



Analytic Formulation of 21 cm Signal from Cosmic Dawn: Ly α Fluctuations

Janakee Raste^{1,2}  and Shiv Sethi¹¹ Raman Research Institute, Bangalore 560080, India² Joint Astronomy Program, Indian Institute of Science, Bangalore 560012, India

Received 2018 November 2; revised 2019 March 13; accepted 2019 March 25; published 2019 May 3

Abstract

We present an analytic formalism to compute the fluctuating component of the HI signal and extend it to take into account the effects of partial Ly α coupling during the era of cosmic dawn. We use excursion set formalism to calculate the size distribution of randomly distributed self-ionized regions. These ionization bubbles are surrounded by partially heated and Ly α coupled regions, which create spin temperature T_S fluctuations. We use the ratio of number of Ly α to ionizing photons (f_I) and number of X-ray photons emitted per stellar baryon (N_{heat}) as modeling parameters. Using our formalism, we compute the global HI signal, its autocorrelation, and its power spectrum in the redshift range $10 \leq z \leq 30$ for the Λ CDM model. We check the validity of this formalism for various limits and simplified cases. Our results agree reasonably well with existing results from N -body simulations, in spite of following a different approach and requiring orders of magnitude less computation power and time. We further apply our formalism to study the fluctuating component corresponding to the recent observation by the Experiment to Detect the Global Epoch of reionization Signature (EDGES) that shows an unexpectedly deep absorption trough in the global HI signal in the redshift range $15 < z < 19$. We show that, generically, the EDGES observation predicts a larger signal in this redshift range but a smaller signal at higher redshifts. We also explore the possibility of negative real-space autocorrelation of spin temperature and show that it can be achieved for partial Ly α coupling in many cases corresponding to simplified models and a complete model without density perturbations.

Key words: cosmology: theory – dark ages, reionization, first stars

1. Introduction

Our current understanding of the history of the universe suggests that the dark ages of the universe ended around redshift $z \sim 35$ with the formation of the first large-scale structures (epoch of cosmic dawn). These collapsed structures emitted radiation that heated and ionized their surrounding medium (epoch of reionization, or EoR) by $z \sim 8$ (Barkana & Loeb 2001; Morales & Wyithe 2010; Pritchard & Loeb 2012; Natarajan & Yoshida 2014). The physics of first stars and galaxies is only partially understood theoretically and is poorly constrained by observations. The current observational bound on the cumulative history of reionization is provided by the detection of the Gunn–Peterson effect, which indicates that the universe was fully ionized by $z \simeq 6$. The cosmic microwave background (CMB) temperature and polarization anisotropy detections by *WMAP* and Planck give the redshift of reionization, $z_{\text{reion}} = 7.75 \pm 0.73$ (Fan et al. 2000; Becker et al. 2001; Hinshaw et al. 2013; Planck Collaboration et al. 2014, 2016, 2018).

The cleanest probe of the physics of EoR is through the detection of the redshifted hyperfine 21 cm line of neutral hydrogen (HI). This signal carries crucial information about the first sources of radiation in the universe and their spectrum in three frequency bands: ultraviolet (UV) radiation (which ionizes the surrounding medium), Ly α radiation (which determines the relative population of neutral hydrogen atoms in hyperfine states), and X-ray photons (which heat and partially ionize the medium). In addition, the sources that emitted soft radio photons would also affect the observable HI signal considerably (e.g., Ewall-Wice et al. 2018; Feng & Holder 2018). Along with primordial density perturbations given by the Λ CDM model, the inhomogeneities of these radiation fields establish the length scales of the fluctuating component of the signal.

The epoch of cosmic dawn and EoR has been studied in detail using numerical, semianalytic, and analytic methods

(e.g., Pritchard & Furlanetto 2007; Mesinger et al. 2011, 2013; Visbal et al. 2012; Tashiro & Sugiyama 2013; Fialkov et al. 2014, 2017; Pacucci et al. 2014; Ghara et al. 2015). Theoretical estimates based on standard thermal and ionization history suggest that the global signal is observable in both absorption and emission with its strength in the range -200 to 20 mK in a frequency range of 50 – 150 MHz, corresponding to a redshift range $25 > z > 8$ (e.g., Madau et al. 1997; Tozzi et al. 2000; Gnedin & Shaver 2004; Sethi 2005). The fluctuating component of the signal is expected to be an order of magnitude smaller on scales in the range 3 – 100 Mpc, which implies angular scales $\simeq 1$ – 30 arcmin (e.g., Furlanetto et al. 2004a, 2004b; Zaldarriaga et al. 2004; Pritchard & Furlanetto 2007; for comprehensive reviews see e.g., Morales & Wyithe 2010; Pritchard & Loeb 2012; Natarajan & Yoshida 2014). Many of the ongoing and upcoming experiments have the capability of detecting this signal in hundreds of hours of integration (e.g., Mesinger et al. 2014; Ahn et al. 2015). Upper limits on the fluctuating component of the HI signal have been obtained by many ongoing experiments, such as GMRT, MWA, PAPER, and LOFAR (Paciga et al. 2013; Ali et al. 2015; Beardsley et al. 2016; Patil et al. 2017), with the best upper limits of $\simeq (50 \text{ mK})^2$ for $k \simeq 0.1 \text{ Mpc}^{-1}$.

The recent detection of a broad global absorption trough of strength 500 mK by the Experiment to Detect the Global Epoch of reionization Signature (EDGES; Bowman et al. 2018) at $\nu \simeq 80 \pm 10$ MHz is the only positive detection of the HI signal at high redshifts. Observational projects that are attempting to detect the global signal (SARAS, Singh et al. 2018a, 2018b; LEDA, Bernardi et al. 2015; Price et al. 2018; BIGHORNS, Sokolowski et al. 2015; SCI-HI, Voytek et al. 2014) and its fluctuating component (HERA, LOFAR, MWA) might provide more insight into the physics of EoR in the near future. If confirmed, the unexpectedly deep absorption trough

detected by EDGES will also open avenues to investigate exotic physics (e.g., Barkana 2018; Fraser et al. 2018; Lambiase & Mohanty 2018; Muñoz & Loeb 2018).

Numerical simulations can provide us insight into the morphology and evolution of the sources in the early universe. However, given the uncertainty in the astrophysics of this epoch, it is useful to develop fast analytic methods that can analyze the signal for a large range of scales for different combinations of physics inputs and modeling parameters. In our previous work (Raste & Sethi 2018, RS18 from now on), we developed a formalism to analytically compute the autocorrelation and power spectrum of the HI signal in the early phase of cosmic dawn and EoR, when the medium is partially heated and ionized. For simplicity, we had assumed a complete coupling of hydrogen spin temperature T_S with matter kinetic temperature T_K . In this paper, we expand the formalism to include the effect of inhomogeneous Ly α coupling on the HI signal. We also apply our method to study the fluctuating signal that would correspond to the global HI signal observed by the EDGES group.

In the next section, we review the HI signal from the EoR and the expected photoionization and heating of the intergalactic medium (IGM) due to UV and X-ray. In Section 2.3, we model fluctuations in the signal due to inhomogeneous Ly α coupling. In Section 3, we briefly present the formalism for computing the two-point correlation function of the HI signal and discuss a few approximations. In Section 4, we present our results and explore their dependence on our modeling parameters. We summarize the derivation of our formalism in Appendix D. We present our conclusions in Section 5. Throughout this paper, we assume the spatially flat Λ CDM model with the following parameters: $\Omega_m = 0.310$, $\Omega_B = 0.049$, $h = 0.677$, and $n_s = 0.967$, with the overall normalization corresponding to $\sigma_8 = 0.808$ (Planck Collaboration et al. 2018).

2. Cosmic Dawn and EoR

The hyperfine splitting of neutral hydrogen in its ground state causes an energy difference corresponding to wavelength $\lambda = 21.1$ cm. The spin (or excitation) temperature of this line, T_S , is a function of the ratio of atoms in two hyperfine states. This ratio is determined by three processes in the early universe: absorption and stimulated emission of CMB photons at temperature T_{CMB} , collisions with atoms and charged particles, and the mixing of hyperfine levels owing to Ly α photons (Wouthuysen–Field effect). We can express T_S in terms of the color temperature of Ly α photons, T_α , gas kinetic temperature T_K , and T_{CMB} (Field 1958; Pritchard & Loeb 2012):

$$T_S = \frac{T_{\text{CMB}} + y_\alpha T_\alpha + y_c T_K}{1 + y_\alpha + y_c}. \quad (1)$$

Here, $y_c \propto n_H$, n_e (number density of neutral hydrogen atom or electrons; neutral atoms dominate for a small ionized fraction) and $y_\alpha \propto n_\alpha$ (number density of Ly α photons). During the dark ages, $1000 < z < 100$, T_S relaxes to T_{CMB} . After matter thermally decouples from CMB, for $100 < z < 30$, collisions couple T_S to the matter temperature T_K . When the first sources of radiation form during cosmic dawn, the production of Ly α photons couples the spin temperature to the color temperature of Ly α T_α , which is relaxed to T_K by multiple scattering of Ly α

photons with HI (e.g., Field 1959; Rybicki & dell’Antonio 1994; Chen & Miralda-Escudé 2004). If, at any redshift, $y_{\text{tot}} = y_c + y_\alpha \gtrsim T_{\text{CMB}}/T_K$, then T_S is strongly coupled to T_K . Otherwise, in the absence of these coupling mechanisms, it relaxes to T_{CMB} .

The HI is observable in emission or absorption depending on whether its spin temperature T_S is greater than or less than T_{CMB} . The CMB spectral distortion caused by this effect is observable and can be expressed as follows (e.g., Madau et al. 1997; Shaver et al. 1999; Gnedin & Shaver 2004; Sethi 2005; Pritchard & Loeb 2012):

$$\Delta T_b \simeq 26.25 n(1 + \delta) \left(1 - \frac{T_{\text{CMB}}}{T_S}\right) \times \left(\frac{1 + z}{10} \frac{0.14}{\Omega_m h^2}\right)^{\frac{1}{2}} \left(\frac{\Omega_b h^2}{0.022}\right) \text{mK}. \quad (2)$$

Here the redshift-space distortion is ignored. We have expressed HI number density as $n_H = \bar{n}_H n(1 + \delta)$, and the mean density \bar{n}_H has been absorbed in the prefactor of Equation (2). Here, δ is overdensity of the gas. We have assumed that a small volume at any point is either completely neutral or completely ionized, so a variable n is defined that is unity if the medium is neutral and zero otherwise. We further define dimensionless temperature fluctuation as (Zaldarriaga et al. 2004)

$$\psi = n(1 + \delta)(1 - s), \quad (3)$$

which captures the density (δ), ionization (n), and spin temperature, T_S , inhomogeneities. Here, $s = T_{\text{CMB}}/T_S$. These quantities are functions of position, and thus they contribute to the spatial fluctuation of the signal. We suppress this dependence for notational clarity.

At the end of the dark ages, high-density regions of the universe collapse and form structures of a range of masses. The radiation emitted by them change the properties of their surrounding medium. In our work, we assume that the smallest mass that can collapse corresponds to a HI-cooled halo (e.g., Barkana & Loeb 2001; Dayal & Ferrara 2018):

$$M_{\text{min}} = 3.915 \times 10^8 \frac{\Omega_m^{-1/2}}{h(1+z)^{3/2}} M_\odot. \quad (4)$$

We consider fluctuations of hydrogen-ionizing, Ly α , and X-ray radiation fields emitted by the sources on the brightness temperature inhomogeneities.

2.1. Photoionization

The hydrogen-ionizing (ultraviolet—UV) photons emitted from the star within the collapsed structures are absorbed in the immediate vicinity of the sources and carve out H II regions around them in the IGM. We use excursion set formalism to compute a size distribution of the ionization bubbles by defining self-ionized regions (Furlanetto et al. 2004b). Such regions have enough sources to ionize all gas within them. Their sizes are determined by the ionization efficiency factor, defined as $\zeta = 1/f_{\text{coll}} = M_{\text{tot}}/M_{\text{coll}}$. Here, f_{coll} is the fraction of collapsed mass inside the self-ionized region, and ζ is a

function of the properties of the sources as well as the surrounding halo (Furlanetto et al. 2004b):

$$\zeta = f_* f_{\text{esc}} N_{\text{ion}} N_{\text{rec}}^{-1}. \quad (5)$$

Here, f_* is the fraction of collapsed baryons that are converted into stars, f_{esc} is the fraction of ionizing photons that escape the source halo, N_{ion} is the number of UV photons created per stellar baryon, and N_{rec} is the number of recombinations. We assume ζ to be independent of redshift in this paper, even though it can evolve with time owing to the evolution of quantities used to define it. For higher values of ζ , the reionization is completed at higher redshift (e.g., RS18). These self-ionized regions are larger than the H II regions of a single source because they are created by highly clustered multiple sources in the early universe (for the Λ CDM model). For our work, we assume the region to be spherical.

2.2. X-Ray Heating

Photons of energy $E \gg 13.6$ eV (X-rays) escape the H II region into the surrounding medium. These photons ionize the neutral gas up to 10% and can impart up to 20% of their energy into heating the gas through photoionization and secondary collisional ionization and excitation (e.g., Shull & van Steenberg 1985; Venkatesan et al. 2001). In our study, we have assumed the medium outside the ionized region to be completely neutral ($n = 1$) because the fraction of ionization due to this process is generally small.

