A Halo Model Approach to the 21 cm and Lyα Cross-correlation
Chang Feng¹, Asantha Cooray¹, and Brian Keating²
¹Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA; chang.feng@uci.edu
²Department of Physics, University of California, San Diego, CA 92093, USA
Received 2017 January 31; revised 2017 June 30; accepted 2017 July 12; published 2017 August 28

Abstract
We present a halo-model-based approach to calculate the cross-correlation between 21 cm H I intensity fluctuations and Lyα emitters (LAE) during the epoch of reionization (EoR). Ionizing radiation around dark matter halos are modeled as bubbles with the size and growth determined based on the reionization photon production, among other physical parameters. The cross-correlation shows a clear negative-to-positive transition, associated with transition from ionized to neutral hydrogen in the intergalactic medium during EoR. The cross-correlation is subject to several foreground contaminants, including foreground radio point sources important for 21 cm experiments and low-z interloper emission lines, such as Hα, O II, and O III for Lyα experiments. Our calculations show that by masking out high fluxes in the Lyα measurement, the correlated foreground contamination on the 21 cm–Lyα cross-correlation can be dramatically reduced. We forecast the detectability of 21 cm–Lyα cross-correlation at different redshifts and adopt a Fisher matrix approach to estimate uncertainties on the key EoR parameters that have not been well constrained by other observations of reionization. This halo-model-based approach enables us to explore the EoR parameter space rapidly for different 21 cm and Lyα experiments.

Key words: dark ages, reionization, first stars – ISM: lines and bands – large-scale structure of universe

1. Introduction
The early universe, initially filled with hot plasma, became neutral as hydrogen ions captured electrons that were decoupled from cosmic microwave background (CMB) photons at a redshift of 1100. A cosmic “dark age” subsequently ensued in the universe until the linear density fluctuations seeded by inflation were amplified, forming the first stars and galaxies (Loeb & Barkana 2001). The X-rays from mini quasars and ultraviolet radiation from the massive stars in first-light galaxies heated and ionized the neutral hydrogen, and the universe gradually transformed from completely neutral to fully ionized during the epoch of reionization (EoR). Today the EoR still remains largely unexplored, as signatures imprinted on the intergalactic medium (IGM) in the early universe are too faint to be detected.

The physical processes present during the EoR are of extreme importance to our understanding of the universe and the structure that formed in it. The clustering of neutral hydrogen (H I) down to the Jeans length scale contains a wealth of information about certain fundamental physics, including dark matter. The H I tomography is not subject to small-scale physical effects such as photon diffusion damping present in the CMB power spectrum. The timing and duration of the EoR can help interpret other cosmological measurements, such as the kinetic Sunyaev–Zel’dovich (kSZ) effect (McQuinn et al. 2005). Moreover, some exotic physics such as primordial magnetic fields (Schleicher et al. 2009) and decaying dark matter (Furlanetto et al. 2006b) could be probed during the EoR. To date, the neutral fraction during the EoR was measured from quasar absorption spectra (Fan et al. 2006) and Lyα-emitting galaxy luminosity functions (Konno et al. 2014; Malhotra & Rhoads 2004, 2006) around z ~ 6. Another important quantity of the EoR, the Thomson scattering optical depth, is constrained to τ = 0.088 ± 0.014 by WMAP (Komatsu et al. 2011) and τ = 0.058 ± 0.012 by Planck satellites (Planck Collaboration et al. 2016).

The best way to measure the H I content prior to and during reionization is through the 21 cm H I fine-structure spin-flip transition. A number of experiments have been targeting the 21 cm emission, such as the Low Frequency Array (LOFAR; van Haarlem et al. 2013), the Murchison Widefield Array (MWA; Tingay et al. 2013), the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons 2010), the Hydrogen Epoch of Reionization Array (DeBoer et al. 2017), and the Square Kilometer Array (SKA; Koopmans et al. 2015). The redshifted 21 cm emission is contaminated by both galactic and extragalactic foregrounds that consist of galactic synchrotron, supernovae remnants, free–free emission, and radio point sources (Furlanetto et al. 2006a). The Galactic synchrotron emission is the dominant contribution, as it is three to four orders of magnitude stronger than the background brightness temperature fluctuations. By performing a component separation or subtracting the 21 cm foreground, the 21 cm brightness fluctuations could be measured (Bonaldi & Brown 2015). This, however, relies on the reliability of the foreground estimation.

The radio point sources are also thought to be another foreground issue for 21 cm experiments, but this signal is very likely to be a subdominant contamination (Liu et al. 2009). The expected 21 cm signal is at the level of 10 mK² at k = 0.3 Mpc⁻¹ (Mesinger et al. 2011), while recent measurements from PAPER set a 2σ upper limit as (22.4 mK)² in the range 0.15 < k < 0.5 h Mpc⁻¹ at z = 8.4 (Ali et al. 2015).

During the EoR the ultraviolet Lyα emission was created by the first stars and galaxies. The Lyα background traces the underlying dark matter distribution and also affects the spin-temperature distribution. By directly measuring the Lyα emissions, we get an additional observable on EoR physics as well (Jensen et al. 2013). However, the Lyα background is contaminated by low-z foregrounds, such as Hα at z = 0.5, O III at z = 0.9, and O II at z = 1.6. These low-z components are much brighter than Lyα, precluding a clean detection. On the other hand, such low-z foregrounds can be easily masked.
out since they are very bright (Pullen et al. 2014; Gong et al. 2014). Therefore, a simple masking procedure would recover the genuine $Ly_\alpha$ background from experiments.

