M, z)}{\int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} L_{\text{gal}}^{\text{Ly}\alpha}}, \quad (14)$$

where  $L_{\text{gal}}^{\text{Ly}\alpha}(M, z)$  is the Ly $\alpha$  luminosity of the galaxy, and  $b(M, z)$  is the halo bias (Sheth & Tormen 1999). Then we can calculate the Ly $\alpha$  clustering power spectrum due to galaxy clustering as

$$P_{\text{Ly}\alpha}^{\text{clus}}(k, z) = \bar{b}_{\text{Ly}\alpha}^2 \bar{I}_{\text{Ly}\alpha}^2 P_{\delta\delta}(k, z), \quad (15)$$

where  $P_{\delta\delta}(k, z)$  is the matter power spectrum which can be estimated from the halo model (Cooray & Sheth 2002). The clustering power spectrum dominates the fluctuations at large scales. At small scales the Poisson noise caused by the discrete distribution of galaxies becomes important. This Poisson or shot-noise power spectrum takes the form (e.g. Visbal & Loeb 2010; Gong et al. 2011):

$$P_{\text{Ly}\alpha}^{\text{shot}}(z) = \int_{M_{\min}}^{M_{\max}} dM \frac{dn}{dM} \left[ \frac{L_{\text{gal}}^{\text{Ly}\alpha}}{4\pi D_L^2} y(z) D_A^2 \right]^2. \quad (16)$$

Therefore, the total power spectrum can be written by  $P_{\text{Ly}\alpha}^{\text{tot}}(k, z) = P_{\text{Ly}\alpha}^{\text{clus}}(k, z) + P_{\text{Ly}\alpha}^{\text{shot}}(z)$ . In Figure 1, we show the Ly $\alpha$  power spectrum at  $z = 7$ . We show the total, clustering and shot-noise power spectrum in solid, dash-dotted and dotted line, respectively. The dashed lines denote the 1-halo and 2-halo terms of the clustering power spectrum. The uncertainty of the total power spectrum is also shown in shaded region, which is derived from the uncertainties of the  $f_{\text{esc}}^{\text{ion}}$ ,  $f_{\text{Ly}\alpha}$ , SFR and IGM clumping factor. We also show the contributions of different halo mass scales to the total power spectrum in colored solid lines. As can be seen, the halos with mass  $10^{10}$ - $10^{11} h^{-1} M_{\odot}$  provide the most contribution on the clustering power spectrum, and the halos

with higher masses ( $10^{11}$ - $10^{13} h^{-1} M_{\odot}$ ) have large shot-noise since they are bright and rare. We find  $\bar{b}_{\text{Ly}\alpha} = 5.0$  and  $P_{\text{Ly}\alpha}^{\text{shot}} = 4.6 \times 10^3 (\text{Jy/sr})^2 (\text{Mpc/h})^{-3}$  at  $z = 7$ , and the shot-noise power spectrum dominates the total power spectrum at  $k \gtrsim 1 \text{ Mpc}^{-1} h$ .

## 3. THE ESTIMATION OF THE FOREGROUND EMISSION LINES

Since we can not resolve individual sources with intensity mapping, the measurements of the Ly $\alpha$  emission during the EoR can be contaminated by emission lines from lower redshifts. Here we consider three low-redshift emission lines, H $\alpha$  at 6563 Å, [OIII] at 5007 Å and [OII] at 3727 Å. At  $z \sim 7$ , the frequency of the Ly $\alpha$  line is about 300 THz, which can be then contaminated by H $\alpha$  at  $z \sim 0.5$ , [OIII] at  $z \sim 0.9$  and [OII] at  $z \sim 1.6$ , respectively. We will estimate the mean intensity and power spectra of these lines in this Section.

### 3.1. The mean intensity from the SFR

The H $\alpha$  line at 6563 Å, [OIII] line at 5007 Å and [OII] line at 3727 Å are good tracers of the SFR of galaxies. The luminosity of these lines is related to the SFR as

$$\text{SFR} (M_{\odot} \text{yr}^{-1}) = (7.9 \pm 2.4) \times 10^{-42} L_{\text{H}\alpha}, \quad (17)$$

$$\text{SFR} (M_{\odot} \text{yr}^{-1}) = (1.4 \pm 0.4) \times 10^{-41} L_{[\text{OII}]}, \quad (18)$$

$$\text{SFR} (M_{\odot} \text{yr}^{-1}) = (7.6 \pm 3.7) \times 10^{-42} L_{[\text{OIII}]}. \quad (19)$$

The luminosity here is in  $\text{erg s}^{-1}$ , and the relations for H $\alpha$  and [OII] are from Kennicutt (1998), and the [OIII] relation is given in Ly et al. (2007). The SFR- $L_{\text{H}\alpha}$  relation assumes the initial mass function from Salpeter (1955), and we add 30% uncertainty to it. These conversions are also in good agreement with other works<sup>3</sup> (e.g. Hopkins et al. 2003; Wijesinghe et al. 2011; Drake et al. 2013). With the help of the relation of the SFR and the halo mass as a function of redshift  $\text{SFR}(M, z)$ , we can use these conversions to derive  $L(M, z)$  and compute the mean intensities for these lines using Eq. (4).

To estimate the  $\text{SFR}(M, z)$ , we can make use of the  $\text{SFRD}(z)$  from the observations and assume the SFR is proportional to the halo mass  $M$ . We take the  $\text{SFRD}(z)$  given by Hopkins & Beacom (2006) with the fitting formula from Cole et al. (2001)

$$\text{SFRD}(z) = \frac{a + bz}{1 + (z/c)^d} (h M_{\odot} \text{yr}^{-1} \text{Mpc}^{-3}), \quad (20)$$

where  $a = 0.0118$ ,  $b = 0.08$ ,  $c = 3.3$  and  $d = 5.2$ . This fitting formula is consistent with the observational data very well, especially at  $z \lesssim 2$ , which is good enough for our estimations here. Then assuming

$$\text{SFR}(M, z) = f_*(z) \frac{\Omega_b}{\Omega_M} \frac{1}{t_s} M, \quad (21)$$

where  $t_s = 10^8 \text{ yr}$  is the typical star formation timescale, and  $f_*(z)$  is the normalization factor which can be determined by  $\text{SFRD}(z) = \int dM \frac{dn}{dM} \text{SFR}(M, z)$ . After obtaining  $\text{SFR}(M, z)$ , we can derive  $L(M, z)$  for the H $\alpha$ , [OII]

<sup>3</sup> The SFR- $L_{[\text{OII}]}$  relation has a larger uncertainty, and we will discuss it in Section 3.2.