The photoionization cross section of X-ray photons falls as $\sigma_i(\nu) = \sigma_{i0}(\nu/\nu_i)^{-3}$, with ν_i being the ionization threshold frequency of species i . In this work, we only consider two species: neutral hydrogen and neutral helium, with their relative fractions $x_i = 12/13$ and $1/13$ of the baryon number, respectively. This is a valid assumption because the metallicity in the early universe is very low. Because the low-energy X-ray photons are absorbed with higher probability, they contribute to heating the medium immediately surrounding the H II region, whereas the higher energy photons free-stream through the medium and might get absorbed far away from any source. These photons uniformly heat up the whole IGM to some background temperature T_{bg} .

We assume the X-ray photon source luminosity to be given by a power law (e.g., Mesinger et al. 2011 and references therein), $\dot{N}_\nu = \dot{N}_i(\nu/\nu_{\text{min}})^{-\alpha}$, where ν_{min} is the lowest frequency of X-ray photons escaping from source halos, and N_{heat} is the number of X-ray photons emitted per stellar baryon. We assume that $f_H = 0.15$ is the fraction of energy of emitted photoelectrons that goes into heating the medium (Shull & van Steenberg 1985; Venkatesan et al. 2001). Other than the adiabatic expansion of the universe, we neglect all other cooling processes.

We divide the neutral hydrogen volume into two zones. In the *near zone*, the heating is dominated by X-ray photons from an individual self-ionized region. In the *far zone*, the contribution from all of the faraway background sources is taken into account. For more details and explanations, see RS18, where the increase in temperature due to a self-ionized region of radius R_x at a distance R_0 from the center of

the ionized region was calculated at redshift z'_c :

$$\Delta T' = \frac{hf_H \alpha N_{\text{heat}} f_* n_0 \nu_{\text{min}}^\alpha R_x^3}{3k_B \zeta R_0^2} (1 + z'_c)^2 \int_{t(z_*)}^{t(z_0)} dt' \frac{f_i(t') \dot{f}_{\text{coll,g}}}{f_i(t) f_{\text{coll,g}}} \left(\frac{1 + z'}{1 + z} \right)^{\alpha+1} \int_{\nu'_{\text{min}}}^{\infty} d\nu' \nu'^{-\alpha-4} e^{-\tau(R_0, \nu')} \sum_i (\nu' - \nu_i) x_i \sigma_{i0} \nu_i^3. \quad (6)$$

The primed quantities are calculated at the receiving point (point P), unprimed quantities are at the source (point S), and quantities with 0 subscript are comoving quantities. Here, α is the X-ray spectra power-law index, and f_i are global ionization fractions.

2.3. Ly α Radiation

For EoR studies, all of the radiation between Ly α and the Lyman limit is referred to as Ly α radiation emitted from the source, and we shall follow this convention. The Ly α contribution at any point arises from two main factors: Ly α emitted from the sources and Ly α created by X-ray photoelectrons (Venkatesan et al. 2001). The latter is generally small, and we neglect it in this paper.

Ly α photons emitted from the sources escape the surrounding H II region and redshift until their frequency nearly equals the resonant frequency of one of the Lyman series lines. When a photon redshifts to the frequency corresponding to one of the Lyman series lines, it gets scattered by the neutral hydrogen³ and eventually cascades to a frequency corresponding to Ly α . Given the complicated frequency structure of Lyman series lines, these photons are absorbed at varying distances from the source. Thus, the coupling depends on two factors: the region of influence of the Ly α radiation and the coupling coefficient y_α . Photons between Ly α and Ly β frequencies can be absorbed by H_2 molecules, but we ignore this effect as the density of H_2 is very low in the IGM.

We define the Ly- n influence region as the distance traveled by the Ly- $(n + 1)$ photons to redshift to Ly- n frequency. If these photons were emitted at $z = z_e$ and absorbed at $z = z_a$ with $\nu_e = \nu_{n+1}$ and $\nu_a = \nu_n$, then the comoving distance traveled by the photon before it is absorbed in an expanding universe is ($n \geq 2$)

$$R_{\text{max}}(n) \simeq \frac{16040 \text{ Mpc}}{(1 + z_e)^{1/2}} \left[\left(\frac{n^3(n+2)}{(n+1)^3(n-1)} \right)^{1/2} - 1 \right]. \quad (7)$$

These influence regions become smaller with increasing n . We note that $R_{\text{max}}(2)$ (Ly α influence region) is much larger than the mean distance between ionization bubbles at any redshift. For $\zeta = 7.5$, the values of mean comoving distance between bubbles for redshifts 25, 20, and 15 are 7.85 Mpc, 2.29 Mpc, and 0.96 Mpc, respectively. Therefore, Ly α regions are very large and merge very early. However, this would create homogeneous coupling to H I atoms only if the Ly α coupling coefficient, y_α , is high enough (Equation (1)); y_α is a function of Ly α photon (physical) number density n'_α (Field 1958;

³ The scattering cross section falls with increasing n , so this scenario is applicable if the optical depth of scattering in the expanding medium exceeds unity. The requirement is readily met for the transitions of interest ($n < 20$) in the paper.

Chen & Miralda-Escudé 2004):

$$y_\alpha = \frac{\sqrt{2}hc^4}{36\pi\nu_\alpha^2 k} \sqrt{\frac{m_H}{2kT_k}} \frac{A_\alpha}{A_{10}} \frac{\nu_{10}}{\nu_\alpha} \frac{n'_\alpha}{T_K}. \quad (8)$$

To calculate the number density of Ly α photons (n'_α) received at any point from ionizing sources, we use the method used to calculate X-ray heating in RS18. Assuming a flat spectrum between Ly α and the Lyman limit, the number density of Ly α photons at a comoving distance R_0 from the source is

$$n'_{\alpha,*} = \frac{\dot{N}_\alpha}{4\pi c R_0^2} \frac{2\Delta\nu_\alpha}{\nu_\beta - \nu_\alpha} \frac{(1+z')^3}{1+z}.$$

Here, $\Delta\nu_\alpha = \sqrt{2kT_k/m_H c^2} \nu_\alpha$ is the Doppler line width. This factor arises because at the source the photons are emitted with

$$k_{10}^e = \begin{cases} \exp\left(-9.607 + 0.5 \log(T_K) \exp\left(-\frac{(\log(T_K))^{4.5}}{1800}\right)\right) \text{ cm}^3 \text{ s}^{-1} & T_K \leq 10^4 \text{ K} \\ k_{10,e}(T_K = 10^4 \text{ K}) & T_K > 10^4 \text{ K} \end{cases} \quad (11)$$

frequencies between ν_β and ν_α , but the only frequencies that are absorbed at redshift z' are in the range of $\Delta\nu_\alpha$ around ν_α . The Ly α luminosity, \dot{N}_α , can be expressed in terms of the size of the ionization halo, assuming that the Ly α luminosity scales with ionizing luminosity with a factor f_L . Using the balance between ionization and recombination in the ionizing region, we have

$$\dot{N}_\alpha = f_L \frac{4\pi}{3} n_0^2 \alpha_B C R_x^3 (1+z)^3.$$

It would be reasonable to assume that $f_L > 1$ because Ly α photons escape the halo more easily than ionizing photons. However, for the sake of completeness, we take $0.1 < f_L < 1000$ in this paper. Combining all these, we get

$$y_{\alpha,*} T_k \simeq \frac{S_\alpha}{54\pi} \frac{hc^2 A_\alpha \nu_{10} n_0^2 \alpha_B}{\nu_\alpha^2 k A_{10} (\nu_\beta - \nu_\alpha)} f_L C \frac{R_x^3}{R_0^2} (1+z')^3 (1+z)^2. \quad (9)$$

Here, S_α is a correction factor of order unity as defined in Chen & Miralda-Escudé (2004), which depends on the photon spectrum around the Ly α line (Hirata 2006; Pritchard & Loeb 2012). This expression has been derived in a somewhat different manner in Pritchard & Loeb (2012).

We find contributions to y_α from higher-order Lyman transitions (e.g., from Ly γ to Ly β) in a similar way, by counting the number of Ly n influence regions the point falls within. The effect of higher-order transitions is expected to be subdominant because, for a continuum source, the total number of photons between Ly β to the Lyman limit is smaller than in the frequency range Ly β and Ly α . However, these photons can have a substantial impact near an ionizing source because they are absorbed closer to the source given their smaller influence regions. If the distance of a point from the source R_0 is such that $R_{\max}(n+1) < R_0 < R_{\max}(n)$, then the point in question will have n Ly α photons in its vicinity rather than one photon if only the transition between Ly α and Ly β is considered. This means that the Ly α flux from the source center generally falls more rapidly than $1/r^2$ when this effect is taken into account.

In this paper, we only consider Lyman series lines that have influence regions larger than the ionization bubble radius.⁴

2.4. Collisional Coupling

The collisional coupling of spin temperature T_S and matter kinetic temperature T_K due to scattering of neutral hydrogen and electrons (Equation (1)) can play an important role at lower redshift ($z \lesssim 30$) too. The coupling coefficient is proportional to the number density of colliding particles:

$$y_c = \frac{(n_H k_{10}^H + n_e k_{10}^e) T_x}{A_{10} T_K}. \quad (10)$$

For collision rate coefficients, we use the following fits (Zygelman 2005; Pritchard & Loeb 2012):

$$k_{10}^H = \begin{cases} 3.6 \times 10^{-16} T_K^{3.640} \exp\left(\frac{6.035}{T_K}\right) \text{ cm}^3 \text{ s}^{-1} & T_K \leq 10 \text{ K} \\ 3.1 \times 10^{-11} T_K^{0.357} \exp\left(-\frac{32}{T_K}\right) \text{ cm}^3 \text{ s}^{-1} & T_K > 10 \text{ K} \end{cases} \quad (12)$$

These rates increase with temperature, which means that there is stronger collisional coupling for hotter gas than for cool gas. This effect is important if the gas was colder during the cosmic dawn, due to unknown physics: the preionization absorption trough might be shallower instead of steeper, in spite of having a larger contrast of matter temperature from the CMB temperature. During EoR, the electron scattering is more effective near the sources where there is partial ionization and high temperature. However, since we assume the ionization fraction to be just the residual fraction outside of ionization bubbles, this effect is negligible in our work.

3. Autocorrelation of Dimensionless Brightness Temperature ψ

In our work, we wish to find the autocorrelation of ψ (Equation (3)), which can be defined as

$$\begin{aligned} \mu &= \langle \psi_1 \psi_2 \rangle - \langle \psi \rangle^2 \\ &= \langle n_1(1+\delta_1)(1-s_1)n_2(1+\delta_2)(1-s_2) \rangle \\ &\quad - \langle n_1(1+\delta_1)(1-s_1) \rangle^2. \end{aligned} \quad (13)$$

Here, (n_1, δ_1, s_1) and (n_2, δ_2, s_2) are values of ionization, overdensity, and heating (T_{CMB}/T_S) at point 1 (\mathbf{r}_1) and at point 2 (\mathbf{r}_2), respectively. Since the process of reionization is statistically

⁴ Photons with influence regions smaller than the ionization bubble will redshift to lower and lower Lyman series lines until they cross the ionization region boundary. However, their effect would be very small and very close to the boundary of the ionization region. It is not useful to model them in more detail because several other assumptions would break down so close to the boundary (e.g., sphericity of bubbles, sharp boundary of ionization regions).

homogeneous and isotropic, the autocorrelation function μ is a function of $r = |\mathbf{r}_2 - \mathbf{r}_1|$. To calculate μ , we need to find all of the pairs of points that are separated by a distance r , and take their weighted average over the entire space. To calculate the probability of a pair with certain values, we use geometric arguments as described in RS18 and briefly summarized in Appendix B.

In this paper, we assume that the density has no correlation with ionization or heating ($\langle n\delta \rangle = \langle s\delta \rangle = 0$).⁵ This gives us

$$\begin{aligned}\langle \psi_1 \psi_2 \rangle &= (1 + \xi) \langle n_1 n_2 (1 - s_1)(1 - s_2) \rangle, \\ \langle \psi \rangle &= \langle (1 + \delta) \langle n(1 - s) \rangle = f_n - \langle ns \rangle\end{aligned}$$

where $\langle \delta \rangle = 0$ and $\xi = \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle$ is the autocorrelation function of the HI density perturbation. We compute ξ using the Λ CDM model power spectrum and assume the relative bias between the dark matter and the HI, $b = 1$. Here, $f_n = \langle n \rangle$ is defined as the average neutral volume fraction at that redshift. We also define $\phi = n(1 - s)$ to explore simplifying cases where we temporarily ignore the effect of density correlation ξ . In such case, the two-point correlation function is

$$\mu = (1 + \xi) \langle n_1 n_2 (1 - s_1)(1 - s_2) \rangle - \langle n(1 - s) \rangle^2 \quad (14)$$

$$= (1 + \xi) \langle \phi_1 \phi_2 \rangle - \langle \phi \rangle^2. \quad (15)$$

The correlation functions we calculate in RS18 and this paper are analytically derived. However, a function $\mu(r)$ is a valid correlation function only if it follows certain properties: (1) it should be a finite positive value at $r = 0$, $\mu(0)$; (2) at any $r > 0$, the value of the correlation function $\mu(r) < \mu(0)$; (3) the correlation should go to zero at very large distance; (4) the Fourier transform of the correlation function should be a positive definite function (power spectrum); (5) between $r = 0$ and $r \rightarrow \infty$, the correlation function can be positive or negative, with the condition that its integration over all space must be zero. We do not satisfy this condition.⁶ However, we check for consistency of our formalism with the rest of the conditions.