The 21 cm and $Ly_\alpha$ emission is anti-correlated at large angular scales because they originate from IGM and galaxies, respectively, and ionized bubbles around $Ly_\alpha$ galaxies are devoid of H I that is seen with 21 cm experiments. The transition in the cross-correlation from negative to positive indicates a characteristic size for the average of H II regions around halos. Therefore, the cross-correlation between 21 cm and $Ly_\alpha$ can be viewed as a complementary probe of EoR physics. The cross-correlation could be more advantageous in terms of foreground removal as the two sets of mentioned foregrounds would be largely uncorrelated, potentially allowing a higher signal-to-noise detection and an easy confirmation of the EoR signature. Previously, the cross-correlation between 21 cm experiments and galaxies was studied for 21 cm experiments such as MWA and LOFAR, using both analytical and numerical calculations (Furlanetto & Lidz 2007; Lidz et al. 2009; Wyithe & Loeb 2007), as well as for LOFAR and Subaru’s Hyper Suprime-cam (Virbanc et al. 2016). The cross-correlation between 21 cm and CO/kSZ also shows a similar transition in the correlation sign (Lidz et al. 2011; Jelić et al. 2010).

So far, different approaches have been used to model reionization. The large-scale $N$-body and radiative transfer simulations, while desirable, are challenging because it is computationally intensive due to the large dynamic range (Iliev et al. 2006, 2015). Another approach involves semi-analytical/semi-numerical models by taking a halo catalog generated from $N$-body simulations and generating a reionization field by smoothly filtering the halo field (Zahn et al. 2007, 2011). A more simplified idea of this semi-numerical simulation is to make the density field from Gaussian random variables instead of relying on the $N$-body simulations. The simulation can be done efficiently within a small box for EoR (Mesinger & Furlanetto 2007; Mesinger et al. 2011). However, this numerical solution becomes ineffective when the box is too large or the simulated epoch is far beyond the EoR when the CMB temperature $T_{\text{cmb}}$ is coupled to the spin temperature $T_s$ and the assumption $T_s >> T_{\text{cmb}}$ breaks down. An upgraded version of this implementation uses a very similar algorithm to extend to large boxes (Santos et al. 2010). Here, we apply a very simple ionizing bubble model (Furlanetto et al. 2004) to the calculations of 21 cm brightness temperature anisotropy and its cross-correlation with $Ly_\alpha$ analytically, so we can quickly forecast the detectability of the signal for different combinations of 21 cm and $Ly_\alpha$ experiments, and explore the EoR parameter space without significant computational cost. This approach would be very beneficial when the cross-correlation measurements with different experiments and major foreground or instrumental issues need to be identified in the early stage of the development.

This paper is organized as follows. In Section 2, we introduce the halo model for the ionizing bubble as well as the cross-correlation. In Section 3, the $Ly_\alpha$ luminosity is discussed. Then we focus on the low-$z$ foregrounds for both 21 cm and $Ly_\alpha$ measurements in Section 4 and estimate signal-to-noise for the detectability of different experiments, as well as the uncertainties on the EoR parameters in Section 5. We conclude in Section 6. We use the Planck cosmological parameters: $\Omega_b h^2 = 0.02230$, $\Omega_c h^2 = 0.1188$, $H_0 = 67.74$ km/s/Mpc, $Y_p = 0.249$, $\ln(10^{10} A_s) = 3.064$ at $k_s = 0.05$ Mpc$^{-1}$, $n_s = 0.9667$, and $\tau = 0.058$.

2. Theoretical Model of the Cross-correlation

Here we describe the basic ingredients of our halo model. Since the mean ionizing fraction is not precisely constrained by current observations, we use the CAMB’s reionization model (Lewis 2008); i.e.,

$$\xi_c(z) = \frac{1}{2} \left[ 1 + \tanh \left( \frac{(1 + z_{re})^{3/2} - (1 + z)^{3/2}}{\Delta y} \right) \right],$$

where the redshift $z_{re}$ is derived from the optical depth $\tau$ today; i.e.,

$$\tau = \int_0^{z_{re}} dl n_c(\xi') \sigma_T,$$

and $\Delta y = 1.5 \sqrt{1 + z_{re}}$. Here, $\sigma_T$ is the Thomson cross-section, the electron density is $n_e = (1 - 3/4 Y_p) n_{H_0}/m_{H_0} a^{-3} \xi_c$, the comoving length $dl = a dy'$, the Helium fraction is $Y_p$, the proton mass is $m_{H_0}$, and the mean neutral hydrogen fraction is $\xi_H = 1 - \xi_c$. We assume that helium is singly ionized along with hydrogen, while the double ionization of helium is neglected.

The 21 cm brightness temperature can be split into two components, $T_b(z)\psi(x, z)$, in which the isotropic background temperature is

$$T_b(z) = 27 \left( \frac{1 - Y_p}{1 - 0.248} \right) \left( \frac{\Omega_b}{0.044} \right) \times \left[ \left( \frac{0.27}{\Omega_m} \right) \left( \frac{1 + z}{10} \right) \right]^{1/2} \text{mK},$$

and spatial fluctuation $\psi$ is (Meerburg et al. 2013)

$$\psi(x, z) = \delta H(1 + \delta)(1 + \delta) = \delta_H(1 + \delta_x + \delta + \delta_x \delta).$$

Here $\delta_x$ is the density contrast of ionizing field $(\delta)$ and we neglect perturbations introduced by spin-temperature fluctuations and peculiar velocities. For the 21 cm field, we only consider the signals from IGM as galaxy contributions are $10^{-4}$ times smaller (Gong et al. 2011), and model the ionizing field with “bubbles” (Meerburg et al. 2013). From Equation (4), the two-point correlation functions for ionizing and matter density contrasts are

$$\left< \delta_x \delta_x \right> = \xi_{xx}/\xi_{H,H}^2, \left< \delta_x \delta_y \right> = \xi_{xy}/\xi_{H,H}, \text{and} \left< \delta_x \delta_y \right> = \xi_{xy}.$$ 

The auto-correlation function of the isotropic 21 cm spatial fluctuation field (Zaldarriaga et al. 2004; Furlanetto et al. 2004) is

$$\xi_{\psi\psi} = \xi_{xx}(1 + \xi_{xx}) + \xi_{H,H}^2 \xi_{xx} + \xi_{xx} + 2\xi_{H,H} + \xi_{xx} + \xi_{H,H}.$$ 