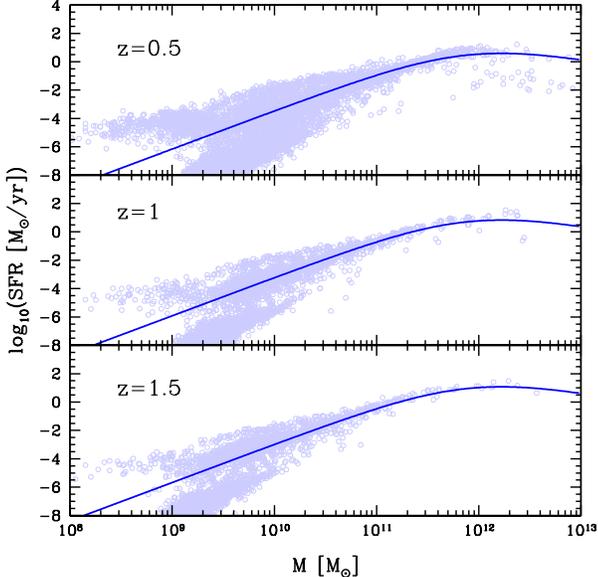


FIG. 2.— The SFR vs. halo mass  $M$  at  $z = 0.5$ ,  $z = 1$  and  $z = 1.5$  from the simulations. The dots are from the galaxy catalog in the simulations of Guo et al. (2011) and the curves are the best fits of the SFR( $M$ ).

and [OIII] lines using Eq. (17), (18) and (19) respectively, and then calculate their mean intensities with Eq. (4). We find  $\bar{I}_{\text{H}\alpha} = 31.2 \pm 9.4$  Jy/sr,  $\bar{I}_{[\text{OII}]} = 17.5 \pm 5.0$  Jy/sr and  $\bar{I}_{[\text{OIII}]} = 35.0 \pm 17.0$  Jy/sr. The errors are derived from the uncertainties in the conversions of the SFR and luminosity given by Eq. (17), (18) and (19).

Another way to obtain the SFR( $M, z$ ) is with simulations. Here we use the galaxy catalog from Guo et al. (2011), which is obtained by Millennium II simulation with a volume of  $100 (\text{Mpc}/h)^3$  and particle mass resolution  $\sim 6.9 \times 10^6 h^{-1} M_{\odot}$  (Boylan-Kolchin et al. 2009). We fit the SFR( $M, z$ ) derived from this catalog for  $z \lesssim 2$  with

$$\text{SFR}(M, z) = 10^{a+bz} \left(\frac{M}{M_1}\right)^c \left(1 + \frac{M}{M_2}\right)^d, \quad (22)$$

where  $a = -9.097$ ,  $b = 0.484$ ,  $c = 2.7$ ,  $d = -4.0$ ,  $M_1 = 10^8 M_{\odot}$  and  $M_2 = 8 \times 10^{11} M_{\odot}$ . The simulation results and the best fits of the SFR( $M, z$ ) at  $z = 0.5$ ,  $z = 1$  and  $z = 1.5$  are shown in Figure 2.

We find that the SFR increases quickly with halo mass but becomes flatter for the large halos with  $M \gtrsim 10^{12} M_{\odot}$ . Also, there is relatively large scattering in the relation at small halo mass with  $M \lesssim 10^{10} M_{\odot}$ , which could provide additional uncertainty in the intensity calculation. The SFRD derived from this SFR( $M, z$ ) are  $2.2 \times 10^{-2}$ ,  $3.4 \times 10^{-2}$  and  $6.6 \times 10^{-2} M_{\odot} \text{yr}^{-1} \text{Mpc}^{-3}$  at  $z \sim 0.5$ ,  $z \sim 0.9$  and  $z \sim 1.6$ , respectively. For comparison, Hopkins & Beacom (2006) based values are  $3.7 \times 10^{-2}$ ,  $5.9 \times 10^{-2}$  and  $9.7 \times 10^{-2} M_{\odot} \text{yr}^{-1} \text{Mpc}^{-3}$  at the corresponding redshifts. Thus the SFRDs from the simulation are lower than the SFRDs from the observations. Then we estimate the  $L(M, z)$  and obtain  $\bar{I}_{\text{H}\alpha} = 16.8 \pm 5.0$  Jy/sr,  $\bar{I}_{[\text{OII}]} = 9.0 \pm 2.6$  Jy/sr and  $\bar{I}_{[\text{OIII}]} = 16.6 \pm 8.1$  Jy/sr using Eq. (17), (18), (19) and Eq. (4). These values are lower than what we have from

TABLE 1  
THE MEAN INTENSITIES IN JY/SR OF THE  $\text{H}\alpha$ , [OIII] AND [OII] AROUND  $z = 0.5$ ,  $z = 0.9$  AND  $z = 1.6$ , DERIVED FROM BOTH OF THE LF AND SFR METHODS.

	$\bar{I}_{\text{H}\alpha}(z \sim 0.5)$	$\bar{I}_{[\text{OIII}]}(z \sim 0.9)$	$\bar{I}_{[\text{OII}]}(z \sim 1.6)$
SFRD <sub>obs</sub>	$31.2 \pm 9.4$	$35.0 \pm 17.0$	$17.5 \pm 5.0$
SFR <sub>sim</sub>	$16.8 \pm 5.0$	$16.6 \pm 8.1$	$9.0 \pm 2.6$
LF <sub>L07</sub>	$12.2^{+17.4}_{-7.1}$	$11.8^{+11.1}_{-5.7}$	$23.5^{+8.3}_{-5.9}$
LF <sub>D13</sub> <sup>1</sup>	8.4	12.7	24.2
LF <sub>D13</sub> <sup>2</sup>	23.0	6.0	50.6
LF <sub>G13</sub>	17.6	—	—

— L07, D13 and G13 denote Ly et al. (2007), Drake et al. (2013) and Gunawardhana et al. (2013).

the observational SFRD( $z$ ), but they are still consistent within  $1\sigma$  error.

### 3.2. The mean intensity from the luminosity functions

A direct way to estimate the mean intensities of the  $\text{H}\alpha$ , [OII] and [OIII] lines is to make use of observed LFs. In Figure 3, we show the observed LFs of  $\text{H}\alpha$ , [OII] and [OIII] lines around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$  that can contaminate the Ly $\alpha$  emission from  $z \sim 7$ . The squares, pentagons, circles and triangles are LFs from Ly et al. (2007)<sup>4</sup>, Drake et al. (2013) and Gunawardhana et al. (2013), respectively. The LFs data points shown here are dust extinction-corrected, except for the LFs from Drake et al. (2013) which are lower than the results of the other two works.

The LF is usually fitted by the Schechter function (Schechter 1976)

$$\Phi(L)dL = \phi_* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-\frac{L}{L_*}\right) \frac{dL}{L_*}, \quad (23)$$

where  $\phi_*$ ,  $L_*$  and  $\alpha$  are free parameters that are obtained by fitting the Schechter function with the data. The LFs from Ly et al. (2007) and Drake et al. (2013) are fitted by this function<sup>5</sup>. Then the mean intensity can be estimated by

$$\bar{I}_{\nu}(z) = \int_{L_{\min}}^{L_{\max}} dL \frac{dn}{dL} \frac{L}{4\pi D_L^2} y(z) D_A^2, \quad (24)$$

where  $dn/dL = \Phi(L)$  is the luminosity function, and  $L_{\min} = 10^5 L_{\odot}$  and  $L_{\max} = 10^{12} L_{\odot}$  are the lower and upper luminosity limits. We find the result is not changed if choosing smaller  $L_{\min}$  and larger  $L_{\max}$  in our calculation. We use the extinction-corrected LFs to calculate the mean intensity for the  $\text{H}\alpha$ , [OII] and [OIII] lines. Note that, in Drake et al. (2013), they just correct the  $L_*$  for the dust extinction to get the extinction-corrected LFs, and the other two parameters  $\phi_*$  and  $\alpha$  are still fixed to the values fitted by the observed LFs.

Using Eq. (24) and the extinction-corrected LFs, we calculate the mean intensities of the  $\text{H}\alpha$ , [OIII] and [OII]

<sup>4</sup> We note that the  $\text{H}\alpha$  and [OII] LFs from Ly et al. (2007) are well consistent with recent observations with larger survey volumes (e.g. Sobral et al. 2012, 2013). So we just show the LF data from Ly et al. (2007) here.