We first discuss a few limiting cases (for details see RS18). For scales greater than the largest bubbles, the fluctuations due to the ionization, heating, and Ly α coupling inhomogeneities vanish, and the HI correlation function is determined by only density perturbations. In this limit we get

$$\mu = \xi (f_n - \langle s \rangle)^2. \quad (16)$$

Here the correlation function scales as the density correlation function ξ . We also note that the correlation function at large scale vanishes when $f_n = \langle s \rangle$ (close to the global heating transition). If the heating and Ly α coupling are uniform, the neutral gas of the IGM is at the same spin temperature T_{bg} . The correlation function simplifies in this case:

$$\mu = (1 - s_b)^2 ((1 + \xi) \langle n_1 n_2 \rangle - f_n^2). \quad (17)$$

Here, $s_b = T_{\text{CMB}}/T_{\text{bg}}$. If the ionization fraction is very small, $\langle n_1 n_2 \rangle \simeq f_n^2 \simeq 1$. This gives us $\mu = \xi(1 - s_b)^2$. Here the density fluctuations are enhanced by the temperature contrast

⁵ These cross-correlations can be computed using excursion set formalism, e.g., Furlanetto et al. (2004b), for density-ionization cross-correlation. These terms are generally subdominant to other terms we retain (for discussion, see, e.g., RS18 and references therein).

⁶ As discussed in RS18, to compute ionization inhomogeneities, we assume that the probability of finding an ionized region outside an ionization bubble is the global ionization fraction f_i . While this is an excellent assumption for computing the correlation function on scales of interest to us, it violates the integral constraint on the correlation function.

between the IGM and the CMB. At late times, $T_{\text{bg}} \gg T_{\text{CMB}}$ owing to X-ray heating, which drives s_b to zero. This reduces the correlation function to the one dominated by density and ionization inhomogeneities:

$$\mu = -f_n^2 + (1 + \xi) \langle n_1 n_2 \rangle \quad (18)$$

$$\langle n_1 n_2 \rangle = f_n - f_n \sum_{R_x} N(R_x) \frac{4\pi}{3} R_x^3 C(r, 0, R_x, R_x). \quad (19)$$

This result is derived and explored in Zaldarriaga et al. (2004) and RS18.

3.1. Modeling and Notations

Our aim in this paper is to analytically model the correlation function of HI brightness temperature fluctuations from the early phase of EoR owing to scales that emerge from ionization, heating, and Ly α coupling inhomogeneities. These inhomogeneities are caused by bubbles of a given size distribution, which evolves with time. These bubbles determine the scales of correlation. The details of our modeling and main assumptions are discussed extensively in RS18. We briefly summarize them here for the sake of completeness, as only a subset of the details are given in this paper. We have assumed a topology where there are isolated, spherical self-ionized bubbles, surrounded by isotropic, smooth T_S profiles that might overlap with one another and smoothly merge with the background. Given the statistical isotropy (we neglect redshift-space distortion) and homogeneity of the process of reionization and heating, our assumption of spherical bubbles and isotropic spin temperature profiles holds even though the individual bubbles might not be spherical.

We compute ionization and spin temperature autocorrelation and ionization–spin temperature cross-correlation. We neglect, as noted above, the cross-correlation of density with ionization and spin temperature inhomogeneities because these contribute negligibly on the scales of interest. We neglect the clustering of self-ionized regions; a self-ionized region already accounts for clustering of ionizing sources at smaller scales. In RS18 we present cases that account for the clustering of self-ionized regions. This does not substantially alter our main results, but it might introduce new scales in the problem that correspond to the correlation scales of bubble centers. We note that this assumption is better for higher redshifts as the mean bubble separation is larger.

At any redshift, using excursion set formalism (Section 2.1) and the matter power spectrum given by the Λ CDM model, we generate a size distribution of self-ionized bubbles. Using Equations (1), (6), (9), and (10), we calculate the spin temperature in shells around these bubbles. The ionization volume fraction and volume fraction due to these shells is, respectively,

$$f_i = \sum_{R_x} \frac{4\pi}{3} N(R_x) R_x^3 \quad (20)$$

$$f_{hb} = \sum_{R_x} \frac{4\pi}{3} N(R_x) (R_h^3 - R_x^3), \quad (21)$$

where R_x are radii of ionization bubbles, $N(R_x)$ is the number density of bubbles with ionization radius R_x , and R_h corresponds to the outer radius of the spin temperature profile for a given R_x (Figure 1). During the initial phase, the ionization bubbles and surrounding temperature profiles are

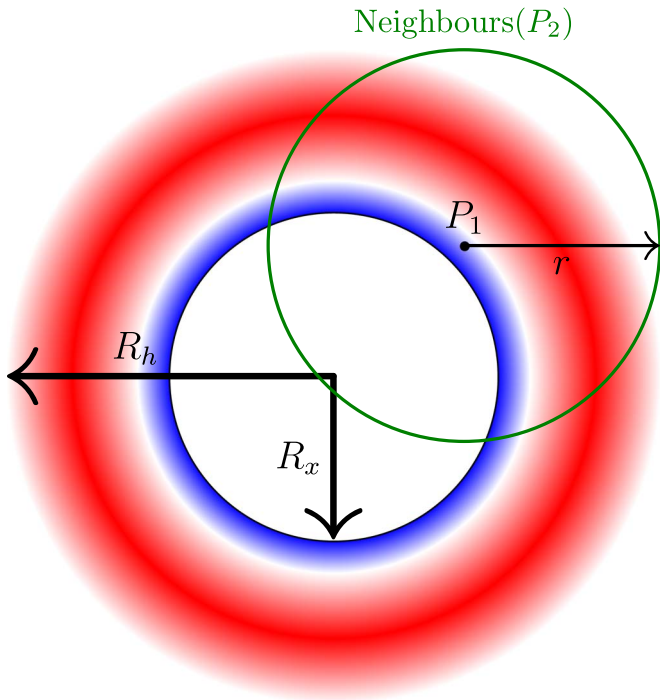


Figure 1. Cartoon for topology of the ionized region and its surrounding IGM. The color scheme shows the dimensionless brightness temperature, ψ ; $\psi_i = 0$ in the ionized region of size R_x (with a sharp boundary). It might be positive in the neutral, heated, and coupled region, and negative in the neutral, nonheated, and coupled region. The profile of radius R_h merges smoothly with the background, which is not coupled in this case.

nonoverlapping; we discuss the case of overlapping heating and Ly α coupled regions later in this section.

For any bubble of size R_x , we take shells of thickness $\Delta R(R_x, s)$ between radii R_x and R_h , having temperature $s = T_{\text{CMB}}/T_S$. A detailed description of notations followed in this paper is given in Table 1. To compute correlations, we assume two random points separated by a distance r , as shown in Figure 1. The formalism used for the computation of correlation functions is described in Appendix D.

3.2. Overlap

At low redshift, spin temperature profiles around ionization bubbles are very large and overlap significantly. Thus, in such cases, f_{hb} can be much larger than 1. In our previous paper (RS18), we had discussed a method to consistently take into account the overlap of heated regions. In this paper, our principal aim is to model the fluctuation of Ly α radiation. Unlike the heated regions that merge after the heating transition, the mean free path of lower Lyman series photons is larger than the mean interbubble distance at all times for $z < 30$, the starting redshift of our study. However, when the overlapped volume is very large ($f_{hb} \gg 1$), our formalism fails to generate a valid correlation function, as its Fourier transform (Appendix B) could fluctuate and yield negative values at large k . To avoid such unphysical results, we have taken a different approach to modeling overlaps in this paper.

We calculate kinetic temperature and Ly α profiles up to very large distances and start shedding outermost shells until f_{hb} approaches 1. The energy and Ly α photons from excess shells are uniformly distributed in the neutral universe (background as well as remaining shells). The bubbles are still likely to overlap

Table 1
Notations

Symbols	Explanation
δ	Overdensity of H I gas
n	Ionization state of H I gas: neutral point $n = 1$ and ionized point $n = 0$
s	Temperature state defined as $s = T_{\text{CMB}}/T_S$
ψ	Dimensionless brightness temperature: $\psi = n(1 + \delta)(1 - s)$
ξ	Autocorrelation of overdensity δ : $\xi = \langle \delta_1 \delta_2 \rangle$
ϕ	$\phi = n(1 - s)$
μ	Autocorrelation of dimensionless brightness temperature ψ : $\mu = \langle \psi_1 \psi_2 \rangle - \langle \psi \rangle^2$
f_i	Average ionized volume fraction
f_n	Average neutral volume fraction
f_{hb}	Total volume fraction due to T_S profiles (without correcting for the overlaps)
f_h	Average T_S profile volume fraction after correcting for the overlaps
f_b	Average background volume fraction
R_x	Radius of given ionization bubble
R_h	Outer radius of given T_S profile: $R_h = R_i(R_x)$
R_s	Inner radius of the shell with spin temperature $T_S = T_{\text{CMB}}/s$ around a given bubble
ΔR_s	Thickness of the shell with spin temperature $T_S = T_{\text{CMB}}/s$ around a given bubble
$N(R_x)$	Number density of ionization bubbles of radius R_x

because of randomness in their positions. To account for that, we use (Appendix C)

$$f_h \simeq \left(1 + \frac{f_i}{f_{hb}}\right)(1 - e^{-f_{hb}}) - f_i \quad (22)$$

$$f_b = 1 - f_i - f_h, \quad (23)$$

where f_h corresponds to the actual volume fraction occupied by heating bubbles (single counting of overlapped region), and f_b is the background fraction. Note that f_h approaches f_{hb} when the T_S profile volume fraction and ionization fraction are small. However, f_h remains less than unity even if the value of f_{hb} becomes much larger than unity.

3.3. Modeling Parameters

In our analysis, we explore two parameters to model X-ray heating and Ly α coupling:

1. N_{heat} : Number of X-ray photons emitted per stellar baryon. For our study, we assume N_{heat} in the range 0.1–10.0. For larger values of N_{heat} , the heating is stronger, with higher values of T_K .
2. f_L : Ratio of source luminosity of photons between Ly α and the Lyman limit to the luminosity of UV photons. We take $0.1 < f_L < 1000$ in this paper. As the value of f_L increases, the coupling between T_S and T_K is stronger. This increases the absolute value of ΔT_B in both emission and absorption.

In this paper, we have not considered scenarios in which the modeling parameters evolve with time.

4. Results

In the early phase of EoR, the brightness temperature fluctuations are determined by perturbations on the scales of

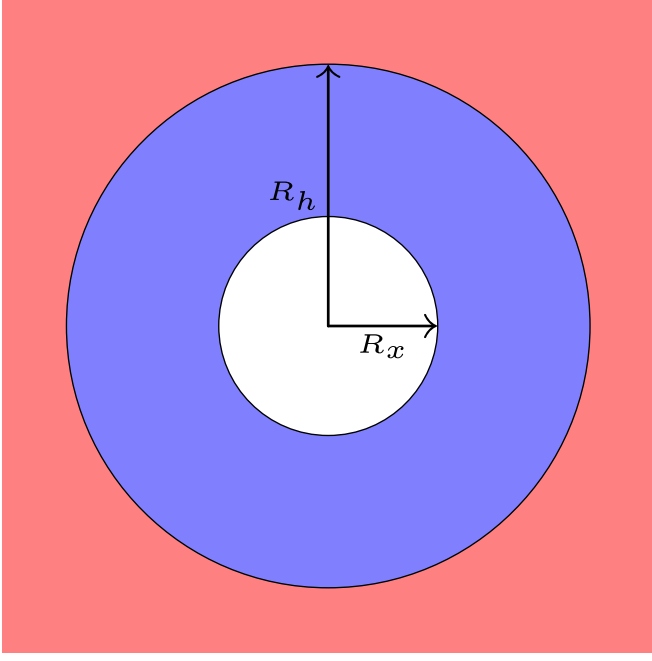


Figure 2. A simple case: an ionization bubble ($\psi_i = 0$) has one heated T_S shell around it. The background is unheated but coupled.

heated and Ly α coupled regions. In the complete model, these regions have a range of sizes and have diffuse profiles (Appendix D). This makes it difficult to disentangle the effect of input physics on the observable quantity. Therefore, to delineate the essential aspects of our formalism, we first consider two simple models.

4.1. A Simple Model: One Bubble Size, Flat Heating Profile

We can study a simple model with a single bubble size and a shell with uniform spin temperature around it (flat profile). There are small ionization bubbles embedded in larger heated bubbles (Figure 2). Ignoring density perturbations, there are only three values of $\phi = n(1 - T_{\text{CMB}}/T_S)$ in this universe: $\phi_i = 0$, $\phi_h = 1 - T_{\text{CMB}}/T_{\text{heat}}$, and $\phi_b = 1 - T_{\text{CMB}}/T_{\text{bg}}$. When we include the density perturbations, the correlation function can be written as

$$\begin{aligned} \mu = & (1 + \xi)(\psi_h^2(f_h - f_h f_b C(r, R_x, R_h, R_h)) \\ & - f_i f_h C(r, 0, R_x, R_h) - f_i(1 - f_i) \\ & \times (C(r, 0, R_x, R_x) - C(r, 0, R_x, R_h))) \\ & + 2\psi_h \psi_b f_h f_b C(r, R_x, R_h, R_h) \\ & + \psi_b^2 f_b (1 - f_h C(r, R_x, R_h, R_h) \\ & - f_i C(r, 0, R_x, R_h)) - (\psi_b f_b + \psi_h f_h)^2, \end{aligned} \quad (24)$$

where R_x is the ionization bubble radius and R_h is the heating bubble radius. Here, $C(x, P, Q, R)$ is a function with value between 0 and 1 (Appendix B). This result was derived and analyzed in detail for various limiting cases in RS18. This simple case does not allow for a negative correlation because, within the bubble, the ψ is positively correlated, and it is not correlated outside the bubble.