We should note that this 21 cm auto-correlation function is only an approximation in that the three-point correlation terms neglected are generally substantial, as Lidz et al. (2007) pointed out. However, we mainly focused on the cross-correlation calculations, for which we did consider all of the higher order terms. Although the assumption that the density field is Gaussian is a reasonable approximation on most of the scales of interest, we create the ionizing field from a Poisson process, as we will discuss later, to account for its non-Gaussianity. We
only use power spectrum to do the statistics so that the non-Gaussianity of the field is not captured (Zaldarriaga et al. 2004). We neglect the redshift distortions and make use of the fact that the spin temperature is significantly higher than CMB at $z < 10$ (Thomas & Zaroubi 2011). We Fourier transform the correlation function, assuming that the quadratic terms are negligibly small. The 21 cm power spectrum is

$$P_{\nu\nu}^{(2h)} = P_{xx}^{(2h)} + \delta_H^2 P_{bb}^{(2h)} + 2\delta_H P_{x\delta}^{(2h)} + P_{xbb}^{(2h)} \approx P_{xx}^{(2h)} + \delta_H^2 P_{bb}^{(2h)} + 2\delta_H P_{x\delta}^{(2h)}. \quad (6)$$

The power spectrum can be calculated from a halo model by describing HII regions as bubbles. The two-halo term of $P_{x\delta}$ is higher order and negligible (Furlanetto et al. 2004).

The virial temperature of halo $T_{\text{vir}} = 5 \times 10^4 \text{K}$ (suggested by Meierberg et al. 2013) sets the minimum halo mass

$$T_{\text{vir}} / 10^4 \text{K} = 1.1 \left(\frac{\Omega_m h^2}{0.15}\right)^{1/3} \left(\frac{1+z}{10}\right) \left(\frac{M_{\text{th}}}{10^8 M_\odot}\right)^{2/3}. \quad (7)$$

With this threshold mass, we can calculate the mean number density of bubble $n_b = \int_{M_{\text{th}}}^\infty dM \frac{dn}{dM}$ and the average bubble size from $\bar{V}_b = -(\ln \bar{\delta}_H) / \bar{V}_b$. When it is compared to the predicted value

$$\bar{V}_b = \int dRP(R) V_{\nu}(R) = \frac{4\pi R^3}{3} e^{-\frac{1}{2} \ln(\bar{\delta}_H)^2/2},$$

the bubble radius is constrained. The bubble radius $R$ is assumed to satisfy a logarithmic distribution (Lidz et al. 2009) as

$$P(R) = \frac{1}{R \sqrt{2\pi \sigma_{\ln R}^2}} e^{-\frac{1}{2} \left(\ln(R/R)\right)^2}. \quad (9)$$

In Figure 1, we show the distribution of the bubble radius at different redshifts.

**Figure 1.** Bubble size radius $R$ is assumed to satisfy a logarithmic distribution in Equation (9). We show the distributions at redshifts $z = 6, 7, 8, 9$, and 10. The mean ionization fractions $\bar{\delta}_H$ at these redshifts are given in parentheses.

Given the average volume and number density, the ionizing field is generated from a Poisson process

$$\langle x_{\nu}(x) \rangle = 1 - e^{-n_b(x)} V_b, \quad (10)$$

and its number density is

$$n_b(x) = n_b(1 + b \delta_L(x)), \quad (11)$$

where the density contrast $\delta_L$ is the matter density $\delta$ smoothed by a top-hat window of radius $R$. The top-hat window in Fourier space is

$$\tilde{W}_R^2(k) = \frac{1}{V_b} \int_0^\infty dR P(R) |V_b(R) W(kR)|^2. \quad (12)$$

The shape factor for the ionizing field $x$ is defined as

$$X_i^x(k, M, z) = \bar{\delta}_H \ln \bar{\delta}_H b_{\text{bubble}} \tilde{W}_R(k) u_1, \quad (13)$$

with the bubble bias given by

$$b_{\text{bubble}} = \frac{1}{\bar{n}_b} \frac{d}{d \ln M} b(M, z) \frac{dn}{dM} \frac{dM}{d \ln M}. \quad (14)$$

Here, $u_1$ is the Fourier transform of the NFW profile (Navarro et al. 1996); i.e., $u_1 = M / \bar{P}_b u$. The NFW Fourier transform is

$$u(k, M, z) = \frac{1}{M} \int_0^r dr \frac{4\pi r^2 \sin kr}{kr} \rho_{\text{NFW}}, \quad (15)$$

which is derived from a standard NFW profile

$$\rho_{\text{NFW}} = \rho_s \left(\frac{r}{r_s}\right)^{-1} \left(1 + \frac{r}{r_s}\right)^{-2}. \quad (16)$$

The detailed discussions of $\rho_s$, $r_s$ and $r_{\text{vir}}$ can be found in Feng et al. (2016).

The 1-halo term of the 21 cm field (Mortonson & Hu 2007; Wang & Hu 2006) is

$$P_{\nu\nu}^{(1h)} = P_{xx}^{(1h)} + \delta_H^2 P_{bb}^{(1h)} + 2\delta_H P_{x\delta}^{(1h)} + P_{xbb}^{(1h)} \approx P_{xx}^{(1h)} + \delta_H^2 P_{bb}^{(1h)} + P_{x\delta}^{(1h)}, \quad (17)$$

where $P_{xx}^{(1h)} = (\delta_b - \xi_b^2) V_b \tilde{W}_R^2$ and $P_{x\delta}^{(1h)} = (\xi_b - \xi_b^2) P_{\delta}$. The term $P_{bb}^{(1h)}$ is zero as the bubble is completely ionized. Here $\sigma_R^2 = \frac{1}{2} \tilde{W}_R(k) P_{\delta}(k)$ and $\tilde{P}_{\delta} = P_{\delta} V_b \sigma_R^2 \sqrt{\tilde{P}_R^2 + (V_\delta \sigma_\delta^2)^2}$. The two-halo term can also be easily calculated with the shape factor in Equation (13). On the other hand, the intensity mapping (IM) of Ly$\alpha$ emitters (LAEs) is a biased tracer of the same dark matter distribution, i.e., $\sim (1 + \delta_b)$. For simplicity, we use $\delta_{\text{Ly} \alpha} = \delta_b$. The cross-correlation between 21 cm and LAEs is

$$\xi_{\nu \eta} = \bar{\delta}_H (\xi_{\delta_b \delta} + \xi_{\delta, \delta} \bar{\delta}) + \xi_{\delta_b}, \quad (18)$$

and the 3D power spectrum is

$$P_{\nu \eta} = \bar{\delta}_H (P_{\delta_b \delta} + P_{\delta, \delta} \bar{\delta}) + P_{\delta_b \delta}. \quad (19)$$