<sup>5</sup> The  $\text{H}\alpha$  LF at  $0.24 < z < 0.35$  from Gunawardhana et al. (2013) agrees well with the LF at  $z = 0.4$  from Ly et al. (2007) at the faint end, but has flatter slope at the bright end. So it is fitted by an alternative/different function with different form given by Saunders et al. (1990).

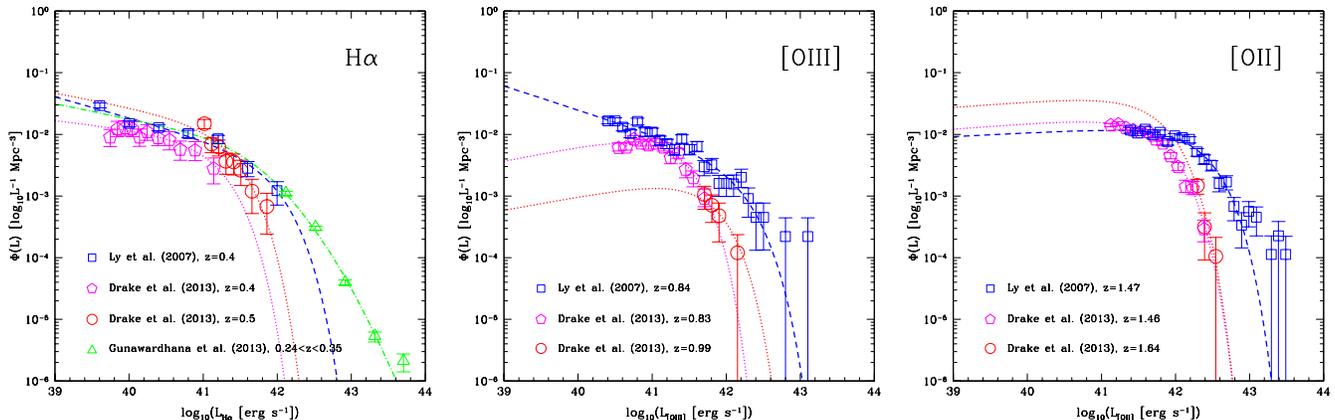


FIG. 3.— The luminosity functions and the best fits for the H $\alpha$ , [OIII] and [OII] lines around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$  that could contaminate the Ly $\alpha$  emission at  $z \sim 7$ . The blue squares, pink pentagons, red circles and green triangles are the observational LFs from Ly et al. (2007), Drake et al. (2013) and Gunawardhana et al. (2013), respectively. The dashed and dotted curves are the best fits of the corresponding LFs. Note that the data points of LFs from Drake et al. (2013) are the observed LFs (without dust extinction correction) which seem lower than the results of the other two works.

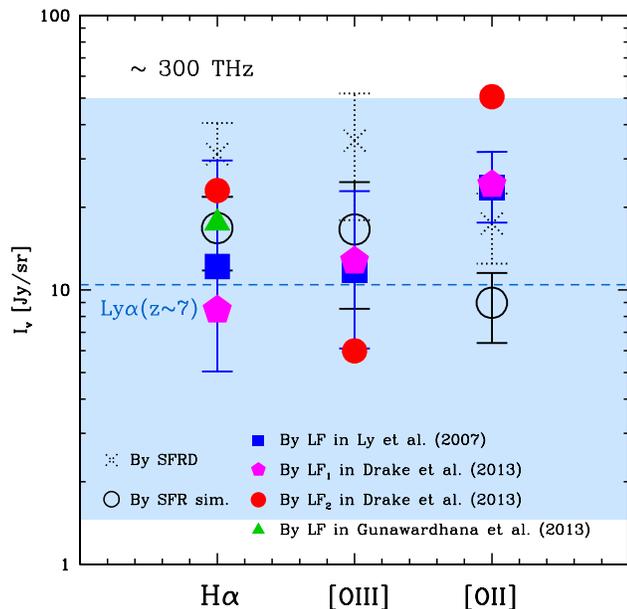


FIG. 4.— The mean intensity of H $\alpha$ , [OIII] and [OII] using both SFR and LF methods around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$ . The dashed line denotes the mean intensity of Ly $\alpha$  at  $z \sim 7$  with uncertainties (shaded region) derived in the last section. The results from these two methods are basically consistent with each other. The mean intensities of the H $\alpha$ , [OII] and [OIII] are generally higher than the central value of the Ly $\alpha$  mean intensity, which can provide considerable contamination on the Ly $\alpha$  emission.

around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$ . We list and plot the mean intensities from both of the LF and SFR methods in Table 1 and Figure 4 for comparison. We also estimate the errors of the LFs from Ly et al. (2007) based on the errors of the fitted values of the  $\phi_*$ ,  $L_*$  and  $\alpha$ , and then derive the error for the mean intensity. In Table 1, the SFRD<sub>obs</sub> and SFR<sub>sim</sub> denote the methods of SFRD from observations and SFR from the simulations respectively. The LF<sub>L07</sub>, LF<sub>D13</sub> and LF<sub>G13</sub> denote the LF method using the LFs from Ly et al. (2007), Drake et al. (2013) and Gunawardhana et al. (2013) respectively. The LF<sub>D13</sub><sup>1</sup> and LF<sub>D13</sub><sup>2</sup> denote to use the

LFs from Drake et al. (2013) shown in pink pentagons and red circles in Figure 3 respectively. In Figure 4, the dotted crosses and open circles are the values for SFRD<sub>obs</sub> and SFR<sub>sim</sub>, and the blue squares, pink pentagons, red circles and green triangles are the values for LF<sub>L07</sub>, LF<sub>D13</sub><sup>1</sup>, LF<sub>D13</sub><sup>2</sup> and LF<sub>G13</sub> respectively. For comparison, we also shown the mean intensity of the Ly $\alpha$  at  $z \sim 7$  (blue dashed line) with uncertainties (shaded region) derived in the last section.

For the mean intensity of the H $\alpha$ , we find the results from the different LF observations are consistent with each other, and they are safely in the  $1\sigma$  error of the result from LF<sub>L07</sub>. Comparing the results from the SFR and LF methods, the mean intensity from SFR<sub>sim</sub> is in a good agreement with that from LF<sub>L07</sub>, LF<sub>D13</sub><sup>2</sup> and LF<sub>G13</sub> which give an average H $\alpha$  intensity around 15 Jy/sr. The SFRD<sub>obs</sub> and LF<sub>D13</sub><sup>1</sup> here give a bit higher and lower results, respectively.

For the [OIII] line, the result of LF<sub>L07</sub> agrees with the LF<sub>D13</sub><sup>1</sup> result, and they are also consistent with the result from SFR<sub>sim</sub> in  $1\sigma$  error. This gives an average intensity around 13 Jy/sr. The SFRD<sub>obs</sub> gives a higher intensity considering its error. The LF<sub>D13</sub><sup>2</sup> provides a low [OIII] intensity about 6 Jy/sr, and this can be caused by the poor model fitting of the LF data in Drake et al. (2013) (circles in the middle panel of Figure 3).