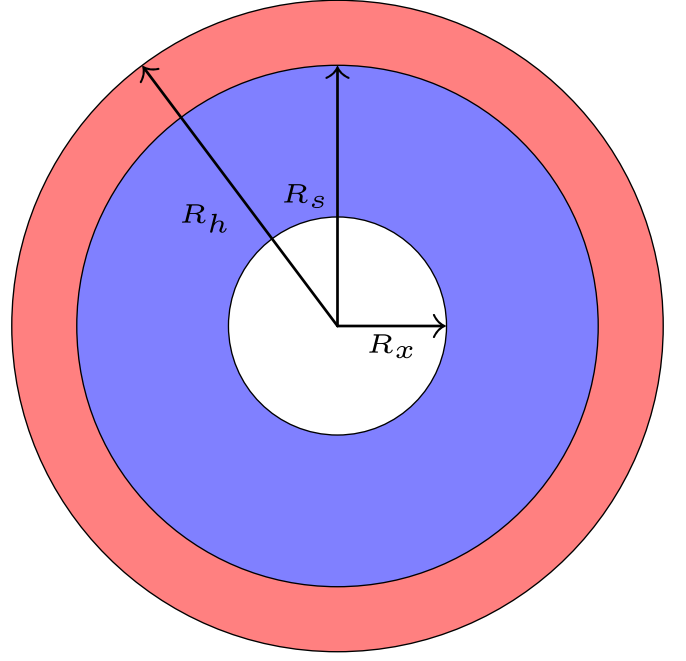


Figure 3. A simple case: an ionization bubble ($\psi_i = 0$) has two T_S shells around it. The background is uncoupled ($\psi_b = 0$).

4.2. A Simple Model: One Bubble Size, Two Shells, $\langle \psi \rangle = 0$

In the case where Ly α coupling is inhomogeneous, we can construct another simple scenario where there is only one ionized bubble size and the profile around the sources has two shells. The first shell is heated and coupled through Ly α radiation. The second shell is nonheated but still coupled. The region outside the second shell (background region) is neither heated nor coupled (Figure 3). This gives three values of $\phi = n(1 - T_{\text{CMB}}/T_S)$ when ignoring density fluctuations: $\phi_i = 0$ in the ionized regions, $\phi_1 = \phi > 0$ in the first shell, $\phi_2 = -\phi < 0$ in the second shell, and $\phi_b = 0$ in the background region. For simplicity, we have assumed that the volume occupied by the two shells is the same, $f_1 = f_2 = f_h/2$ and $\phi_1 = -\phi_2 = \phi$. Therefore, $\langle \phi \rangle = f_1 \phi_1 + f_2 \phi_2 = 0$.

From Figure 3, we can see that the correlation will be positive if both points are either in shell 1 or both in shell 2. If one point is in shell 1 and another in shell 2, then the correlation is negative. The two-point correlation function without density perturbation is

$$\begin{aligned} \mu = & \phi_1^2 P(\phi_1 \cap \phi_1) + \phi_2^2 P(\phi_2 \cap \phi_2) + 2\phi_1 \phi_2 P(\phi_1 \cap \phi_2) \\ = & \phi^2 (P(\phi_1) - P(\phi_1 \cap \phi_1) - P(\phi_1 \cap \phi_2) - P(\phi_1 \cap \phi_2) \\ & + P(\phi_2) - P(\phi_2 \cap \phi_1) - P(\phi_2 \cap \phi_2) \\ & - P(\phi_2 \cap \phi_1) - 2P(\phi_1 \cap \phi_2)), \end{aligned}$$

which can be expanded using the same logic used in RS18 and Appendix D:

$$\begin{aligned} \frac{\mu}{\phi^2} = & f_h (1 - C(r, 0, R_x, R_h)) - f_i (1 - f_i) \\ & \times [C(r, 0, R_x, R_x) - C(r, 0, R_x, R_h)] - (1 - f_i) \frac{f_h}{2} \\ & \times [C(r, R_s, R_h, R_h) - C(r, R_x, R_s, R_h)] \\ & - (1 - f_i) f_h [C(r, R_x, R_s, R_s) + C(r, R_s, R_h, R_x) \\ & - C(r, R_s, R_h, R_s) - C(r, 0, R_x, R_h)]. \end{aligned} \quad (25)$$

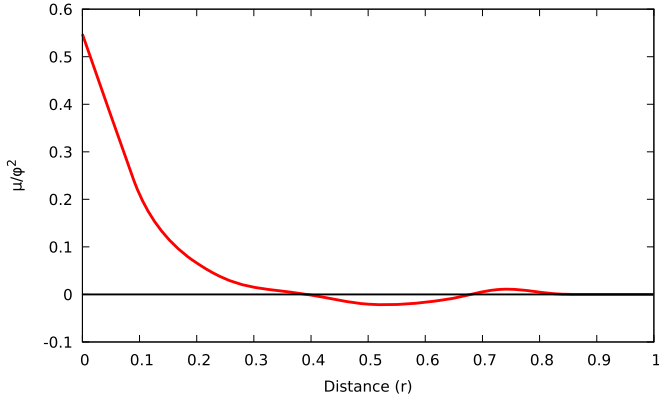


Figure 4. Case with a negative correlation at certain scales. Here, the ionization fraction $f_i = 0.01$ and the heated fraction is $f_h = 0.55$. The distance scale would depend on the size of the ionization bubble, which is chosen to be 0.1 Mpc.

This result can also be obtained by simplifying the complete model (Equation (41)) for approximations used in this section. At large scales, all of the functions $C(., ., ., .)$ tend to unity. In this case, Equation (25) approaches zero. This is expected because, at large scale ($r > 2R_h$), the two-point correlation should vanish if we ignore density correlation. When $r = 0$, $\mu = f_h \phi^2$. This is also as expected because there is no probability of two points being in different shells at $r = 0$. This expression takes a negative value for a range of values of r , where the two points are more likely to be in different shells than in the same shell.

In Figure 4, we have shown a case for this simple model where the autocorrelation of dimensionless brightness temperature is negative at certain scales. For scales $R_x + R_s < r < R_s + R_h$, depending on the values of various radii, two randomly chosen points can have a higher probability of being in different shells than in the same shell, driving the correlation to a negative value. For larger scales, $R_s + R_h < r < 2R_h$, both points have a finite probability of being in the outer shell and a zero probability of being in different shells of the same bubble, which leads the overall correlation at these scales to be positive. On scales $r > 2R_h$, the correlation function is zero. We note that the possibility of negative correlation at any scale depends on values of R_x , R_s , and R_h and their differences. It is entirely possible to have models where the correlation function remains positive at all scales of interest.

For the complete model (discussed in detail in the next section and Appendix D), identifying the scales is more difficult because the ionization bubbles have a range of sizes, and the spin temperature profiles have a number of shells. However, if we ignore ξ , which is positively correlated at all scales of interest, we can still identify cases where the autocorrelation of ϕ goes negative. We show one such case in Figure 5. This constitutes the first important result of our analysis, which underlines the importance of correlation analysis in real space to extract the relevant physics. If the correlation function is negative at certain scales, it will point to a very specific geometry of the heated and Ly α coupled regions surrounding the self-ionized bubbles.

4.3. Complete Model

In this paper, we have explored the correlation function and power spectrum evolution due to inhomogeneity of spin

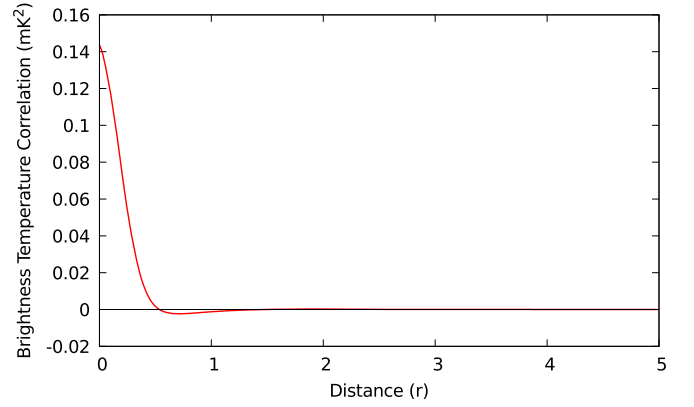


Figure 5. Neglecting density perturbations ($\xi = 0$), the autocorrelation of H I brightness temperature is shown at $z = 20$ for the complete model for $\zeta = 7.5$, $f_L = 1.0$, and $N_{\text{heat}} = 1.0$.

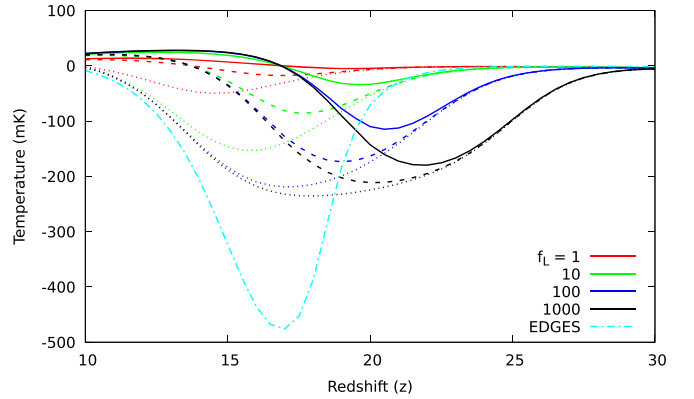


Figure 6. Global brightness temperature as a function of redshift for various values of f_L ranging from 0.1 to 1000. The solid lines are for $N_{\text{heat}} = 10$, long dashed lines for $N_{\text{heat}} = 1.0$, and short dashed lines for $N_{\text{heat}} = 0.1$. All plots have $\zeta = 7.5$. The dotted–dashed line represents a fiducial model that matches EDGES observations.

temperature using two modeling parameters: number of X-ray photons per stellar baryon N_{heat} and the ratio of luminosity of Ly α photons to ionizing photons of the sources f_L . We have taken $\zeta \sim 7.5$ (Equation (5)), which is in agreement with the reionization optical depth ($\tau_{\text{reion}} \simeq 0.055$) given by Planck (Planck Collaboration et al. 2018). We first present results for the Λ CDM model in the redshift range 10–30, without any additional cooling mechanism needed to explain the recent EDGES results (Bowman et al. 2018). The results below redshift $z \simeq 12$ are not entirely reliable because several approximations used in our formalism (including excursion set formalism) become less valid and eventually break down when the ionization volume fraction is large ($f_i > 0.1$; Furlanetto & Oh 2016; Giri et al. 2018). As noted above, to compute the brightness temperature mean and fluctuations, we have taken a size distribution of ionization bubbles, given by excursion set formalism and calculated spin temperature profiles surrounding these bubbles with a number of shells.

In Figure 6, we show the evolution of global H I brightness temperature as a function of redshift for various combinations of modeling parameters. At $z \simeq 30$, the global signal starts slightly negative as weak collisional coupling drives the spin temperature toward the matter temperature, which is below the CMB temperature (Equations (1) and (2)). At smaller values of z , the behavior is completely determined by the modeling

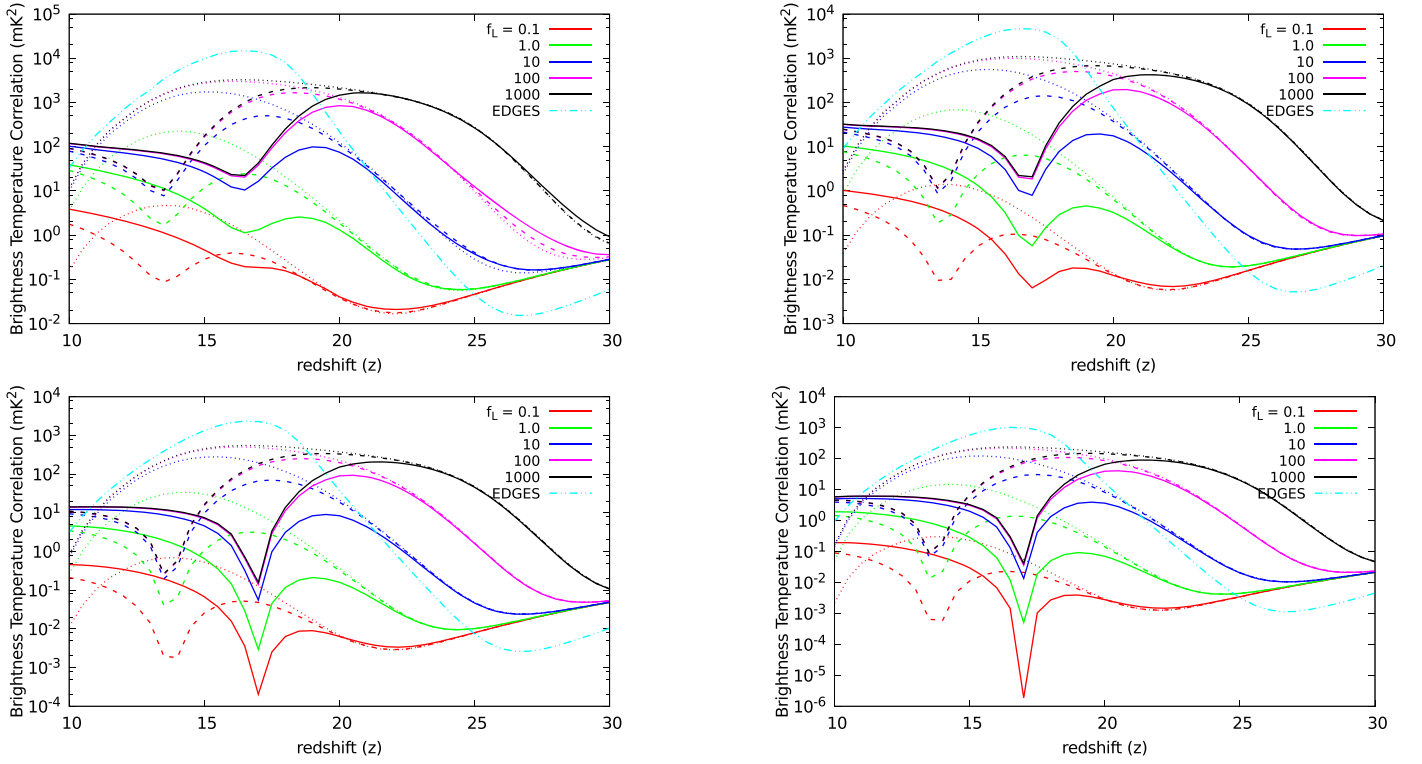


Figure 7. Evolution of autocorrelation of H I brightness temperature for $r = 0.5$ Mpc (top left panel), $r = 2.0$ Mpc (top right panel), $r = 4$ Mpc (bottom left panel), and $r = 8$ Mpc (bottom right panel) for $\zeta = 7.5$, a range of f_L varying from 0.1 to 1000 and three values of N_{heat} : 10 (solid lines), 1.0 (long dashed lines), and 0.1 (short dashed lines). The dotted–dashed lines represent a fiducial model that matches EDGES observations.