Here, the two-halo and one-halo terms are given by

$$P_{\nu \eta}^{(2h)} = \bar{\delta}_H (P_{\delta_b \delta}^{(2h)} + P_{\delta, \delta}^{(2h)} + P_{\delta_b \delta}^{(2h)}) \approx \bar{\delta}_H P_{\delta_b \delta}^{(2h)} + P_{\delta_b \delta}^{(2h)} \quad (20)$$

and its number density is

$$n_b(x) = n_b(1 + b \delta_L(x)), \quad (11)$$
and
\[ P_{\nu}^{(1h)} = \tilde{x}_H(P_{\nu}^{(1h)} + P_{\nu,ax}^{(1h)}) + P_{\nu,ax}^{(1h)} \]
\[ \approx 0, \]  \quad (21)
respectively. The subscript \( \delta, \chi \delta \) essentially refers to \( \delta, \chi \) and the “,” is omitted for simplicity.

On small scales \( P_{\delta,\chi}^{(1h)} \) cancels out the term \( P_{\nu,\chi}^{(1h)} \), so the summation is almost zero (Lidz et al. 2009). Also, the large-scale information of \( P_{\delta,\chi}^{(2h)} \) should be very negligible. With all of these approximations, the final power spectrum of the 21 cm–\( \text{Ly}\alpha \) cross-correlation is
\[ P_{\nu} \approx P_{\nu,ax}^{(2h)} + \tilde{x}_H P_{\nu,ax}^{(2h)}. \]  \quad (22)

The halo-model approach, i.e.,
\[ P_{1h,\chi}(k, z) = \int dM \frac{dn}{dM} X_I(k, M, z) Y_I(k, M, z) \]  \quad (23)
and
\[ P_{2h,\chi}(k, z) = P_{\text{lin}}(k, z) \int dM \frac{dn}{dM} b(M, z) \tilde{X}_I(k, M, z) \]
\[ \times \int dM \frac{dn}{dM} b(M, z) \tilde{Y}_I(k, M, z), \]  \quad (24)
can be used to calculate each power spectrum in Equation (22). In these equations, \( dn/dM \) is the mass function and \( b(M, z) \) is the bias. The linear matter power spectrum is \( P_{\text{lin}} \). We will work out the shape factors \( X_I(k, M, z) \) and \( Y_I(k, M, z) \) (or \( \tilde{X}_I(k, M, z) \) and \( \tilde{Y}_I(k, M, z) \)) for \( \text{Ly}\alpha \) in the next section.

3. \( \text{Ly}\alpha \) Emission

The UV radiation emitted from massive and short-lived stars can ionize the neutral hydrogen in the interstellar medium (ISM) in galaxies and the number of ionizing photons closely depends on the star formation rate (SFR). The fitted SFR (Silva et al. 2013) is
\[ \frac{\text{SFR}(M, z)}{M_\odot \text{ yr}^{-1}} = 2.8 \times 10^{-28} M^a \left(1 + \frac{M}{M_1}\right)^b \times \left(1 + \frac{M}{M_2}\right)^d, \]  \quad (25)
where \( a = 2.8, \ b = -0.94, \ d = -1.7, \ M_1 = 10^9 M_\odot \) and \( M_2 = 7 \times 10^{10} M_\odot \).

The ionizing photons could escape the galaxies with a fraction \( f_{\text{esc}}(M, z) = e^{-\alpha(M, z) r^3} \), but the remains will ionize the hydrogen and 66% of the ionization will result in a recombination process that produces \( \text{Ly}\alpha \) photons. The dust in the ISM can also absorb the \( \text{Ly}\alpha \) emissions, and the remaining fraction that survives the dust extinction is \( \tilde{f}_{\text{Ly}\alpha}(z) \). The luminosity due to the recombination is then calculated as
\[ L_{\text{rec}}^{\text{GAL}}(M, z) = 1.55 \times 10^{42} (1 - f_{\text{esc}}) \tilde{f}_{\text{Ly}\alpha} \]
\[ \times \frac{\text{SFR}}{M_\odot \text{ yr}^{-1}} (\text{erg s}^{-1}). \]  \quad (26)

The ionizing radiation can heat the gas so that the process of hydrogen excitation and cooling produces \( \text{Ly}\alpha \) emission as well. The luminosities due to excitation and cooling are
\[ L_{\text{exc}}^{\text{GAL}}(M, z) = 4.03 \times 10^{44} (1 - f_{\text{esc}}) f_{\text{Ly}\alpha} \]
\[ \times \frac{\text{SFR}}{M_\odot \text{ yr}^{-1}} (\text{erg s}^{-1}), \]  \quad (27)
and
\[ L_{\text{cooling}}^{\text{GAL}}(M, z) = 1.69 \times 10^{35} f_{\text{Ly}\alpha} \left(1 + \frac{M}{10^8}\right)^2 \]
\[ \times \left(1 + \frac{M}{2 \times 10^{10}}\right)^2 \left(1 + \frac{M}{3 \times 10^{13}}\right)^3 (\text{erg s}^{-1}), \]  \quad (28)
respectively.

Besides these line emissions, the continuum produces \( \text{Ly}\alpha \) photons through stellar radiation, free–free (ff), free–bound (fb), and two-photon (2γ) processes. Among these contributions, the stellar emission with a blackbody spectrum below the Lyman limit is dominant and its luminosity is
\[ L_{\text{stellar}}^{\text{GAL}}(M, z) = 5.12 \times 10^{40} f_{\text{Ly}\alpha} \frac{\text{SFR}}{M_\odot \text{ yr}^{-1}} (\text{erg s}^{-1}). \]  \quad (29)

Our calculation takes all of these continuum lines into account, and the detailed line luminosity can be found in Silva et al. (2013).