For the [OII] emission at  $z \sim 1.6$ , the results have larger discrepancy compared to the H $\alpha$  and [OIII] cases. The LF<sub>L07</sub> and LF<sub>D13</sub><sup>1</sup> still give the same intensity which is around 24 Jy/sr, and their results agree with SFRD<sub>obs</sub> in  $1\sigma$ . But the SFR<sub>sim</sub> suggests a low intensity around 9 Jy/sr, while the LF<sub>D13</sub><sup>2</sup> gives a much higher value of  $\sim 50$  Jy/sr. It can be the same reason for the LF<sub>D13</sub><sup>2</sup> which gives a different result from the others as in the [OIII] case, that there are not enough LF data for the fitting of the LF. In the right panel of Figure 3, we find there are just three data points around  $2 \times 10^{42}$  erg s<sup>-1</sup> and no data is observed in the faint end (see red circles). For the SFR<sub>sim</sub>, it actually has large discrepancy between different observations for the SFR- $L_{[OII]}$  relation. For example, Kewley et al. (2004) proposes another relation

$$\text{SFR} (M_{\odot} \text{yr}^{-1}) = (6.58 \pm 1.65) \times 10^{-42} L_{[OII]},$$

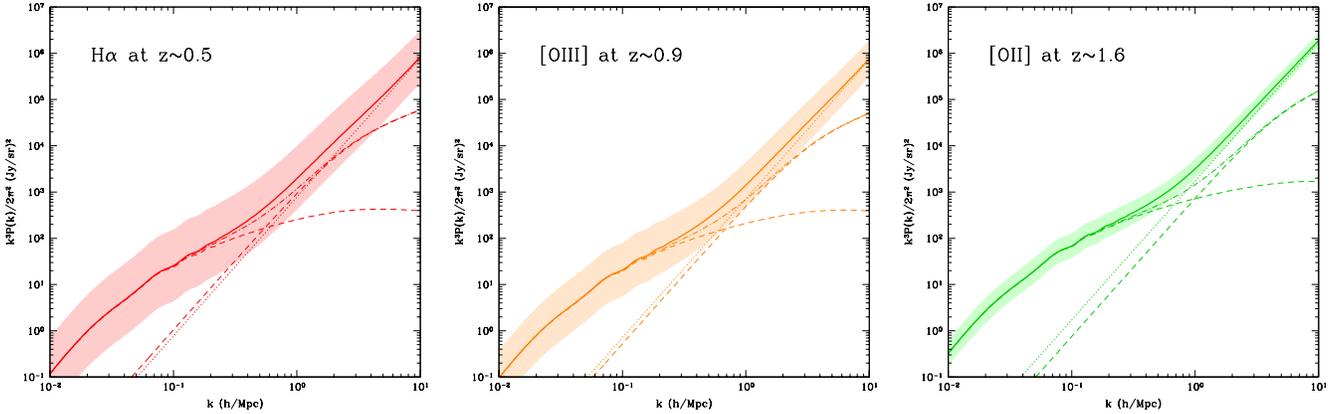


FIG. 5.— The power spectra of the H $\alpha$ , [OIII] and [OII] lines at  $z \sim 0.5$ ,  $z \sim 0.9$  and  $z \sim 1.6$ , which can contaminate the Ly $\alpha$  emission at  $z \sim 7$ . The solid, dash-dotted and dotted lines denote the total, clustering and shot-noise power spectrum respectively. The short dashed lines are the 1-halo and 2-halo terms from the halo model (Cooray & Sheth 2002). The uncertainty of the total power spectrum is shown in shaded region, which is estimated by the uncertainty of the mean intensity. Here we adopt the LFs and errors in Ly et al. (2007) to estimate the mean intensity and errors for the three emission lines.

which gives a higher [OII] intensity  $\sim 17.4$  Jy/sr using the SFR<sub>sim</sub> method. This substantially reduces the tension between the SFR<sub>sim</sub> and the other methods.

Generally, we find that the intensities of the H $\alpha$ , [OII] and [OIII] around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$  are larger than that of the Ly $\alpha$  at  $z \sim 7$ , which provide considerable contamination on the Ly $\alpha$  emission from the EoR. Also, the result derived by the LFs from Ly et al. (2007) is well consistent with the other results from both of the LF and SFR method, hence we would adopt it as the foreground line contamination in our following estimation and discussion.

Next, we calculate the anisotropy power spectra of the H $\alpha$ , [OII] and [OIII] emissions using their mean intensities. Using Eq. (15) and (16) and replacing the  $\bar{b}_{\text{Ly}\alpha}$ ,  $\bar{L}_{\text{Ly}\alpha}$  and  $L_{\text{Ly}\alpha}$  to be the bias, mean intensity and line luminosity of the H $\alpha$ , [OII] and [OIII], we obtain their clustering and shot-noise power spectra. Note that we assume the line luminosity is proportional to the halo mass to get the mean bias of these three lines, i.e. replacing  $L_{\text{Ly}\alpha}$  to be  $M$  in Eq. (14). This approximation is good enough to estimate the mean bias given the uncertainty in the mean intensity. In Figure 5, we show the power spectrum of the H $\alpha$ , [OIII] and [OII] lines at  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$ , which contaminate the Ly $\alpha$  emission at  $z \sim 7$ . The total, clustering and shot-noise power spectra are in solid, dash-dotted and dotted lines respectively. The 1-halo and 2-halo terms from the halo model are also shown in short dashed lines (Cooray & Sheth 2002). The uncertainty of the total power spectrum in shaded region is estimated by the uncertainty of the mean intensity. As mentioned, we use the LFs and errors from Ly et al. (2007) to compute the power spectrum and the uncertainty for the three emission lines.

#### 4. THE REMOVAL OF THE FOREGROUND EMISSION LINE CONTAMINATION

##### 4.1. The observed power spectrum

So far we have analyzed the expected intensity from contaminating lines. In practice, experiments (at least first generation ones) will try to make a statistical measurement of the 3-d power spectrum. Therefore, this is the quantity that we should compare to in order to access

the level of foreground contamination. Because the signal and foregrounds will be emitted at different redshifts there will be an extra factor multiplying the foreground power spectrum. In order to calculate this, we need to take into account the effect of the observed light cone in the 3-D power spectrum.

The intensity that we measure,  $I(\Omega, \nu)$  will correspond to a sum of the signal we are interested in,  $I_s$  (the Ly $\alpha$  emission from redshift  $z_s$ ), and the foreground emission,  $I_f$ , from lower redshifts  $z_f$ , which is contributing to the same frequency:

$$I(\Omega, \nu) = I_s(\Omega, \nu) + I_f(\Omega, \nu). \quad (25)$$

In the flat sky approximation, displacements in angle  $\Delta\theta$  and frequency  $\Delta\nu$  about a central reference point, can be transformed into a position in 3-D comoving space

$$\begin{aligned} x_1, x_2 &= r(z)\Delta\theta \\ x_3 &= -y(z)\Delta\nu, \end{aligned} \quad (26)$$

where we are already making a translation along the  $x_3$  direction and  $x_1, x_2$  are assumed perpendicular to the line of sight while  $x_3$  is taken to be parallel to it. Note that, besides the flat sky, we are also neglecting cosmic evolution, e.g. assuming that points along the line of sight are all emitted at the same redshift  $z$ .