parameters. For higher values of N_{heat} , the heating starts earlier and the absorption troughs are shallower. For higher values of f_L , the coupling starts earlier as well as stronger, and the overall strength of the signal is larger at all redshifts before complete coupling is achieved. The redshift of the heating transition (when the average gas temperature $T_K = T_{\text{CMB}}$ and the signal goes from absorption to emission) is only dependent on the heating (N_{heat}) and is independent of Ly α coupling (f_L). It happens sooner for higher values of N_{heat} .

In Figure 7, we show the evolution of the correlation function at scales $r = 0.5$ –8 Mpc for different values of N_{heat} and f_L , using Equations (2), (6), (9), and (41). The correlation functions are large at small scales and decrease as the distance between two points increases. At very large scale, they approach Equation (16), where the density inhomogeneity is enhanced by the average of $(1 - T_{\text{CMB}}/T_S)$. On intermediate scales, the structure of the correlation function is determined by the size distribution of bubbles and their surrounding T_S profiles.

As the main aim of this paper is to model brightness temperature fluctuations owing to incomplete Ly α coupling, we first discuss it qualitatively. As in the case of incomplete heating, the Ly α coupling can also be separated into near and far zones. In the near zone, the emission from a nearby source dominates the coupling. The strength of this coupling is determined by Equation (9), which shows that there always is a distance from the source at which complete Ly α coupling ($\nu_{\alpha,*} T_k > T_{\text{CMB}}$) can be established. Two conditions have to be met to ensure this creates a new length scale in the correlation function. First, this distance must exceed the size of the ionization region. Because the coupling strength scales as R_x^3 , larger ionization bubbles meet this requirement more easily.

Second, the Ly α flux at this point from all of the background sources must be smaller than the flux from the nearby source, or we are in the limit of homogeneous coupling. The background intensity at any point is proportional to $R_{\text{max}}(n)$ (Equation (7)) multiplied by the number density of sources at any redshift. In the initial phase, the background intensity is not high enough to cause complete Ly α coupling. During this phase, there are regions around individual sources that attain complete coupling, and these regions are surrounded by a background in which only partial coupling has been attained. This creates inhomogeneities in Ly α coupling on the scales of these regions. As the intensity builds in the background owing to the birth of new self-ionized regions and it reaches levels sufficient to cause complete coupling, these inhomogeneities disappear. The nature of these inhomogeneities is determined by f_L , the atomic structure of neutral hydrogen, and the excursion set formalism.

At $z \simeq 30$, only the collisional coupling is effective. Because it is weak and the Ly α intensity is small, all of the curves shown in the figure start with small values of the correlation function with similar strength. The fluctuations are dominated by density perturbations because the number density of ionizing sources is small and the ionization, heating, and Ly α coupled fractions are tiny. For small f_L , the correlation declines with time as the collisional coupling weakens, owing to the fall of the number density of particles with the expansion of the universe. This situation is only reversed when Ly α coupling becomes efficient. For higher values of f_L , this coupling occurs sooner, leading to an increase in fluctuations with time until complete coupling is achieved. After this period, the fluctuations are determined by heating and ionization inhomogeneities. The position of the first peak in

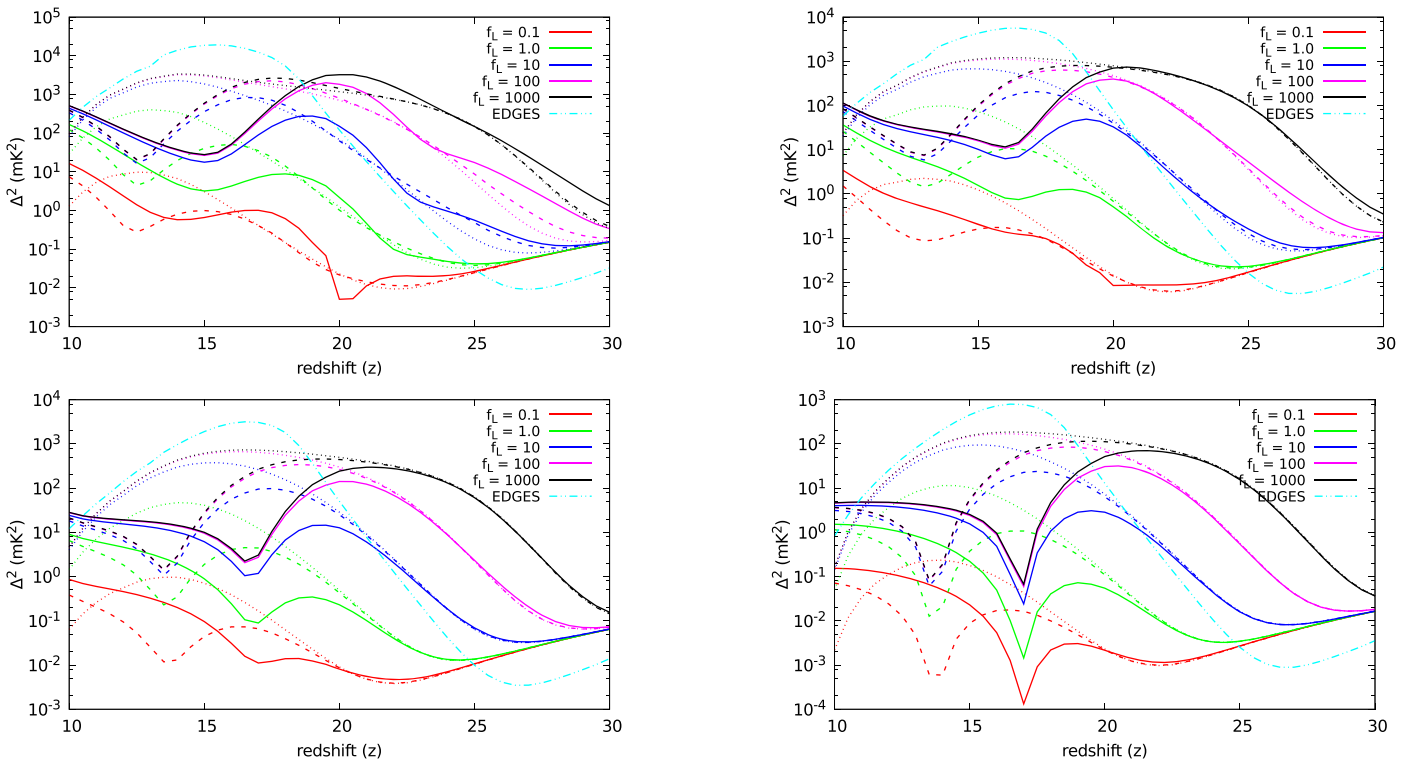


Figure 8. Evolution of $\Delta^2 = k^3 P(k)/2\pi^2$ of H I brightness temperature for $k = 2 \text{ Mpc}^{-1}$ (top left panel), $k = 1 \text{ Mpc}^{-1}$ (top right panel), $k = 0.5 \text{ Mpc}^{-1}$ (bottom left panel), and $k = 0.125 \text{ Mpc}^{-1}$ (bottom right panel) for $\zeta = 7.5$, a range of f_L varying from 0.1 to 1000, and three values of N_{heat} : 10 (solid lines), 1.0 (long dashed lines), and 0.1 (short dashed lines). The dotted–dashed lines represent a fiducial model that matches EDGES observations.

Figure 7 depends strongly on the heating parameter as the peak is determined by the thermal evolution of the background; the correlation only starts decreasing when the background is sufficiently heated. This phase is described in detail in RS18.

The correlation function approaches zero near the heating transition. At large scales, the signal vanishes completely when T_{CMB}/T_K approaches f_n (Equation (16)). We again see the effect of N_{heat} on the redshift of the heating transition. Inhomogeneous collisional coupling and the shape of the temperature profiles, which are determined by the spectrum of X-ray photons (α and ν_{min} , see RS18), can potentially change the redshift and depth of the heating transition by a small amount. We do not study this effect in the current paper. After the heating transition, the effect of inhomogeneous T_S decreases, and the main source of fluctuations is ionization inhomogeneities. Their effect is suppressed if the heating or coupling is not saturated (very small values of N_{heat} and f_L). In general, larger N_{heat} creates larger heating profiles and causes correlation at larger scales. However, these profiles also merge sooner and wipe out heating fluctuations at those scales.

The HI signal from the epoch of cosmic dawn and reionization has been extensively studied in the literature using semianalytic methods and large-scale simulations (e.g., Pritchard & Furlanetto 2007; Santos et al. 2008, 2010; Baek et al. 2010; Visbal et al. 2012; Mesinger et al. 2013, 2016; Tashiro & Sugiyama 2013; Pacucci et al. 2014; Fialkov et al. 2015, 2017; Ghara et al. 2015; Ross et al. 2017). Our formalism has been built on geometric arguments, which are intuitively easier to visualize in real space. The entire correlation structure can be written in terms of a single polynomial function: $C(., ., ., .)$ (Appendix B). Taking a Fourier transform with respect to the

first argument r of this function yields the power spectrum. In Figure 8, we show the evolution of the power spectrum $\Delta^2 = k^3 P(k)/2\pi^2$ for a range of k . These figures show an evolutionary trend similar to the correlation functions (Figure 7).

While comparing our results with simulations, we have focused on three features: (1) the number of peaks in the power spectrum, (2) the amplitude of $\Delta^2(k)$ for a range of scales $k \sim 0.1\text{--}0.5 \text{ Mpc}^{-1}$, and (3) the difference between the redshift of the heating transition and the redshift of the power spectrum minimum.

Existing results in the literature show that, for $k \simeq 0.1\text{--}0.5 \text{ Mpc}^{-1}$, there are generally two or three peaks of a power spectrum as a function of redshift (Santos et al. 2008; Baek et al. 2010; Mesinger et al. 2013, 2016; Ghara et al. 2015; Fialkov et al. 2017). At high redshifts, when the Ly α coupling and X-ray heating commence, they create fluctuations of T_S in the medium. If fluctuations in these two fields dominate at widely different times, there will be two distinct peaks at high redshift: one due to coupling inhomogeneities and the other (generally at lower redshift than the former) due to heating inhomogeneities (Pritchard & Furlanetto 2007; Chen & Miralda-Escudé 2008; Ahn et al. 2015). After the heating transition, there is a third, smaller peak at low redshifts, when the power spectrum is dominated by ionization inhomogeneities (e.g., Pritchard & Furlanetto 2007; Fialkov et al. 2014; Ghara et al. 2015). We studied the two peaks that are due to heating and ionization inhomogeneities in RS18.

The third peak owing to the inhomogeneous Ly α coupling is expected at early times for the following reason: as large self-ionizing regions are born, these fluctuations should initially build and then diminish as the contrast between Ly α coupling

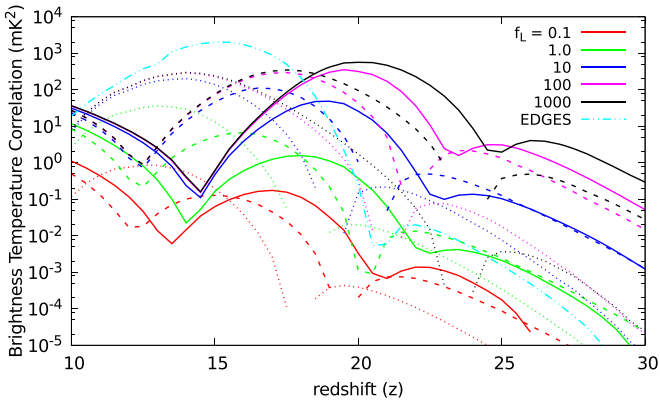


Figure 9. Neglecting density perturbations, the evolution of autocorrelation of HI brightness temperature is shown for $r = 0.5$ Mpc for $\zeta = 7.5$, a range of f_L varying from 0.1 to 1000, and three values of N_{heat} : 10 (solid lines), 1.0 (long dashed curves), and 0.1 (short dashed lines). The dotted-dashed line represents a fiducial model that matches EDGES observations.

in the near and far zone reduces, finally disappearing when complete Ly α coupling is established. We see this additional peak in Figure 9 in which the effect of density perturbations is neglected. However, in the complete model, we find a weak peak owing to this effect only at small scales in the power spectrum ($k = 2 \text{ Mpc}^{-1}$ in Figure 8). Generally this possible additional peak is masked by density perturbations. The strength of this peak can be understood in terms of the evolution of number density of large self-ionized regions at early times and the influence region of Ly α photons. For $z < 30$, the number density of the self-ionizing bubble builds exponentially in the excursion set formalism. While this creates inhomogeneities owing to the geometry seen in Figure 3, it also causes a rapid build-up of the background Ly α photons, rapidly destroying the contrast between the near and far zones. At any point, the background flux gets a nearly equal contribution from sources within the (comoving) radius $R_{\text{max}}(n)$ (Equation (7)). This radius is close to 600 comoving Mpc for $n = 2$ at $z \simeq 25$. This large influence region contributes to wiping out the contrast in a short span. We note that $R_{\text{max}}(n)$ (for small principal quantum numbers $n > 2$) determines the length scale at which the physics needs to be captured to study the Ly α generated inhomogeneities. It might be difficult to achieve it using an N -body simulation as the box size is generally smaller than this length scale. In our results for large N_{heat} and small f_L , we get three peaks at large k (top left panel of Figure 8) because the heating inhomogeneities, and by extension collisional coupling inhomogeneities, dominate before the Ly α inhomogeneities commence.