The total \( \text{Ly}\alpha \) luminosity \( L(M, z) \) from a galaxy is a summation of all of the above components and the shape factor for \( \text{Ly}\alpha \) field is
\[ X_I(k, M, z) = \frac{L(M, z)}{4\pi D_L^2} y d\lambda u(k, M, z). \]  \quad (30)

Here the conversion factor from frequency to comoving distance is \( y = dx/dv = \lambda/(c a^3) d\lambda/dz \), \( \lambda \) is the line rest frame wavelength, and \( D_L \) and \( D_\lambda \) are luminosity and angular comoving distances, respectively.

This construction of shape factor is only a mathematical definition that facilitates the halo-model calculations in Equations (23) and (24). We should note that for individual point-like \( \text{Ly}\alpha \) emitters, it would appear to be extended due to spatial diffusion of \( \text{Ly}\alpha \) photons (Zheng et al. 2010, 2011). Also, the scattering of photons on the red-side of \( \text{Ly}\alpha \) in the IGM would further damp the \( \text{Ly}\alpha \) flux along the line of sight (Miralda-Escudé 1998). We first consider that the \( \text{Ly}\alpha \) emission is a biased tracer of the underlying dark matter distribution and phenomenologically account for the extended structure by the mass- and redshift-dependent quantity \( b(M, z) \) in Equation (24) and the luminosity function \( L(M, z) \). Next we will discuss some dominating effects of IGM on the \( \text{Ly}\alpha \) emissions, but effects such as the damping wing of \( \text{Ly}\alpha \) have to rely on a numerical simulation.

The mean \( \text{Ly}\alpha \) intensity varies at different redshifts as
\[ \tilde{I}_{\text{Ly}\alpha}(z) = \int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn}{dM} L(M, z) \frac{4\pi D_L^2}{4\pi D_\lambda^2} y D_\lambda^2, \]  \quad (31)
where \( M_{\text{min}} = 10^8 M_\odot \) and \( M_{\text{max}} = 10^{13} M_\odot \).

The escaped photons from galaxies can ionize the IGM, which can also emit \( \text{Ly}\alpha \) photons due to the recombination process. The recombination rate is
\[ n_{\text{rec}} = \alpha n_e n_{\text{HII}}. \]  \quad (32)
space becomes \( k_{\text{f-s}} = (x_a/x_b)k_{\text{l}}, y_a/y_b k_{\text{l}} \), which is not radially symmetric. The low-\( z \) foregrounds are identified as H\( \alpha \) [6563 Å, \( z = 0.5 \)], O \( \text{III} \) [5007 Å, \( z = 0.9 \)], and O \( \text{II} \) [3727 Å, \( z = 1.6 \)] with luminosities \( L_{\text{H}\alpha} = 1.3 \times 10^{41} \text{ SFR}_{\text{M}_{\odot} \text{yr}^{-1}} \), \( L_{\text{O \text{III}}} = 7.1 \times 10^{40} \text{ SFR}_{\text{M}_{\odot} \text{yr}^{-1}} \), and \( L_{\text{O \text{II}}} = 1.3 \times 10^{41} \text{ SFR}_{\text{M}_{\odot} \text{yr}^{-1}} \), respectively. The low-\( z \) SFR is exclusively modeled as

\[
\frac{\text{SFR}(M_{\odot}, z)}{M_{\odot} \text{yr}^{-1}} = 10^{a+bz}\left(\frac{M_1}{M_2}\right)^c \left(\frac{M_1}{M_2}\right)^d
\]

(33)

for the foreground line emissions. This SFR model is fitted to the numerical simulations below \( z = 2 \) and the parameters are constrained as \( a = -9.097, b = 0.484, c = 2.7, d = -4.0 \), \( M_1 = 10^8 M_{\odot} \), and \( M_2 = 8 \times 10^{11} M_{\odot} \) (Gong et al. 2014).

The projected power spectrum of the foreground is then expressed as

\[
P_{\text{f-s}}(k_{\text{l}}, k_{\text{f}}, z_0) = \left(\frac{x_a}{x_b}\right)^2 \frac{y_a}{y_b} P(k_f, z_f)
\]

(34)

and \( k_f = \sqrt{(x_a/x_b)^2 k_{\text{l}}^2 + (y_a/y_b) k_{\text{l}}^2} \). In Figure 3, we show the unprojected and projected power spectra for H\( \alpha \), O \( \text{III} \) and O \( \text{II} \) show similar contours so they are neglected. In Figure 4, the blue and red curves are radially averaged from the anisotropic power spectra.

In Figures 4 and 5, we show the power spectra for \( \text{Ly} \alpha \) and the foreground lines H\( \alpha \), O \( \text{III} \), and O \( \text{II} \) with projection and flux masking at redshifts \( z = 7 \) and 9. The projected foreground emissions are much higher than the \( \text{Ly} \alpha \) lines. By selecting the brightest sources at the flux detection threshold and forming a mask, we can effectively remove those “hot” pixels which only account for a very tiny fraction of the sky coverage (Pullen et al. 2014; Gong et al. 2014). As can be seen in Figure 6, we show the percentage of the removed pixels as the threshold flux changes. We find that a flux cut at \( 10^{-18} \text{ W m}^{-2} \) can significantly lower amplitudes of the foreground power spectra while removing less than 0.1% of the pixels. Therefore, the flux masking procedure makes the low-\( z \) foregrounds negligible.