We then write the observed signal as

$$\begin{aligned} I(\Omega, \nu) &= \bar{I}_s(z_s)[1 + b_s(z_s)\delta(z_s, \mathbf{x}_s)] \\ &+ \bar{I}_f(z_f)[1 + b_f(z_f)\delta(z_f, \mathbf{x}_f)] \end{aligned} \quad (27)$$

where  $\delta()$  is the dark matter perturbation and  $\mathbf{x}_s, \mathbf{x}_f$  is the 3-D position of the signal and foreground emission respectively. If we Fourier transform the above signal with respect to  $\mathbf{x}_s$ , the corresponding foreground Fourier mode will be offset with respect to the true one. The observed power spectrum,  $P_{\text{obs}}$  will then be:

$$\begin{aligned} P_{\text{obs}}(k_{\perp}, k_{\parallel}) &= \bar{I}_s^2(z_s)b_s^2(z_s)P(z_s, k_s) \\ &+ \bar{I}_f^2(z_f)b_f^2(z_f)\left(\frac{r_s}{r_f}\right)^2\left(\frac{y_s}{y_f}\right)P(z_f, k_f), \end{aligned} \quad (28)$$

where  $|\vec{k}_s| = \sqrt{k_{\perp}^2 + k_{\parallel}^2}$  is the 3-D  $k$  at the redshift of

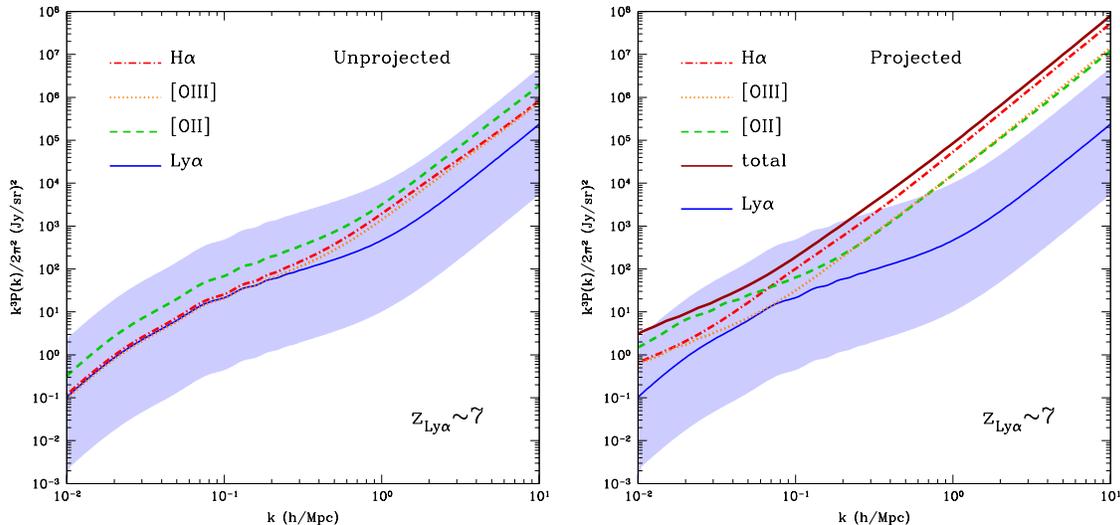


FIG. 6.— *Left*: the comparison of the 3-D total power spectrum of the Ly $\alpha$  at  $z \sim 7$  to that of the H $\alpha$  at  $z = 0.5$ , [OIII] at  $z = 0.9$  and [OII] at  $z = 1.6$ . The uncertainty of the Ly $\alpha$  power spectrum is shown in shaded region. *Right*: the same as the left panel, but considering the projection effect on the power spectra of the H $\alpha$ , [OII] and [OIII] lines. The total projected power spectrum is shown in dark red line.

the signal, and  $|\vec{k}_f| = \sqrt{(r_s/r_f)^2 k_\perp^2 + (y_s/y_f)^2 k_\parallel^2}$  is the 3-D  $k$  at the redshift of the foreground line. The factor  $(r_s/r_f)^2 (y_s/y_f)$  comes from the distortion of the volume element when Fourier transforming the foreground correlation function to the power spectrum (Visbal & Loeb 2010). This process can be considered to project the foreground power spectrum to the high redshift where the intensity signal comes from. We then see that, even when looking at the spherical averaged power spectrum, e.g.  $P_{\text{obs}}(k = \sqrt{k_\perp^2 + k_\parallel^2})$ , the foreground contribution to the power spectrum will be boosted by this projection. Moreover, there is an anisotropy in the  $\mathbf{k}$  space which provides a potential method to distinguish the foregrounds from the signal. We discuss the details of this effect in the Appendix.

In Figure 6, we compare the 3-D power spectrum of H $\alpha$ , [OII] and [OIII] with and without the projection. In the left panel of Figure 6, we show the total power spectrum of the H $\alpha$  at  $z = 0.5$ , [OIII] at  $z = 0.9$  and [OII] at  $z = 1.6$  without projection. The  $P_{\text{Ly}\alpha}^{\text{tot}}$  and its uncertainty are also shown in blue line and shaded region for comparison. In the right panel of Figure 6, we show the 3-D projected power spectra of the H $\alpha$ , [OII] and [OIII] lines to the redshift of the Ly $\alpha$  emission ( $z \sim 7$ ). To estimate foreground 3-D projected power spectrum from Eq. (28), we simply assume  $k_1 = k_2 = k_\parallel$ , where  $\mathbf{k} = k_1 \hat{\mathbf{i}} + k_2 \hat{\mathbf{j}} + k_\parallel \hat{\mathbf{k}}$  and  $k_\perp = \sqrt{k_1^2 + k_2^2}$ . We also ignore the redshift distortion here. We find the projected power spectra of the three lines become a bit higher than the case where we ignore the projection of the foregrounds to the background cosmological signal of intensity during the EoR.

#### 4.2. Foregrounds masking

To remove the contamination from these foreground emission lines, we need to mask the bright sources at low redshifts. However, we cannot identify the individual sources and their redshifts in the intensity mapping, and

all the signals that lie in the same survey pixel, which is defined by the spectral and angular resolutions, will be mixed together and observed as only one signal.

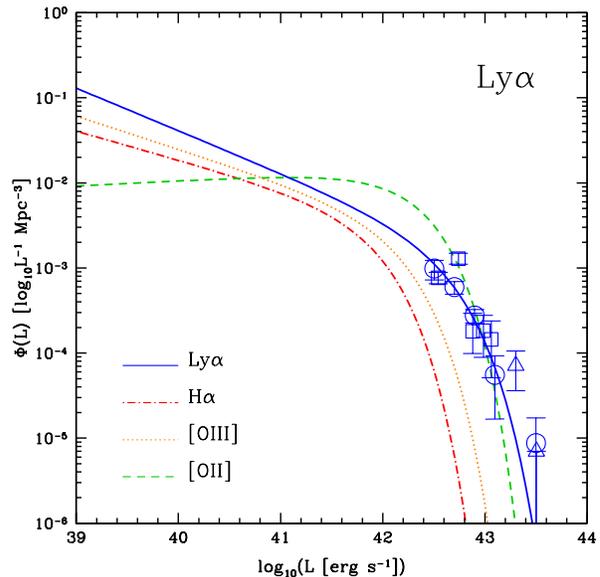


FIG. 7.— The LF of the Ly $\alpha$  at  $z \sim 7$  (actually around  $z=6.6$ ) from the observations of the LAEs. The circles, squares and triangles are the LF data points from Ouchi et al. (2010), Kashikawa et al. (2006) and Hu et al. (2005), respectively. The blue solid curve is the best fit of the LF data using the fitting results given in Ouchi et al. (2010). For comparison, the LFs of the H $\alpha$ , [OIII] and [OII] around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$  from Ly et al. (2007) are also shown.