While $\Delta^2(k)$ agrees with the results of simulations for smaller scales, for $k \simeq 0.1 \text{ Mpc}^{-1}$, our results give less power as compared to simulations (bottom right panel of Figure 8). Given the small sizes of ionization bubbles at high redshifts, the only contribution to fluctuations at $k \sim 0.1 \text{ Mpc}^{-1}$ is due to density and spin temperature fluctuations. For the models of heating and Ly α coupling used in this paper, the contribution due to heating and coupling is dominated by faraway sources, which diminishes fluctuations at large scales. Therefore, at these scales, the spin temperature inhomogeneities are negligible, and only the density fluctuations are enhanced by the average contrast of HI spin temperature with CMB temperature. While it is conceivable that the higher power at large scales in the simulation is due to the finite box size (for a

discussion, see Zahn et al. 2011; Ghara et al. 2015), which might not allow one to take into account the contribution of faraway sources whose effect tends to homogenize the fluctuations of T_s at large scales, a more detailed comparison with simulations is hard as the parameter range used is generally not the same.

When the average spin temperature equals the CMB temperature during the heating transition, the power spectrum at small k reaches a minimum value (Figure 8). However, for larger k (small r), even during the heating transition there are significant fluctuations due to inhomogeneities of spin temperature and ionization, which delay the minima of the power spectrum. Figure 8 shows that, for $k = 2 \text{ Mpc}^{-1}$, the minimum of the power spectrum depends on the value of f_L even though the heating transition is independent of it. In general, the minimum of the power spectrum occurs during or after the global heating transition, depending on the scales and modeling parameters. This is in agreement with simulations that explore the dependence of these inhomogeneities on modeling parameters (Mesinger et al. 2013, 2016; Ghara et al. 2015; Fialkov et al. 2017).

4.4. Implication of EDGES Detection

Recent EDGES observations (Bowman et al. 2018) reported a sky-averaged absorption feature of strength $\Delta T \simeq -500$ mK in the frequency range 70–90 MHz, corresponding to a redshift range 15–19 for the redshifted HI line. It can be shown that for a standard recombination and thermal history, the minimum temperature of the gas at $z \simeq 17$ is $T_K \simeq 6$ K. It follows from Equations (1) and (2) that the absorption trough should not have been deeper than -180 mK.

One possible explanation of the EDGES result is that there is additional radio background in the redshift range $15 < z < 19$; in this case, we can replace T_{CMB} with the $T_{\text{CMB}} + T_{\text{radio}}$ in Equation (1) in the relevant redshift range (Ewall-Wice et al. 2018; Feng & Holder 2018; Sharma 2018). With this replacement and a suitable choice of T_{radio} , we can rederive all our results in this paper for compatibility with the EDGES result.

Another plausible explanation invokes the additional cooling of baryons that is due to interaction between dark matter and baryons.⁷ In this case, we can explain the EDGES detection using Equations (1) and (2) if (1) Ly α photons globally couple the spin temperature to matter temperature, that is, $T_K y_{\alpha} \gg T_{\text{CMB}}$, such that $T_s = T_K$ at $z \simeq 19$; and (2) $T_K \simeq 2.5$ K.⁸ We consider this theoretical extension to explain the EDGES result in this paper.⁹

We consider a fiducial model that roughly fits the EDGES results. We assume dark matter–baryon interaction of the form

⁷ It might be possible to distinguish this scenario from the one invoking a higher radio background if the radio background leaves signatures of its characteristic fluctuations due to radio sources. We hope to return to this issue in a future work.

⁸ EDGES detection implies a sharp trough in the signal at $z \simeq 19$ and an equally sharp rise at $z \simeq 15$. As the noise level for the detection is $\simeq 20$ mK, the drop at higher redshift can arise from complete Ly α coupling being established close to $z \simeq 19$, with the rapid heating being responsible for the sharp rise at smaller redshift. It should be noted that one of the implications of the EDGES results is that f_L cannot be too large, otherwise the complete Ly α coupling would be established at a higher redshift, and even in the absence of additional cooling the signal would be close to -200 mK at $z > 19$, which should be observable but is not seen by EDGES.

⁹ Other possible explanations of this result include a possible systematic error (Hills et al. 2018) and absorption from spinning dust grains in the Galactic ISM (Draine & Miralda-Escudé 2018), among others.

described in Barkana (2018), with a cross section $\sigma_1 = 5 \times 10^{-24} \text{ cm}^2$ and the ratio of dark matter to proton mass $m_{dm}/m_p = 0.001$. Such an interaction helps cool the baryon gas temperature sufficiently to explain the EDGES result. For $f_L=2$ and $N_{\text{heat}} = 0.08$, we get an absorption trough in the global signal similar to the one observed by EDGES data (Figure 6). In Figures 7 and 8, we show the correlation functions and power spectrum for a range of scales, taking into account our fiducial model to replicate the EDGES results.

The main effect of EDGES observations on the expected correlations is to boost the signal by nearly an order of magnitude in the redshift range $15 < z < 19$, even as compared to the most optimistic models¹⁰ (low N_{heat} and high f_L) in the usual case.¹¹ This is entirely owing to a decrease in T_S that boosts s in Equation (3). The correlation function scales roughly as $(1 - s)^2$. Even though this result has been arrived at within the framework of a model involving baryon cooling, this result holds when excess radio background is responsible for the deep absorption feature because s increases by a similar factor in that case too.

This result is also robust to change in other modeling parameters f_L and N_{heat} . Therefore, an important prediction of the EDGES detection is that the deeper absorption trough in the global signal gives a corresponding increase in the fluctuating component too. Even when the spin temperature field is very uniform, with Ly α -coupled and unheated gas, the low spin temperature would enhance the underlying density inhomogeneities (Equation (17)). As we have shown in Section 4.2, the spin temperature field can be negatively correlated at some scales. In such scenarios, the fluctuating component of the signal could be less than given by Equation (17). We tried to produce such models for parameters needed to explain the EDGES data and found it very difficult to anticorrelate the spin temperature field. Hence we infer that the minimum value of the correlation function that corresponds to the EDGES signal is given by Equation (17), using the temperature derived from the trough in the global signal. If the global signal has a trough of $\sim 500 \text{ mK}$ at $z \simeq 17$, then assuming complete coupling and no heating (no fluctuations due to these two fields), the autocorrelation function at $r = 2 \text{ Mpc}$ and $r = 4 \text{ Mpc}$ should be $\sim 5100 \text{ (mK)}^2$ and $\sim 2500 \text{ (mK)}^2$, respectively.

Many currently operational (e.g., LOFAR, MWA, PAPER, GMRT) and upcoming radio interferometers (HERA, SKA) have the capability to detect the fluctuating component of the HI signal in the redshift range $8 < z < 25$ (for details, see, e.g., Mesinger et al. 2014; Ahn et al. 2015; Koopmans et al. 2015). We discuss the detectability of this signal especially in light of the recent EDGES results. These radio interferometers directly measure visibilities and their correlations, which can be related to the power spectrum of the HI signal (e.g., Bharadwaj & Sethi 2001; Zaldarriaga et al. 2004). This data analysis can readily be extended to the image plane (which is a byproduct of the analysis pipeline; e.g., Patil et al. 2017 for LOFAR). This means real-space correlation functions can also be used for computation of the signal (e.g., Sethi & Haiman 2008). As an

approximate rule, one could use $r \simeq \pi/k$ to shift from Fourier to real space.

SKA1-LOW is expected to detect the HI signal at $z \simeq 16$ with a signal-to-noise ratio varying from 100 to 10 for $0.1 < k < 0.6 \text{ Mpc}^{-1}$ for a signal strength $\Delta^2(k) \simeq 10^2 \text{ (mK)}^2$ (for details, see, e.g., Koopmans et al. 2015). EDGES results predict a signal strength nearly a factor of five larger, which means the signal-to-noise ratio would be significantly higher (Figure 8). The ongoing experiment LOFAR's best upper limit corresponds to $(80 \text{ mK})^2$ ($k \simeq 0.05 \text{ Mpc}^{-1}$) in the redshift range $9.6 < z < 10.6$ in 10 hr of integration.¹² LOFAR has the frequency range to probe the redshift range of EDGES detection. If the noise properties at smaller frequencies ($\simeq 80 \text{ MHz}$) behave roughly as the one observed at higher frequencies ($\simeq 110 \text{ MHz}$), LOFAR might be able to detect the signal in a few hundred hours of integration.

SKA1-LOW (and SKA2) have the capability to detect the EoR signal in the redshift range $20 < z < 25$. A signal of 100 (mK)^2 at $z \simeq 25$ is detectable by a deep SKA1-LOW survey with signal-to-noise ratio of five for $k < 0.1 \text{ Mpc}^{-1}$ (Koopmans et al. 2015). However, we notice that the EDGES result places a significant constraint on the signal in this era. This, as noted above, is because the EDGES results imply a smaller value of f_L , which results in a smaller signal for $z > 19$ as compared to models with a larger f_L , which give a significantly higher signal (Figure 8).

5. Summary and Conclusions

The main aim of this paper is to extend the analytic formalism presented in our previous work (Raste & Sethi 2018) to explore the epoch of partial Ly α coupling. Following RS18, we generate the size distribution of self-ionized regions using excursion set formalism for the Λ CDM model. Around these bubbles, we create spin temperature profiles due to X-ray heating and Ly α coupling, which merge smoothly with a uniform background. We model these spin temperature T_S profiles using two parameters: N_{heat} , the number of X-ray photons per stellar baryon, and f_L , the ratio of Ly α to ionizing photons. Our analytic formulation allows us to explore relevant physical processes and study both their individual and combined effect on the HI signal.

We study the evolution of correlation properties of HI in the redshift range $10 \leq z \leq 30$ for many possible scenarios, with a greater focus on higher redshifts at which partial Ly α coupling plays an important role (Figures 4, 9, 7, 8). We find a reasonable agreement with existing semianalytic and N -body simulation results. As we compute correlation functions in both real and Fourier space, we find a possible case where the correlation function in real space is negative owing to partial heating and Ly α coupling (Figure 4).

We also analyze the implications of the recent EDGES detection of the global HI signal in the redshift range $15 < z < 19$. Generically, EDGES detection results in a higher correlation signal in the redshift range of the detection but a lower signal at higher redshifts, as compared to the most

¹⁰ We have not incorporated the enhancement in the signal due to the inhomogeneous velocity-dependent cooling of gas within this model (Fialkov et al. 2018; Muñoz et al. 2018).

¹¹ We notice a decrease in the signal in this case at $z \simeq 30$. This model-dependent decrement is due to cooler baryons causing a decrease in the efficiency of coupling from collisions.

¹² The angular scale above which the HI signal can be reliably measured for most ongoing and upcoming radio interferometers is a few arcminutes; $1'$ corresponds to nearly 3 Mpc (comoving) at $z \simeq 15$, or these telescopes are sensitive to linear scales larger than 5–10 Mpc (comoving). However, these telescopes have frequency resolutions that correspond to much smaller linear scales; for example, MWA's frequency resolution of 40 kHz corresponds to nearly 1 Mpc (comoving) along the line of sight. Or the 3D HI signal is probed with a different resolution on the sky plane as compared to the line of sight.

optimistic models, which do not take into account this detection (Figures 7 and 8).

In this paper, we study the implications of the standard Λ CDM model. The resultant size distribution of self-ionized regions and hence the fluctuation scales of brightness temperature would be different for an extension of this model, which we hope to explore in future work (e.g., Sethi & Subramanian 2009; Sarkar et al. 2016).