The synchrotron radiation dominates the 21 cm signals but its smooth spectral feature can easily be used to isolate this component in frequency domain, also the galactic foregrounds are not correlated with extragalactic line emissions at low-\( z \). So we do not expect any noticeable cross-correlations between galactic synchrotron and \( \text{Ly} \alpha \) foregrounds. However, the radio point sources that are too faint to be resolved are indeed correlated with the low-\( z \) foregrounds within \( 0.5 < z < 1.6 \), so this component would be picked up in the 21 cm–\( \text{Ly} \alpha \) cross-correlation, making it not as systematic-free as expected. To estimate its contribution, we use the model in Gleser et al. (2008), Singal et al. (2010), and Serra et al. (2008). The model is described as

\[
\delta_{\text{radio}}(x) = \left(\frac{\partial B}{\partial T}\right)^{-1} I_0^{\text{radio}} \delta_\nu(x)
\]

(35)

based on the fact that the radio point source is a tracer of underlying density field. Here we have defined \( I_0^{\text{radio}} = \int_0^{S_{\text{cut}}} dS dN/dS \) and assume a flux limit \( S_{\text{cut}} = 1 \text{ mJy} \), above which the radio point sources are bright enough to be resolved. The flux distribution is a simple power-law, i.e.,

Figure 2. IGM contribution to both the auto- and cross-power spectra at \( z = 7 \). The IGM component is negligible, compared to the galaxy. The dashed portion is negative. In the y axis, \( \Delta^2(k) = k^3 / (2\pi^2) P(k) \).

where \( n_e = e_n e_b, n_{\text{H}2} = x_n n_b C, C = (1 - Y_p)/(1 - 3Y_p/4) \), and \( \alpha \) is a case A comoving recombination coefficient. The luminosity function of IGM is \( L_{\text{IGM}} = f_{\text{rec}} n_{\text{rec}} E_{\text{Ly} \alpha} \) and the fraction \( f_{\text{rec}} \) is spin-temperature dependent. We show the contribution of IGM in Figure 2 and it is seen that the IGM contribution is negligible for both auto- and cross-power spectra.

Another IGM contribution to \( \text{Ly} \alpha \) emission comes from the scattering of Ly\( n \) photons escaping from galaxies. From the previous calculations (Silva et al. 2013, 2016; Pullen et al. 2014) it is found that the diffuse IGM contribution is a few orders of magnitude smaller than the galaxies. Therefore, we ignore this contribution to the overall \( \text{Ly} \alpha \) signal.

4. Low-\( z \) Foregrounds

The \( \text{Ly} \alpha \) emission at the EoR can be significantly contaminated by low-\( z \) foregrounds. The foreground at \( z_f \) projected onto the source plane \( z_s \) becomes anisotropic as the wave vector of the foreground power spectrum in Fourier
\[ \frac{dN}{dS} = A(S/S_0)^\alpha, \]
where \( A = 4 \, \text{mJy}^{-1} \, \text{sr}^{-1}, \) \( S_0 = 880 \, \text{mJy}, \) and \( \alpha = -1.75 \) (Liu et al. 2009). Also, \( \frac{\partial B_c}{\partial T} = 99.27 \, \text{Jy} \, \text{sr}^{-1} / (\mu K)x^2e^x/(e^x - 1)^2 \) and \( x = h\nu/k_bT_{\text{CMB}} = \nu/56.84 \, \text{GHz} \). In Fourier space, the shape functions of the halo model for the radio sources are described as

\[ X_k(k, M, z) = \sqrt{2N_cN_a(k, M, z) + N_c^2u^2(k, M, z)} \frac{\hat{n}_g}{\bar{n}_g}, \]
(36)

and

\[ \tilde{X}_k(k, M, z) = \frac{N_d(k, M, z)}{\bar{n}_g}, \]
(37)

which are directly inserted into Equations (23) and (24) to obtain the one-halo and two-halo terms for the point-source clustering. Both \( X_k(k, M, z) \) and \( \tilde{X}_k(k, M, z) \) describe the Fourier-space profile of a point source with mass \( M \) located at redshift \( z \). Here the central and satellite galaxy numbers are

\[ N_c(M) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_10 M - \log_10 M_{\text{min}}}{\sigma_M} \right) \right] \]
(38)

and

\[ N_s(M) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log_10 M - \log_10 M_{2\text{min}}}{\sigma_M} \right) \right] \times \left( \frac{M}{M_s} \right)^\alpha. \]
(39)

The mean galaxy number density is

\[ \bar{n}_g(z) = \int dMn(M, z)N_c(M), \]
(40)

where \( N_c(M) = N_c(M) + N_s(M) \). The parameters determined from luminosity and color dependence of galaxy clustering in

---

**Figure 3.** Power spectrum projection of Hα at \( z = 7 \). O III and O II power spectrum projections have similar patterns.

**Figure 4.** Lyα power spectrum at \( z = 7 \) (green). We show the low-z foreground power spectra with no projection (black), projection (blue), and masking (red). The instrumental noise (orange) is derived from the proposed CDM specification listed in Table 2.

**Figure 5.** Lyα power spectrum at \( z = 9 \) (green). The description of other lines is the same as Figure 4.
the SDSS DR7 main galaxy sample are $M_{\text{min}} = 10^9 M_\odot$, $\sigma_M = 0.2$, $M_a = 5 \times 10^{10} M_\odot$, and $\alpha_a = 1$ (Zehavi et al. 2011).

We estimate that the radio foreground contributions at the Ly\$\alpha$ foreground redshifts and the raw power spectra are a few orders of magnitude higher than the 21 cm signal as revealed by Liu et al. (2009) and Alonso et al. (2014). Therefore, the foreground suppression is very crucial; the spectral fitting procedure studied in Liu et al. (2009) demonstrated that the radio foregrounds can be reduced by six orders of magnitude in map space, and it has been validated that this is true from flux cut 0.1–100 mJy. Consequently, the radio foreground contamination becomes negligible, and we show all of the power spectra in Figures 7 and 8 at redshifts $z = 7$ and 9. We see that the resulting radio point sources have very negligible contaminating power on the 21 cm measurements. Finally, we show the 21 cm–Ly\$\alpha$ cross-power spectra at redshifts $z = 7$ and 9 in Figures 9 and 10 with both foreground separation schemes incorporated. In Figure 11, we show both the evolutions and cross-correlation coefficients of the 21 cm–Ly\$\alpha$ cross-power spectrum as a function of $k$ from $z = 6$ to $z = 10$.