During the EoR, the galaxies are averagely smaller and fainter than the present galaxies, so their LFs should be relatively higher at the faint end than the galaxies at the low redshifts. In Figure 7, we show the Ly $\alpha$  LF at  $z \sim 7$  from the observations of the Ly $\alpha$  emitters (LAEs) (Ouchi et al. 2010; Kashikawa et al. 2006;

Hu et al. 2005)<sup>6</sup>. The best fit of the LF is shown in blue solid line using the Schechter function (see Eq. (23)) with fitting values given by Ouchi et al. (2010). The LFs of the H $\alpha$ , [OIII] and [OII] lines are also shown for comparison. We can see that, the LFs of the Ly $\alpha$  is higher than the others at the faint end, but basically lower than the [OII] LF at the bright end. Therefore, we can mask the whole bright pixels whose fluxes are above some threshold in the intensity mapping to reduce the contamination of low-redshift line emission, and then the remainder pixels should be dominated by the Ly $\alpha$  emission at high redshifts.

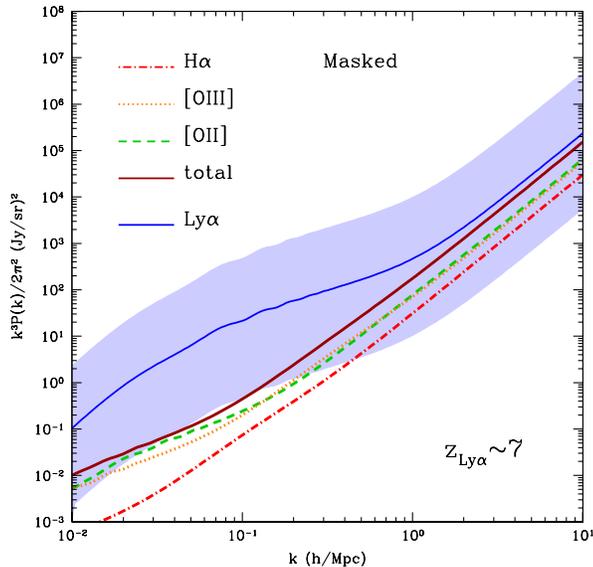


FIG. 8.— The power spectra of the H $\alpha$ , [OII] and [OIII] with masking and projection. The deep red line denotes the total foreground power spectrum after the masking. We apply a flux cut at  $1.4 \times 10^{-20} \text{ W/m}^2$  here, which could make the total foreground power spectrum  $\sim 100$  times smaller than the Ly $\alpha$  around  $k = 0.1 \text{ h Mpc}^{-1}$ .

In Figure 8, the power spectra of the H $\alpha$ , [OII] and [OIII] lines with masking and projection are shown. Here we try to make the total power spectrum of H $\alpha$ , [OII] and [OIII] smaller by a factor of  $\sim 100$  around  $k = 0.1 \text{ h Mpc}^{-1}$  where the shot-noise is small. We find we need to mask sources with fluxes greater than  $1.4 \times 10^{-20} \text{ W/m}^2$ . The corresponding line luminosities are  $L_{\text{H}\alpha} \simeq 3.5 \times 10^6 L_{\odot}$ ,  $L_{[\text{OII}]} \simeq 6.3 \times 10^7 L_{\odot}$  and  $L_{[\text{OIII}]} \simeq 1.5 \times 10^7 L_{\odot}$  respectively. Note that these luminosity cuts are close to the lower luminosity limit ( $L_{\text{min}} = 10^5 L_{\odot}$ ) we take in Eq. (24), hence our masking results are dependent on the faint-end slopes of the LFs which are not well constrained by observational data.

In Figure 9, we show the number of sources whose flux is greater than some value in a survey volume pixel. Here

<sup>6</sup> Using Eq. (24) and the Ly $\alpha$  LF (blue solid curve in Figure 7), we note that the Ly $\alpha$  mean intensity is just  $\sim 1 \text{ Jy/sr}$  around  $z = 7$ , which is lower than the value of  $\sim 9 \text{ Jy/sr}$  from the galaxy we get in Section 2 (see also Pullen et al. 2013). This can be caused by the absence of the Ly $\alpha$  LF data at the faint end, since only very bright individual sources can be detected at this high redshift. This also means that there could be more faint LAEs around  $z = 7$ , and the real LF could be higher at the faint end.

we assume a survey with  $6'' \times 6''$  beam size and 190–330 THz frequency range with R=40 frequency resolution. As can be seen, the H $\alpha$  sources are dominant at the bright end, but are subdominant in the survey since all of these sources are from low redshifts ( $z \sim 0.5$ ) where the survey volume is small. On the other hand, [OII] sources are the main contamination, and are brighter than  $1.4 \times 10^{-20} \text{ W/m}^2$  in approximately 2% of the pixels. In total, we need to mask about 3% survey pixels. Here we should notice the large uncertainty of the Ly $\alpha$  emission. The number of the Ly $\alpha$  sources can be lower by a factor of  $\sim 10$  with the uncertainty, which would lead to a larger masking percentage.

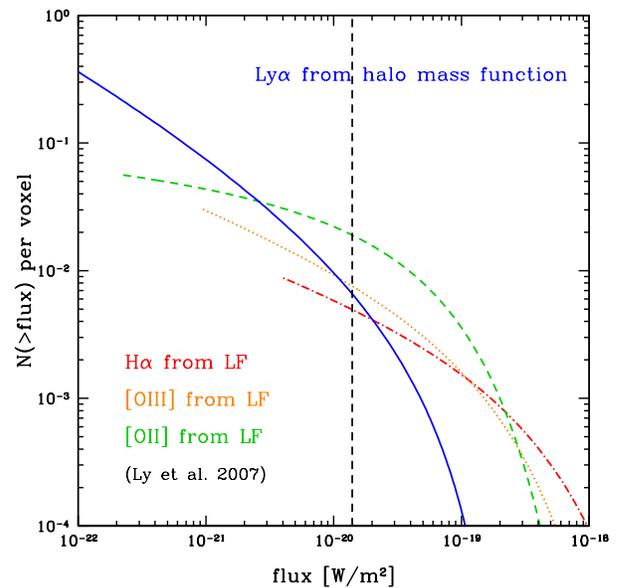


FIG. 9.— The number of sources per volume pixel whose fluxes are greater than the x-axis flux value. The Ly $\alpha$  curve is derived from the halo mass function and the calculation in Section 2, and the curves of H $\alpha$ , [OII] and [OIII] are computed by the LFs in Ly et al. (2007). The vertical dashed line denotes the flux cut at  $1.4 \times 10^{-20} \text{ W/m}^2$ . We find this corresponds to remove 3% of the total pixels.

There is a second method that can be used to eliminate the foreground line contamination, which is making use of the cross-correlation between different emission lines at the same redshift. This method is discussed in many previous works and can be used either for the signal (e.g. Visbal & Loeb 2010; Gong et al. 2012; Silva et al. 2013) or the foreground contamination (e.g. Pullen et al. 2013). The idea is that the lines emitted at the same redshift should trace the same underlying matter distribution and hence lead to large cross-correlation on the power spectrum. On the other hand, the emissions at different redshifts, which are far away from each other, would not provide considerable cross power spectrum. Therefore, we can derive the auto power spectrum of one line if we know both the cross power spectrum and the auto power spectrum of the other line. For instance, we can cross-correlate the Ly $\alpha$  line with 21-cm at  $z \sim 7$ , and then the foreground contamination would be reduced and we can estimate the  $P_{\text{Ly}\alpha}$  from the  $P_{\text{Ly}\alpha \times 21\text{cm}}$  and  $P_{21\text{cm}}$  (Silva et al. 2013). Or we can estimate the auto power spectra of the foreground emission lines by cross-

correlating with the 21-cm intensity mapping surveys at  $z \sim 1$ , e.g. the GBT<sup>7</sup>, CHIME<sup>8</sup> and Tianlai projects (Chen 2012). However, although we can detect the cross power spectrum, it is still hard to measure the auto power spectrum for the 21-cm or the other lines due to their own foregrounds removal. So this method is indirect and cannot substitute the masking method discussed above. Even if we can only obtain the cross power spectrum, it could still be a good guide and a secondary check for the intensity mapping experiments.