Understanding theoretically the HI signal from EoR/cosmic dawn remains a challenge. Given the large amount of uncertainty

$$P(A \cap B) = P(A) - P(A \cap \bar{B}). \quad (28)$$

Appendix B Geometry

If point 1 is located between distance P and Q ($P < Q$) from the center of a sphere, then $C(x, P, Q, R)$ is the probability that its neighbor point 2 at distance x from point 1 is located outside the concentric sphere of radius R :

$$C(x, P, Q, R) = \begin{cases} 0 & x \leq R - Q \\ 1 & x \leq P - R \\ 1 & x \geq R + Q \\ \frac{1}{2} \frac{Q^3 - (R-x)^3}{Q^3 - P^3} + \frac{3}{8x} \frac{Q^2 - (R-x)^2}{Q^3 - P^3} \left[\frac{Q^2 + (R-x)^2}{2} + (x^2 - R^2) \right] & R - Q \leq x \leq R - P, x > Q - R \\ \frac{Q^3 - R^3}{Q^3 - P^3} & R - Q \leq x \leq R - P, x \leq Q - R \\ \frac{1}{2} \frac{2Q^3 - (R+x)^3 - P^3}{Q^3 - P^3} + \frac{3}{8x} \frac{(R+x)^2 - P^2}{Q^3 - P^3} \left[\frac{(R+x)^2 + P^2}{2} + (x^2 - R^2) \right] & |P - R| \leq x \leq Q - R, x < P + R \\ 1 - \frac{R^3}{Q^3 - P^3} & P + R \leq x \leq Q - R, \\ \frac{1}{2} \frac{(x-R)^3 + Q^3 - 2P^3}{Q^3 - P^3} + \frac{3}{8x} \frac{Q^2 - (x-R)^2}{Q^3 - P^3} \left[\frac{Q^2 + (x-R)^2}{2} + (x^2 - R^2) \right] & R + P \leq x \leq R + Q, x > Q - R \\ \frac{1}{2} + \frac{3}{8x} \frac{P+Q}{P^2 + PQ + Q^2} \left[\frac{P^2 + Q^2}{2} + (x^2 - R^2) \right] & R - P \leq x \leq P + R, x \geq Q - R. \end{cases} \quad (29)$$

in the physics of ionizing sources, IGM, feedback mechanisms, and so on, it is important to explore a wide set of modeling paradigms. While N -body simulations are important to understand and image the HI field, analytic methods, like the one presented in this paper, are suited to predicting the statistical quantities like correlation function and power spectrum. Since our formalism is not limited by the size of the simulation box, we can easily incorporate a variety of physical processes at very small or very large scale, for example, the influence region of $\text{Ly}\alpha$ photons. Also, this formalism is computationally cheaper, which means we can explore a large set of modeling parameters and their degeneracies at a fraction of the computation resources taken by N -body simulation. Since N -body simulations, semi-analytic, and analytic formalisms each have their own set of assumptions, strengths, and weaknesses, it is beneficial to apply all these methods to unravel the complex physics of reionization.

The authors would like to thank Akash Kumar Patwa for valuable discussions and insights.

Appendix A Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (26)$$

$$P((A \cap B)|C) = P(A|(B \cap C)) P(B|C) \quad (27)$$

Three-dimensional Fourier transform of $1 - C(x, P, Q, R)$ is useful while calculating the power spectrum:

$$\begin{aligned} \text{FT}(1 - C(x, P, Q, R)) &= \frac{12\pi}{k^6(Q^3 - P^3)} \\ &([kQ \text{Cos}(kQ) - \text{Sin}(kQ)] - [kP \text{Cos}(kP) - \text{Sin}(kP)]) \\ &\times [kR \text{Cos}(kR) - \text{Sin}(kR)]. \end{aligned}$$

Appendix C Overlap

Assume that there is a sphere with two shells: R_i is the radius of the inner shell (actually a sphere), and R_o is the outer shell radius. We wish to randomly put N such spheres in a large box of volume V ($V \gg 4\pi/3 R_o^3$), in such a way that the inner shells of any two spheres must not overlap with each other, but they can overlap with the outer shell of another sphere. Outer shells of two or more spheres can overlap with one another. We wish to calculate the total volume fraction occupied by the outer shells of these N spheres.

When $N = 1$, there is only one sphere randomly placed within the box. The fraction of volume of this box occupied by the inner and outer shells is, respectively,

$$g_i = \frac{4\pi R_i^3}{3V} \quad \text{and} \quad g_o = \frac{4\pi R_o^3 - R_i^3}{3V}$$

with the total volume fraction occupied by the sphere $g_i = g_i + g_o$. Here we are only interested in the volume fraction occupied by the outer shells. Thus, we call this fraction (for $N=1$) $g_1 = g_o$. For $N=2$, we need to randomly place the second sphere after the first sphere is already placed in the box. Across multiple experiments, where the second sphere is placed randomly, the average total volume fraction that is occupied by the outer shell of these two spheres is

$$g_2 = g_1 + g_1(1 - g_1) = g_1 + g_1(1 - (g_1 + g_1)),$$

where the first term g_1 is due to the first sphere, and the second term is due to the second sphere, which has $(1 - g_1)$ probability of overlap with the first sphere. Placing a third sphere in this scenario ($N=3$), we have

$$g_3 = g_2 + g_1(1 - (2g_1 + g_2)),$$

where the first term g_2 is due to the first two spheres, and the second term is due to the third sphere, which can overlap with these two spheres. This can be expressed recursively for N spheres as

$$\begin{aligned} g_N &= g_{N-1} + g_1(1 - ((N-1)g_1 + g_{N-1})) \\ &= g_{N-1}(1 - g_1) + g_1(1 - (N-1)g_1) \\ &= \sum_{k=0}^{N-1} g_1(1 - g_1)^k(1 - (N-1-k)g_1) \\ &= \left(1 + \frac{g_1}{g_1}\right)(1 - (1 - g_1)^N) - Ng_1. \end{aligned} \quad (30)$$

Here we can define $f_i = Ng_i$, the total volume fraction occupied by the inner shells of all spheres (since they do not overlap, it is simple addition), and $f_{ob} = Ng_o$, the total sum of the fractions occupied by the outer shells independently, including the multiple counting of overlapped parts. The actual volume fraction occupied by the outer shells is $f_o = g_N$. In the limit where $g_1 \ll 1$ and $N \gg 1$, we have

$$1 - (1 - g_1)^N \simeq 1 - \exp(-Ng_1).$$

Thus, Equation (30) can be written as

$$f_o = \left(1 + \frac{f_i}{f_{ob}}\right)(1 - \exp(-f_{ob})) - f_i. \quad (31)$$

The results match very well with the simulation where a large number of spheres are randomly arranged within a box with the above conditions.

Appendix D

Complete Model ($\mu = \langle \psi_1 \psi_2 \rangle - \langle \psi \rangle^2$)

In this section, we develop a formalism to calculate the correlation of dimensionless brightness temperature ψ for epochs at which the ionization volume fraction is small. This ensures that ionization bubbles are separate and nonoverlapping. However, as described in Section 3.2 and Appendix C, our formulation allows us to deal with overlap of heating bubbles.

Since we have already assumed that the cross-correlation of density with ionization or spin temperature is negligible, we try to find the autocorrelation of $\phi = n(1 - s)$ (henceforth referred to as ‘‘temperature’’ in this section; Equation (15)). Therefore, to calculate the correlation at scale r , we need to find pairs of points such that $\phi_1 \neq 0$ and $\phi_2 \neq 0$, where ϕ_1 and ϕ_2 are temperatures of points 1 and 2 that are separated by distance r :

$$\begin{aligned} \langle \phi_1 \phi_2 \rangle &= \phi_b^2 P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi_b\}) \\ &+ 2\phi_b \sum_{R_x} \sum_{\phi(R_x)} \phi P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi\}) \\ &+ \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p \phi_q P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\}). \end{aligned} \quad (32)$$

Here, ϕ_i and ϕ_b are the temperatures of the ionized region and background regions, respectively; $P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\})$ is the joint probability of point 1 having temperature ϕ_p and point 2 having temperature ϕ_q . In the above equation, the first, second, and third terms correspond respectively to (1) both points being in the background region, (2) one point being in the background and another in a heated bubble, and (3) both points being in some heated bubble. If one or both points are in the ionized region, the term corresponding to that pair will be zero. When both points are in the background, we use Equation (28):

$$\begin{aligned} P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi_b\}) &= P(\phi_1 = \phi_b) - P(\{\phi_1 = \phi_b\} \cap \{\phi_2 \neq \phi_b\}) \\ &= P(\phi_1 = \phi_b) - P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi_i\}) \\ &- \sum_{R_x} \sum_{\phi(R_x)} P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi\}). \end{aligned} \quad (33)$$

In the case where both points are partially heated, these points can be within the same bubble or different bubbles. Within the same bubble, they can be in the same shell ($\phi_1 = \phi_2$) or in different shells. This gives

$$\begin{aligned} &\sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p \phi_q P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\}) \\ &= \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_p\} \cap \{\text{same}\}) \\ &+ \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{\phi_q(R_x) \neq \phi_p} \phi_p \phi_q P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{same}\}) \\ &+ \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p \phi_q P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{diff}\}). \end{aligned} \quad (34)$$

Here, $P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_p\} \cap \{\text{same}\})$ is the probability that both points are in the same bubble with the same temperature ϕ_p . This is straightforward to calculate. However, because we allow overlap of heated profiles, the simplest derivation is not necessarily correct. Instead, we expand it

further:

$$\begin{aligned}
 & \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_p\} \cap \{\text{same}\}) \\
 &= \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 (P(\phi_1 = \phi_p) - P(\{\phi_1 = \phi_p\} \cap \{\phi_2 \neq \phi_p\})) \\
 &= \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 (P(\phi_1 = \phi_p) - P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_b\}) \\
 &\quad - P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_i\})) \\
 &\quad - \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{\phi_q(R_x) \neq \phi_p} \phi_p^2 P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{same}\}) \\
 &\quad - \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p^2 P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{diff}\}).
 \end{aligned} \tag{35}$$

Putting Equations (33)–(35) in (32), we get

$$\begin{aligned}
 \langle \phi_1 \phi_2 \rangle &= \phi_b^2 P(\phi_1 = \phi_b) - \phi_b^2 P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi_i\}) \\
 &+ \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 P(\phi_1 = \phi_p) - \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 \\
 &\quad \times P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_i\}) \\
 &- \sum_{R_x} \sum_{\phi(R_x)} (\phi_b - \phi)^2 P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi\}) \\
 &+ \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{\phi_q(R_x)} \phi_p (\phi_q - \phi_p) P(\{\phi_1 = \phi_p\} \\
 &\quad \cap \{\phi_2 = \phi_q\} \cap \{\text{same}\}) \\
 &+ \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p (\phi_q - \phi_p) P(\{\phi_1 = \phi_p\} \\
 &\quad \cap \{\phi_2 = \phi_q\} \cap \{\text{diff}\}).
 \end{aligned} \tag{36}$$

Now we calculate each term separately. Here, $P(\phi_1 = \phi_b) = f_b$ is the probability of a point being in the background region, and $P(\phi_1 = \phi_p) = 4\pi/3 N(R_x) f_b ((R_{sp} + \Delta R_{sp})^3 - R_{sp}^3)$ is the probability of a point having temperature ϕ_p in a bubble of ionization radius R_x .

The probability of one point being in the background and a second point being ionized can be calculated by assuming that point 1 is in an ionized region of radius R_x , and calculating the probability that its neighbor point 2 at distance r is outside the temperature profile of that bubble ($C(r, 0, R_x, R_h)$). Given this condition, the probability of point 2 being in the background region is equal to the average background fraction f_b . Thus, we have

$$\begin{aligned}
 & P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi_i\}) \\
 &= f_b \sum_{R_x} N(R_x) \frac{4\pi}{3} R_x^3 C(r, 0, R_x, R_h).
 \end{aligned}$$

When point 1 has temperature ϕ_p and point 2 is ionized, they both can be in the same bubble or in different bubbles, which

respectively give the first and second terms on the right-hand side of the following equation:

$$\begin{aligned}
 & \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_i\}) \\
 &= \sum_{R_x} \sum_{\phi_p(R_x)} \phi_p^2 \frac{4\pi}{3} N(R_x) \left(R_x^3 (1 - f_i) \right. \\
 &\quad [C(r, 0, R_x, R_{sp}) - C(r, 0, R_x, R_{sp} + \Delta R_{sp})] \\
 &\quad \left. + f_b ((R_{sp} + \Delta R_{sp})^3 - R_{sp}^3) \right) \\
 &\quad \times \sum_{R'_x} \frac{4\pi}{3} N(R'_x) R_x'^3 C(r, 0, R'_x, R_h').
 \end{aligned} \tag{37}$$

When point 1 is in the background and point 2 has temperature ϕ , we can expand it as

$$\begin{aligned}
 & P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi\}) \\
 &= P(\phi_1(\text{out other}) | (\phi_1(\text{out same}) \cap \{\phi_2 = \phi\})) \\
 &\quad \times P(\phi_1(\text{out same}) \cap \{\phi_2 = \phi\}),
 \end{aligned}$$

where the first term on the right-hand side, $P(\phi_1(\text{out other}) | (\phi_1(\text{out same}) \cap \{\phi_2 = \phi\}))$, is the probability that point 1 is in the background region given that point 2 is partially heated with temperature ϕ and point 1 is not inside the same bubble in which point 2 is. This probability equals the fraction of the universe heated to background temperature, f_b . The second term on the right-hand side, $P(\phi_1(\text{out same}) \cap \{\phi_2 = \phi\})$, is the probability that point 1 is out of the bubble in which point 2 is, and point 2 is partially heated with temperature ϕ . As point 2 can be in a bubble with any ionization radius R_x , $P(\phi_1(\text{out same}) | \{\phi_2 = \phi\}(R_x))$ is the probability that point 1 is outside of the bubble that has ionization radius R_x and that contains point 2 with temperature ϕ . This equals the probability that point 1 is outside of the bubble with outer radius R_h in which point 2 is located between radius R_s and $R_s + \Delta R_s$.