Despite the fact that all of the foreground cross-correlations are small, this has to rely on the assumption that we have very efficient foreground mitigation strategies for both 21 cm and Ly\$\alpha$ measurements. Non-negligible foreground residuals on 21 cm and Ly\$\alpha$ would make a great impact on the power spectrum uncertainties, even if they are uncorrelated.

5. Forecast for the Experiments

In this section, we consider two experiments, SKA and Cosmic Dawn Intensity Mapper (Pritchard et al. 2015; Cooray et al. 2016), and investigate the detectability of the 21 cm–Ly\$\alpha$ cross-correlation with both instrumental noises and foregrounds. We list the experimental specifications in Tables 1 and 2.

The instrumental noise of the 21 cm experiment is entirely determined by key factors such as integration time $t_0$, system temperature $T_{\text{sys}}$, maximum baseline $D_{\text{max}}$, collecting area $A_{\text{tot}}$, antenna number $N_a$, and frequency resolution $d\nu$. The noise is then given by

$$P_N^{21 \text{ cm}} = \chi^2 y \pi \alpha A_{\text{tot}}^2 \frac{\lambda D_{\text{max}} T_{\text{sys}}}{N_a \nu} \frac{1}{t_0}.$$ (41)

For the Ly\$\alpha$ experiments, the noise is

$$P_N^{\text{Ly} \alpha} = V_{\text{pix}}^{-2},$$ (42)

where the comoving volume subtended by the detector pixel

$$V_{\text{pix}} = \chi A_{\text{pix}}^2 \delta_\nu,$$ (43)

depends on the pixel area $A_{\text{pix}}$ and frequency resolution $\delta_\nu$. The general Knox formula (Bowman et al. 2006) for the measured
The noise- and foreground-included power spectra are formed as $\Delta^2(k) = \Delta^2(k)_{\text{signal}}$, and the number of modes in the bin $k$ is $N_m = 2\pi k^2 \Delta k V_s$, and
\begin{equation}
\Delta^2(k)_{\text{signal}} = \frac{\langle P_X(k, z) P_Y(k, z) \rangle}{N_m}.
\end{equation}

Here $X, Y = \{21 \text{ cm}, \text{ Ly} \alpha \}$, the survey volume is $V_s = \chi^2 A \Delta B$, $A$ is the survey area, and $B$ is the bandwidth (BW). For Ly$\alpha$ experiments, the minimum and maximum scales are determined by the survey and pixel areas. For 21 cm, we normally consider modes at scales below $k = 10 \text{ Mpc}^{-1}$ and get the minimum $k$ from the total survey area. For the cross-correlation, the common $k$ range is chosen from two experiments and the minimum volume between the 21 cm and Ly$\alpha$ experiments is taken to calculate the number of modes.

The Astrophysical Journal, 846:21 (11pp), 2017 September 1 Feng, Cooray, & Keating
For example, at $k_{\text{max}} = 10 \, h \, \text{Mpc}^{-1}$, a horizontal cut between 0.05 $h \, \text{Mpc}^{-1} < k_{\parallel} < 0.1 \, h \, \text{Mpc}^{-1}$ introduces negligible changes to the overall signal-to-noise. But for a small $k_{\text{max}}$, such as 0.5 $h \, \text{Mpc}^{-1}$, the total wedge could reduce the signal-to-noise by 11% with a horizontal cut at $k_{\parallel} = 0.1 \, h \, \text{Mpc}^{-1}$.

In Figures 13–15, we show all of the power spectra and their band errors for 21 cm and Lyα at $z = 7$ and 9. As can be seen from Figure 15, the anti-correlations between neutral hydrogen and galaxies can be probed at very high signal-to-noise ratios.
From the forecasted power spectra in Figure 15, we can further try to constrain the EoR parameters defined as $P = \{\tau, \Delta y, \sigma_{\ln R}\} = \{0.058, 6.0, 1.0\}$. The Fisher matrix (Pober et al. 2014) is

$$F_{ij} = \sum_{k,z} \frac{1}{(\Delta \eta(k, z))^2} \frac{\partial \Delta^2(k, z)}{\partial p_i} \frac{\partial \Delta^2(k, z)}{\partial p_j}. \quad (45)$$

Here $\Delta \eta(k, z)$ is the error on the cross-power spectrum $\Delta^2$, and $p_i$ refers to any parameters in the set $P$ and $\Delta^2 = k^3/(2\pi^2)P_{21\text{ cm} - \text{Ly} \alpha}$. All of the 1σ confidence levels, as well as the likelihood functions in Figure 16, are calculated from the 21 cm–Lyα cross-power spectra at $z = 7$ and 9. As can be seen in the figure, the bubble size can be constrained from the cross-correlation while the errors on the optical depth and duration of

**Figure 16.** 1-σ confidence levels for the optical depth $\tau$, the EoR duration $\Delta y$, and the rms of the bubble size $\sigma_{\ln R}$ at $z = 7$ (solid, $\delta_H = 0.80$) and 9 (dotted, $\delta_H = 0.16$).

**Figure 17.** 1-σ confidence levels for the optical depth $\tau$, the EoR duration $\Delta y$, and the rms of the bubble size $\sigma_{\ln R}$ at $6 \leq z \leq 10$. 

Feng, Cooray, & Keating
the reionization transition are large. This is due to the fact that the cross-correlation is proportional to $x_a(z)$ and not $x_a^2(z)$, to which the $21 \text{ cm} - 21 \text{ cm}$ and $21 \text{ cm} - \tau$ (Meerbreg et al. 2013) are proportional. Therefore, the cross-power spectrum at a single redshift is less sensitive to the reionization history. However, the cross-correlations at different redshifts might be able to break the degeneracies among the parameters, and so are useful as complementary probes to cosmological and astrophysical problems. In Figure 17, we reduce the bandwidth by a factor of 5 and combine the cross-power spectra measured at $z = 6, 7, 8, 9$, and 10. It is seen that the Fisher matrix error bars on all of the parameters are significantly reduced, and that multi-redshift measurements within narrow bins can effectively break the parameter degeneracies.