## 5. SUMMARY

We estimate the Ly $\alpha$  mean intensity and power spectrum during the EoR, and explore the foreground contamination from the low-redshift emission lines. We consider the Ly $\alpha$  emission from both of galaxies and IGM for the recombination and collisional emission processes. We find the Ly $\alpha$  emission of galaxies is dominant over the IGM at  $z \sim 7$  with a total mean intensity  $\bar{I}_{\text{Ly}\alpha} \sim 10$  Jy/sr (about 9 Jy/sr from galaxies and 1 Jy/sr from the IGM). We also evaluate the uncertainty of the mean intensity according to the uncertainties of  $f_{\text{esc}}^{\text{ion}}$ ,  $f_{\text{Ly}\alpha}$ , SFR and the clumping factor. With the help of halo model, we also calculate the Ly $\alpha$  clustering, shot-noise power spectrum with the uncertainty given by the mean intensity.

Next we investigate the foreground contamination by low-redshift emission lines. We find H $\alpha$  at 6563 Å, [OIII] at 5007 Å and [OII] at 3727 Å can be the strongest contamination on Ly $\alpha$  emission during the EoR. We estimate the mean intensity of the H $\alpha$ , [OIII] and [OII] lines at  $z \sim 0.5$ ,  $z \sim 0.9$  and  $z \sim 1.6$  respectively, which are the redshifts that can contaminate the Ly $\alpha$  emission at  $z \sim 7$ . We use two methods to do the estimation, i.e. the SFR and LF methods. In the SFR method, both of the SFRD( $z$ ) from the observations and the SFR( $M, z$ )

derived from the simulations are used to compute the intensity. In the LF method, we adopt the LFs of H $\alpha$ , [OIII] and [OII] around  $z = 0.5$ ,  $z = 0.9$  and  $z = 1.6$  respectively from different observations. We find the results from both methods are basically consistent with each other, especially for the H $\alpha$  and [OIII] lines. The mean intensity of the three lines are  $\bar{I}_{\text{H}\alpha} \sim 15$  Jy/sr,  $\bar{I}_{\text{[OIII]}} \sim 13$  Jy/sr and  $\bar{I}_{\text{[OII]}} \sim 24$  Jy/sr, which are larger than the  $\bar{I}_{\text{Ly}\alpha} \sim 10$  Jy/sr. The results from the LFs in (Ly et al. 2007) is in good agreements with the others, and we adopt their LFs and the errors to calculate the power spectrum and uncertainty for the three foreground lines.

At last, we discuss the methods to remove the foreground contamination due to low-redshift emission lines. We first compare the power spectrum of the Ly $\alpha$  at  $z \sim 7$  to the H $\alpha$ , [OIII] and [OII] power spectrum at low redshifts, and consider the projection effect in the real survey. The power spectrum of the foreground lines become larger after the projection. We then propose to mask the whole bright pixels with foreground emission above some flux threshold to reduce the contamination of the foreground lines in the intensity mapping. We find the contamination can be neglected when the flux cut is  $1.4 \times 10^{-20}$  W/m<sup>2</sup>. The cross-correlation method is helpful to reduce or estimate the contamination, but it is indirect and can not substitute the masking method in the survey of the intensity mapping.

This work was supported by NSF CAREER AST-0645427 and AST-1313319. MBS and MGS acknowledge support by FCT-Portugal under grant PTDC/FIS/100170/2008. MBS acknowledges support by FCT-Portugal under grant SFRH/BD/51373/2011.

## REFERENCES

- Barkana, R., & Loeb, A. 2001, Phys. Rep., 349, 125  
 Blanc, G. A., Adams, J. J., Gebhardt, K., et al. 2011, ApJ, 736, 31  
 Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., & Sargent, W. L. W. 2010, MNRAS, 406, 612-625  
 Bolton, J. S., Becker, G. D., Raskutti, S., et al. 2012, MNRAS, 419, 2880-2892  
 Bouwens, R. J., Ilingworth, G. D., Oesch, P. A., et al. 2012, ApJ, 754, 83  
 Boylan-Kolchin, M., Springel, V., White, S. D. M., et al. 2009, MNRAS, 398, 1150  
 Cantalupo, S., Porciani, C., & Lilly, S. J. 2008, ApJ, 672, 48  
 Carilli, C. L. 2011, ApJ, 730, L30  
 Chen, X. 2013, arXiv:1212.6278  
 Cole, S., Norberg, P., Baugh, C. M., et al. 2001, MNRAS, 326, 255-273  
 Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1  
 Cooray, A., Gong, Y., Smidt, J., & Santos, M. 2012, ApJ, 756, 92  
 De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2  
 Dijkstra, M., & Jeeson-Daniel, A. 2013, MNRAS 435, 3333-3341  
 Drake, A. B., Simpson, C., Collins, C. A., et al. 2013, MNRAS, 433, 796-811  
 Fan, X., Carilli, C., & Keating, B. 2006, ARA&A, 44, 415  
 Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Phys. Rep., 433, 181  
 Giovanardi, C., Natta, A., & Palla, F. 1987, A&AS, 70, 269  
 Gong, Y., Cooray, A., Silva, M. B., Santos, M. G. & Lubin, P. 2011, ApJ, 728, L46  
 Gong, Y., Cooray, A., Santos, M., et al. 2012, ApJ, 745, 49  
 Gong, Y., Cooray, A., & Santos, M. 2013, ApJ, 768, 130  
 Gould, A., & Weinberg, D. H. 1996, ApJ, 468, 462-468  
 Gunawardhana, M. L. P., Hopkins, A. M., Bland-Hawthorn, J., et al. 2013, MNRAS, 433, 2764  
 Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633-1641  
 Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101  
 Hayes, M., Schaerer, D., Ostlin, G., et al. 2011, ApJ, 730, 8  
 Hopkins, A. M., Miller, C. J., Nichol, R. C., et al. 2003, ApJ, 599, 971-991  
 Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142  
 Hu, E. M., Cowie, L. L., Capak, P., & Kakazu, Y. 2005, in IAU Colloq. 199, Probing Galaxies through Quasar Absorption Lines, ed. P. R. Williams, C.-G. Shu, & B. Menard (Cambridge: Cambridge Univ. Press), 363  
 Hummer D. G. 1994, MNRAS, 268, 109  
 Kashikawa, N., Shimasaku, K., Malkan, M. A., et al. 2006, ApJ, 648, 7  
 Kennicutt, R. C. 1998, ARA&A, 36, 189-231  
 Kewley, L. J., Geller, M., & Jansen, R. A. 2004, ApJ, 127, 2002-2030  
 Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18  
 Laursen, P., Duval, F., & Ostlin, G. 2013, ApJ, 766, 124  
 Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70  
 Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738-759