$$\begin{aligned}
 & \sum_{R_x} \sum_{\phi(R_x)} P(\{\phi_1 = \phi_b\} \cap \{\phi_2 = \phi\}) \\
 &= f_b \sum_{R_x} N(R_x) \frac{4\pi}{3} \sum_{\phi(R_x)} f_b ((R_s + \Delta R_s)^3 - R_s^3) C \\
 &\quad \times (r, R_s, R_s + \Delta R_s, R_h).
 \end{aligned} \tag{38}$$

Note that $P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{diff}\})$ gives the probability that point 1 has $\phi = \phi_p$, point 2 has $\phi = \phi_q$, and they both belong to different bubbles. We take a simple assumption that if point 2 is outside the bubble in which point 1 is, then its probability of having $\phi = \phi_q$ is equal to the global probability of ϕ_q temperature shell. Here, either p or q could

have been point 1. Therefore,

$$\begin{aligned}
& \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{R'_x} \sum_{\phi_q(R'_x)} \phi_p(\phi_q - \phi_p) \\
& P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{diff}\}) \\
& = \sum_{R_x} N(R_x) \frac{4\pi}{3} \sum_{\phi_p(R_x)} f_b((R_{s_p} + \Delta R_{s_p})^3 - R_{s_p}^3) \\
& \frac{\sum_{R'_x} N(R'_x) \frac{4\pi}{3} \sum_{\phi_q(R'_x)} f_b((R'_{s_q} + \Delta R'_{s_q})^3 - R'_{s_q}{}^3) \phi_p(\phi_q - \phi_p)}{2} \\
& \frac{C(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_h) + C(r, R'_{s_q}, R'_{s_q} + \Delta R'_{s_q}, R'_h)}{2}.
\end{aligned} \tag{39}$$

Here, $P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{same}\})$ gives the probability that point 1 has $\phi = \phi_p$, point 2 has $\phi = \phi_q \neq \phi_p$, and they both belong to the same bubble. If point 1 is located at a distance between R_{s_p} and $R_{s_p} + \Delta R_{s_p}$ from the center of the sphere, then the fraction of its neighbors at distance r that are outside the sphere of radius R_{s_q} and inside the sphere of radius $R_{s_q} + \Delta R_{s_q}$ can be computed. However, because bubbles can overlap, point 2 can be neutral or ionized, which leads to

$$\begin{aligned}
& \sum_{R_x} \sum_{\phi_p(R_x)} \sum_{\phi_q(R_x) \neq \phi_p} \phi_p \phi_q P(\{\phi_1 = \phi_p\} \cap \{\phi_2 = \phi_q\} \cap \{\text{same}\}) \\
& = \sum_{R_x} N(R_x) \frac{4\pi}{3} \sum_{\phi_p(R_x)} f_b((R_{s_p} + \Delta R_{s_p})^3 - R_{s_p}^3) \\
& \sum_{\phi_q(R_x) \neq \phi_p} \phi_p \phi_q (1 - f_i) (C(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_{s_q}) \\
& - C(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_{s_q} + \Delta R_{s_q})).
\end{aligned} \tag{40}$$

Putting all equations together, including the influence of ξ and simplifying in terms of $F(x, P, Q, R) = 1 - C(x, P, Q, R)$, we get

$$\begin{aligned}
\mu & = \xi(f_b \phi_b + \langle \phi_h \rangle)^2 + (1 + \xi) \\
& \times \left(\left(\phi_b^2 f_b + \langle \phi_h^2 \rangle \right) \sum_{R_x} \frac{4\pi}{3} N(R_x) R_x^3 F(r, 0, R_x, R_h) \right) \\
& + f_b \sum_{R_x} \frac{4\pi}{3} N(R_x) \sum_{\phi(R_x)} (\phi_b - \phi)^2 f_b \\
& \times ((R_s + \Delta R_s)^3 - R_s^3) F(r, R_s, R_s + \Delta R_s, R_h) \\
& + \sum_{R_x} \frac{4\pi}{3} N(R_x) R_x^3 \sum_{\phi(R_x)} \phi^2 (1 - f_i) \\
& \times [F(r, 0, R_x, R_s) - F(r, 0, R_x, R_s + \Delta R_s)] \\
& - \sum_{R_x} \frac{4\pi}{3} N(R_x) \sum_{\phi_p(R_x)} f_b((R_{s_p} + \Delta R_{s_p})^3 - R_{s_p}^3) \\
& \times \sum_{\phi_q(R_x)} \phi_p(\phi_q - \phi_p) (1 - f_i)
\end{aligned} \tag{41}$$

$$\begin{aligned}
& [F(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_{s_q}) \\
& - F(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_{s_q} + \Delta R_{s_q})] \\
& - \sum_{R_x} \frac{4\pi}{3} N(R_x) \sum_{\phi_p(R_x)} f_b((R_{s_p} + \Delta R_{s_p})^3 - R_{s_p}^3) \\
& \times \sum_{R'_x} \frac{4\pi}{3} N(R'_x) \sum_{\phi_q(R'_x)} f_b((R'_{s_q} + \Delta R'_{s_q})^3 - R'_{s_q}{}^3) \\
& \phi_p(\phi_q - \phi_p) \frac{1}{2} [F(r, R_{s_p}, R_{s_p} + \Delta R_{s_p}, R_h) \\
& + F(r, R'_{s_q}, R'_{s_q} + \Delta R'_{s_q}, R'_h)].
\end{aligned} \tag{41}$$

where

$$\begin{aligned}
\langle \phi_h \rangle & = \sum_{R_x} \frac{4\pi}{3} N(R_x) \sum_{\phi(R_x)} \phi f_b((R_s + \Delta R_s)^3 - R_s^3) \\
\langle \phi_h^2 \rangle & = \sum_{R_x} \frac{4\pi}{3} N(R_x) \sum_{\phi(R_x)} \phi^2 f_b((R_s + \Delta R_s)^3 - R_s^3).
\end{aligned} \tag{42}$$

We also calculate the correlation at the same point using $\langle \phi_1 \phi_1 \rangle = \langle \phi^2 \rangle$,

$$\mu_0 = (1 + \xi_0)(\phi_b^2 f_b + \langle \phi_h^2 \rangle) - (f_b \phi_b + \langle \phi_h \rangle)^2, \tag{43}$$

where $\xi_0 = \xi(r = 0)$.

Equation (41) matches the expression derived in RS18 for most of the cases. However, this expression is more robust as it gives expected results in all simplifying cases and boundary conditions.

ORCID iDs

Janakee Raste  <https://orcid.org/0000-0001-7451-6139>

References

- Ahn, K., Mesinger, A., Alvarez, M. A., & Chen, X. 2015, in Proc. Advancing Astrophysics with the Square Kilometre Array Conf., ed. T. L. Bourke et al. (Trieste: PoS), 3
- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, *ApJ*, **809**, 61
- Baek, S., Semelin, B., Di Matteo, P., Revaz, Y., & Combes, F. 2010, *A&A*, **523**, A4
- Barkana, R. 2018, *Natur*, **555**, 71
- Barkana, R., & Loeb, A. 2001, *PhR*, **349**, 125
- Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, *ApJ*, **833**, 102
- Becker, R. H., Fan, X., White, R. L., et al. 2001, *AJ*, **122**, 2850
- Bernardi, G., McQuinn, M., & Greenhill, L. J. 2015, *ApJ*, **799**, 90
- Bharadwaj, S., & Sethi, S. K. 2001, *JApA*, **22**, 293
- Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. 2018, *Natur*, **555**, 67
- Chen, X., & Miralda-Escudé, J. 2004, *ApJ*, **602**, 1
- Chen, X., & Miralda-Escudé, J. 2008, *ApJ*, **684**, 18
- Dayal, P., & Ferrara, A. 2018, *PhR*, **780**, 1
- Draine, B. T., & Miralda-Escudé, J. 2018, *ApJL*, **858**, L10
- Ewall-Wice, A., Chang, T.-C., Lazio, J., et al. 2018, *ApJ*, **868**, 63
- Fan, X., White, R. L., Davis, M., et al. 2000, *AJ*, **120**, 1167
- Feng, C., & Holder, G. 2018, *ApJL*, **858**, L17
- Fialkov, A., Barkana, R., & Cohen, A. 2015, *PhRvL*, **114**, 101303
- Fialkov, A., Barkana, R., & Cohen, A. 2018, *PhRvL*, **121**, 011101
- Fialkov, A., Barkana, R., & Visbal, E. 2014, *Natur*, **506**, 197
- Fialkov, A., Cohen, A., Barkana, R., & Silk, J. 2017, *MNRAS*, **464**, 3498
- Field, G. B. 1958, *PIRE*, **46**, 240
- Field, G. B. 1959, *ApJ*, **129**, 551

- Fraser, S., Hektor, A., Hütsi, G., et al. 2018, *PhLB*, **785**, 159
- Furlanetto, S. R., & Oh, S. P. 2016, *MNRAS*, **457**, 1813
- Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004a, *ApJ*, **613**, 16
- Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004b, *ApJ*, **613**, 1
- Ghara, R., Choudhury, T. R., & Datta, K. K. 2015, *MNRAS*, **447**, 1806
- Giri, S. K., Mellema, G., Dixon, K. L., & Iliev, I. T. 2018, *MNRAS*, **473**, 2949
- Gnedin, N. Y., & Shaver, P. A. 2004, *ApJ*, **608**, 611
- Hills, R., Kulkarni, G., Meerburg, P. D., & Puchwein, E. 2018, arXiv:1805.01421
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hirata, C. M. 2006, *MNRAS*, **367**, 259
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in Proc. Advancing Astrophysics with the Square Kilometre Array Conf., ed. T. L. Bourke et al. (Trieste: PoS), **1**
- Lambiase, G., & Mohanty, S. 2018, arXiv:1804.05318
- Madau, P., Meiksin, A., & Rees, M. J. 1997, *ApJ*, **475**, 429
- Mesinger, A., Ewall-Wice, A., & Hewitt, J. 2014, *MNRAS*, **439**, 3262
- Mesinger, A., Ferrara, A., & Spiegel, D. S. 2013, *MNRAS*, **431**, 621
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, **411**, 955
- Mesinger, A., Greig, B., & Sobacchi, E. 2016, *MNRAS*, **459**, 2342
- Morales, M. F., & Wyithe, J. S. B. 2010, *ARA&A*, **48**, 127
- Muñoz, J. B., Dvorkin, C., & Loeb, A. 2018, *PhRvL*, **121**, 121301
- Muñoz, J. B., & Loeb, A. 2018, *Natur*, **557**, 684
- Natarajan, A., & Yoshida, N. 2014, *PTEP*, **2014**, 06B112
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, **433**, 639
- Pacucci, F., Mesinger, A., Mineo, S., & Ferrara, A. 2014, *MNRAS*, **443**, 678
- Patil, A. H., Yatawatta, S., Koopmans, L. V. E., et al. 2017, *ApJ*, **838**, 65
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, *A&A*, **571**, A16
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209
- Price, D. C., Greenhill, L. J., Fialkov, A., et al. 2018, *MNRAS*, **478**, 4193
- Pritchard, J. R., & Furlanetto, S. R. 2007, *MNRAS*, **376**, 1680
- Pritchard, J. R., & Loeb, A. 2012, *RPPh*, **75**, 086901
- Raste, J., & Sethi, S. 2018, *ApJ*, **860**, 55
- Ross, H. E., Dixon, K. L., Iliev, I. T., & Mellema, G. 2017, *MNRAS*, **468**, 3785
- Rybicki, G. B., & dell'Antonio, I. P. 1994, *ApJ*, **427**, 603
- Santos, M. G., Amblard, A., Pritchard, J., et al. 2008, *ApJ*, **689**, 1
- Santos, M. G., Ferramacho, L., Silva, M. B., Amblard, A., & Cooray, A. 2010, *MNRAS*, **406**, 2421
- Sarkar, A., Mondal, R., Das, S., et al. 2016, *JCAP*, **4**, 012
- Sethi, S., & Haiman, Z. 2008, *ApJ*, **673**, 1
- Sethi, S. K. 2005, *MNRAS*, **363**, 818
- Sethi, S. K., & Subramanian, K. 2009, *JCAP*, **11**, 021
- Sharma, P. 2018, *MNRAS*, **481**, L6
- Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, *A&A*, **345**, 380
- Shull, J. M., & van Steenberg, M. E. 1985, *ApJ*, **298**, 268
- Singh, S., Subrahmanyam, R., Shankar, N. U., et al. 2018a, *ExA*, **45**, 269
- Singh, S., Subrahmanyam, R., Udaya Shankar, N., et al. 2018b, *ApJ*, **858**, 54
- Sokolowski, M., Tremblay, S. E., Wayth, R. B., et al. 2015, *PASA*, **32**, 4
- Tashiro, H., & Sugiyama, N. 2013, *MNRAS*, **435**, 3001
- Tozzi, P., Madau, P., Meiksin, A., & Rees, M. J. 2000, *NuPhS*, **80**, 05
- Venkatesan, A., Giroux, M. L., & Shull, J. M. 2001, *ApJ*, **563**, 1
- Visbal, E., Barkana, R., Fialkov, A., Tselikhovich, D., & Hirata, C. M. 2012, *Natur*, **487**, 70
- Voytek, T. C., Natarajan, A., Jáuregui García, J. M., Peterson, J. B., & López-Cruz, O. 2014, *ApJL*, **782**, L9
- Zahn, O., Mesinger, A., McQuinn, M., et al. 2011, *MNRAS*, **414**, 727
- Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, *ApJ*, **608**, 622
- Zygelman, B. 2005, *ApJ*, **622**, 1356