6. Conclusion

In this work, we applied a bubble model to the computation of $21 \text{ cm}$ and Ly$\alpha$ cross-correlation at the EoR. Making use of the empirical relation between Ly$\alpha$ luminosity and mass for the line emissions, we also calculated the power spectra for Ly$\alpha$. The $21 \text{ cm} - \text{Ly} \alpha$ cross-power spectrum in this fast approach can reproduce the key features of the one made by detailed numerical simulations, and we can use it to quickly assess the overall performance of future EoR experiments.

The cross-correlation is contaminated by the low-$z$ foregrounds for both $21 \text{ cm}$ and Ly$\alpha$. We studied the radio galaxies for the $21 \text{ cm}$ experiments at H$\alpha$, O I, and O II line emissions for the Ly$\alpha$ experiments. All of these foregrounds could be a few orders of magnitude higher than the signals we are probing if the foreground mitigation is not incorporated. The map-space spectral fitting can effectively remove the radio point-source contaminations, while a flux masking for the intensity mapping experiments have been shown to be a good and easy foreground-removal method.

We take advantage of this efficient algorithm and estimate the errors on the EoR parameters $\tau$, $\Delta y$, and $\sigma_{\text{m-R}}$, based on the Fisher matrix formalism. For other physical processes during the EoR, such as X-ray heating, supernovae explosion, and shock heating, numerical simulations with these effects or an extension to this work should be devised. We will discuss these in future works.

A.C. and C.F. acknowledge support from NASA grants NASA NNX16A69G, NASA NNX16AF39G and Ax Foundation for Cosmology at UC San Diego.

References

Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, ApJ, 809, 61
Alonso, D., Ferreira, P. G., & Santos, M. G. 2014, MNRAS, 444, 3183
Bonaldi, A., & Brown, M. L. 2015, MNRAS, 447, 1973
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, ApJ, 648, 20
Cooray, A., Bock, J., Burgarella, D., et al. 2016, arXiv:1602.05178
DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
Dillon, J. S., Liu, A., Williams, C. L., et al. 2014, PrRvD, 89, 023002
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Feng, C., Cooray, A., & Keating, B. 2017, ApJ, 836, 127
Furlanetto, S. R., & Lidz, A. 2007, ApJ, 660, 1030
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006a, PRl, 433, 181
Furlanetto, S. R., Oh, S. P., & Pierpaoli, E. 2006b, PrRvD, 74, 103502
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, ApJ, 613, 1
Gleser, L., Nusser, A., & Benson, A. J. 2008, MNRAS, 391, 383
Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46
Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJL, 785, 72
Iliev, I., Santos, M., Mesinger, A., Majumdar, S., & Mel Willie, G. 2015, in Proc. of Advancing Astrophysics with Square Kilometre Array (AASKA14) (Trieste: SISSA), 1
Lewis, A. 2008, PrRvD, 78, 023002
Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70
Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252
Lidz, A., Zahn, O., McQuinn, M., et al. 2007, ApJ, 659, 865
Liu, A., Tegmark, M., & Zaldarriaga, M. 2009, MNRAS, 394, 1575
Loeb, A., & Barkana, R. 2001, ARA&A, 39, 19
Machado, S., & Rhoads, J. E. 2006, ApJL, 117, L5
Mirolo, S., & Rhoads, J. E. 2006, ApJL, 647, L95
McQuinn, M., Furlanetto, S. R., Hernquist, L., Zahn, O., & Zaldarriaga, M. 2005, ApJ, 630, 643
Meerbreg, P. D., Dvorkin, C., & Spengel, D. N. 2013, ApJL, 779, 124
Mesinger, A., & Furlanetto, S. 2007, ApJ, 669, 663
Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955
Miralda-Escudé, J. 1998, ApJ, 501, 15
Mortonson, M. J., & Hu, W. 2007, ApJ, 657, 1
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, AJ, 139, 1468
Planck Collaboration, Adam, R., Aghanim, N., et al. 2016, ApJS, 192, 18
Pober, J. C., Liu, A., Dillon, J. S., et al. 2014, ApJ, 782, 66
Pritchard, J., Ichiki, K., Mesinger, A., et al. 2015, in Proc. of Advancing Astrophysics with Square Kilometre Array (AASKA14) (Trieste: SISSA), 12
Pullen, A. R., Doré, O., & Bock, J. 2014, ApJL, 786, 111
Santos, M. G., Ferramacho, L., Silva, M. B., Amblard, A., & Cooray, A. 2010, MNRAS, 406, 2421
Schlueicher, D. R. G., Banerjee, R., & Klessen, R. S. 2009, ApJ, 692, 236
Serra, P., Cooray, A., Amblard, A., Pagano, L., & Melchiorri, A. 2008, PrRvD, 78, 043004
Silva, M. B., Kooistra, R., & Zaldarriaga, S. 2016, MNRAS, 462, 1961
Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 765, 132
Singal, J., Stawarz, Ł., Lawrence, A., & Petrovian, V. 2010, MNRAS, 409, 1172
Thomas, R. M., & Zaldarriaga, S. 2011, MNRAS, 410, 1377
Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, e007
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Vibraniec, D., Ciardi, B., & Zahn, O., et al. 2016, MNRAS, 357, 666
Wang, W., & Hu, W. 2006, ApJ, 643, 585
Wyithe, J. S. B., & Loeb, A. 2007, MNRAS, 375, 1034
Zahn, O., Lidz, A., McQuinn, M., et al. 2007, ApJ, 654, 12
Zahn, O., Mesinger, A., McQuinn, M., et al. 2011, MNRAS, 414, 727
Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, ApJ, 608, 622
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59
Zheng, Z., Cen, R., Trac, H., & Miralda-Escudé, J. 2010, ApJ, 716, 574
Zheng, Z., Cen, R., Trac, H., & Miralda-Escudé, J. 2011, ApJ, 726, 38