<sup>7</sup> <https://science.nrao.edu/facilities/gbt/>

<sup>8</sup> <http://chime.phas.ubc.ca/>

Matsuda, Y., Yamada, T., Hayashino, T., et al. 2012, MNRAS, 425, 878-883  
 Osterbrock, D. E. 1989, Astrophysics of Gaseous Nebular and Active Galactic Nuclei (Mill Valley, CA: Univ.Sci.)  
 Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2010, ApJ, 723, 869-894  
 Pawlik, A. H., Schaye, J., & van Scherpenzeel, E. 2009, MNRAS, 394, 1812-1824  
 Pullen, A., Dore, O., & Bock, J. 2013, arXiv:1309.2295  
 Raiter, A., Schaerer, D., & Fosbury, R. A. E. 2010, A&A, 523, A64  
 Righi, M., Hernandez-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489  
 Razoumov, A. O., & Sommer-Larsen, J. 2010, ApJ, 710, 1239  
 Salpeter E. E. 1955, ApJ, 121, 161-67  
 Saunders, W., Cannon, R., Sutherland, W. 2004, Anglo-Australian Observatory Epping Newslett., 106, 16  
 Schechter, P. 1976, ApJ, 203, 297  
 Seager, S., Sasselov, D. D., & Scott, D. 1999, ApJ, 523, L1  
 Shimasaku, K., Kashikawa, N., Ota, K., et al. 2006, PASJ, 58, 313  
 Silva, M., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132  
 Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119

Sobral, D., Best, P. N., Matsuda, Y., et al. 2012, MNRAS, 420, 1926-1945  
 Sobral, D., Smail, I., Best, P. N., et al. 2013, MNRAS, 428, 1128-1146  
 Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408, 1628-1648  
 Stark, D. P., Ellis, R. S., & Ouchi, M. 2011, ApJ, 728, L2  
 Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289-322  
 Theuns, T., Schaye, J., Zaroubi, S., et al. 2002, ApJ, 567, L103-L106  
 Tittley, E. R., & Meiksin, A., et al. 2007, MNRAS, 380, 1369-1386  
 Trac, H., Cen, R., & Loeb, A. 2008, ApJ, 689, L81-L84  
 Visbal, E. & Loeb, A. 2010, JCAP, 11, 016  
 Wijesinghe, D. B., Hopkins, A. M., Sharp, R., et al. 2011, MNRAS, 410, 2291-2301  
 Wyithe, J. S. B., & Dijkstra, M. 2011, MNRAS, 415, 3929-3950  
 Zheng, Z., Cen, R., Trac, H., & Miralda-Escude, J. 2010, ApJ, 716, 574  
 Zheng, Z., Cen, R., Trac, H., & Miralda-Escude, J. 2011, ApJ, 726, 38

## APPENDIX

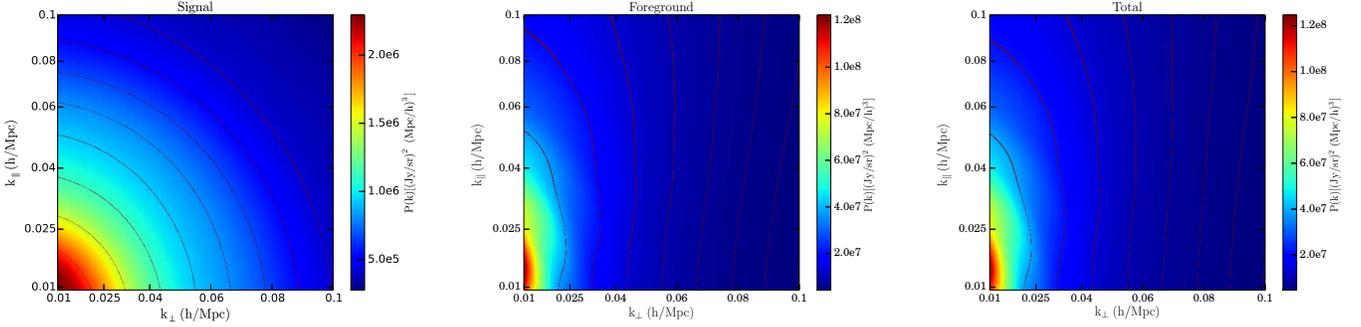


FIG. 10.— The 2-D anisotropic power spectra shown by perpendicular and parallel Fourier modes for the signal ( $\text{Ly}\alpha$ ), foreground (only  $\text{H}\alpha$  here) and total observed signal+foreground. We find the 2-D signal power spectrum is almost symmetric for  $k_{\perp}$  and  $k_{\parallel}$ , and the redshift distortion effect is relatively small. For the foreground power spectrum at low redshift, because the redshift distortion effect is strong and the shift factor on the foreground  $k_{\perp}$  and  $k_{\parallel}$  are different, the shape of the power spectrum is irregular. This effect provides a way to distinguish the signal and the foreground in principle. The total power spectrum is similar with the foreground power spectrum, since the amplitude of the signal power spectrum is much lower than the foreground.

In Eq. (28) of Section 4.1, we see that the observed foreground power spectrum is actually anisotropic, i.e., not simply a function of  $k = \sqrt{k_{\perp}^2 + k_{\parallel}^2}$ . This can in principle be used in the foreground cleaning process. In particular, after the intensity cut, we can check if there is still any strong contamination by looking at the anisotropy of the total power spectrum. In order to check the strength of these anisotropies caused by the projection effect, we calculate the 2-D anisotropic power spectra for the signal and foregrounds. Here we also take into account of the linear redshift-space distortions. In that case, the bias should be replaced by

$$b_s(z_s) \rightarrow b_s(z_s) + \left( \frac{k_{\parallel}}{\sqrt{k_{\perp}^2 + k_{\parallel}^2}} \right)^2 \frac{1}{H(z_s)} \frac{\dot{D}(z_s)}{D(z_s)}$$

$$b_f(z_f) \rightarrow b_f(z_f) + \left( \frac{(y_s/y_f)k_{\parallel}}{\sqrt{(r_s/r_f)^2 k_{\perp}^2 + (y_s/y_f)^2 k_{\parallel}^2}} \right)^2 \frac{1}{H(z_f)} \frac{\dot{D}(z_f)}{D(z_f)}.$$

Here,  $D(z)$  is the growth factor. The normal expression of the second term of the bias is  $f\mu^2$ . Here  $f = d\ln D/d\ln a$  where  $a$  is the scale factor, and  $\mu = \cos\theta$  where  $\theta$  is the angle between the line of sight and the wave-vector  $\mathbf{k}$ . We also notice that the so-called non-gravitational effects can introduce strong redshift-space distortion effect (Zheng et al. 2011; Wyithe & Dijkstra 2011). According to full  $\text{Ly}\alpha$  radiative transfer calculations, the  $\text{Ly}\alpha$  emission is dependent on environment (gas density and velocity) around LAEs. The observed LAE clustering features can be changed by this effect especially at high redshifts. However, this effect relies on “missing”  $\text{Ly}\alpha$  photons scattering in relatively

close proximity to Ly $\alpha$  emission galaxies, which could be recovered by intensity mapping. Therefore we ignore this effect in our discussion.

In Figure 10, we show the 2-D anisotropic power spectrum decomposed into  $k_{\perp}$  and  $k_{\parallel}$  for the signal, foreground and the total observations. We just show the H $\alpha$  foreground here, since it has the lowest redshift among the other two foreground lines and has the largest effect of the power spectrum projection. We find the shape of the signal power spectrum is quite symmetric and the redshift distortion effect is relatively small, since the signal comes from high redshifts. However, for the foreground, the shape of the spectrum is irregular due to the redshift distortion and the different factors on the  $k_{\perp}$  and  $k_{\parallel}$  (i.e.  $r_s/r_f$  on  $k_{\perp}$  and  $y_s/y_f$  on  $k_{\parallel}$ ) in the projection. This effect provides a potential method to distinguish the signal from the foreground, and could be helpful to remove the foreground in future high sensitive experiments.