Ultimate frontier of 21-cm cosmology

Patrick C. Breysse, Yacine Ali-Haïmoud, and Christopher M. Hirata

The ultimate frontier of 21cm cosmology

Patrick C. Breysse,[1, 2, *] Yacine Ali-Haïmoud,[3] and Christopher M. Hirata[4]

[1]*Canadian Institute for Theoretical Astrophysics,*
60 St George St., Toronto, ON, M5S 3H8, Canada
[2]*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218 USA*
[3]*Center for Cosmology and Particle Physics, Department of Physics,*
New York University, NY 10003, New York, USA
[4]*Center for Cosmology and AstroParticle Physics,*
Department of Physics, The Ohio State University,
191 W Woodruff Ave, Columbus OH 43210, USA

We present the most detailed computation to date of the 21-cm global signal and fluctuations at $z \gtrsim 500$. Our calculations include a highly precise estimate of the Wouthuysen-Field (WF) effect and the first explicit calculation of the impact of free-free processes, the two dominant components of the signal at $z \gtrsim 800$. We implement a new high-resolution Lyα radiative transfer calculation, coupled to a state-of-the-art primordial recombination code. Using these tools, we find a global signal from 21-cm processes alone of roughly 0.01mK at $z \sim 1000$, slightly larger than it would be without the WF effect, but much weaker than previous estimates including this effect. We also find that this signal is swamped by a smooth $1-2$ mK signal due to free-free absorption at high redshift by the partially ionized gas along the line of sight. In addition, we estimate the amplitude of 21-cm fluctuations, of order $\sim 10^{-7}$ mK at $z \sim 1000$. Unfortunately, we find that due to the brightness of the low-frequency sky, these fluctuations will not be observable beyond $z \sim$ a few hundred by even extremely futuristic observations. The 21 cm fluctuations are exponentially suppressed at higher redshifts by the large free-free optical depth, making this the ultimate upper redshift limit for 21-cm surveys.

I. INTRODUCTION

In the most distant reaches of the Universe, between the onset of cosmological recombination and the end of reionization, hydrogen atoms provide one of the only available tracers of large-scale structure. In particular, the 21-cm hyperfine transition in neutral hydrogen has become recognized as a powerful probe of these so-called "dark ages" [1–4]. This transition is optically thin, so it is in principle possible to map the structure of the high-redshift Universe in three dimensions by measuring angular and spectral fluctuations in brightness.

Most theoretical and experimental efforts to date have been focused on the 21-cm signal from the epoch of reionization, the last large-scale phase transition in the Universe where the intergalactic medium was ionized by the first galaxies [5]. Radio instruments such as the LOw Frequency ARay (LOFAR, [6]), the Precision Array for Probing the Epoch of Reionization (PAPER, [7]), and the Murchison Wide-field Array (MWA, [8]) have begun placing upper limits on the EoR signal in preparation for planned experiments like the under-construction Hydrogen Epoch of Reionization Array (HERA, [9]), the Square Kilometer Array (SKA, [10]), and the Dark Ages Radio Explorer (DARE, [11]). Other projects, including the Canadian Hydrogen Intensity Mapping Experiment (CHIME, [12]) and the Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX, [13]), will target the re-

maining neutral hydrogen within galaxies after reionization.

While the detection of the 21-cm signal from reionization seems imminent, it depends sensitively on complex astrophysical processes, and its interpretation will be a theoretical challenge [14]. It is also in principle possible to go after the 21-cm signal from beyond reionization [15, 16]. Though atmospheric and astrophysical foreground emission pose a significant technical challenge to observations at these redshifts [17, 18], the signal is expected to be much cleaner of astrophysical uncertainties, in particular at $z \gtrsim 20 - 30$, before the first stars have formed [19]. High-redshift 21-cm tomography promises to become the ultimate cosmological probe by opening a window into the earliest periods of structure formation. To cite only a few applications, it can shed light on the character of initial conditions through the measurement of small-scale perturbations [15, 20], statistical isotropy [21] and primordial non-gaussianity [22–24], and can probe the nature of dark matter [25–28].

The 21-cm signal is the contrast between the brightness temperature in the redshifted 21-cm line and the CMB temperature. This contrast is itself proportional to the difference between the spin temperature, characterizing the relative abundance of ground-state hydrogen in its two hyperfine substates, and the CMB temperature. This is set by three competing processes: absorption of and stimulated emissions by resonant CMB photons, collisions between hydrogen atoms, and spin-flipping Lyman-α scatterings. The latter process, the Wouthuysen-Field (WF) effect [29, 30], indirectly drives the spin temperature towards the gas temperature, as the Lyman-α color

temperature near line center is expected to be very near equilibrium with the kinetic temperature of the gas. The standard lore is that the WF effect is only relevant at low redshifts, once the first galaxies have produced a significant Lyman-α background. Recently, however, Ref. [31] (hereafter FL13) pointed out that the WF effect actually dominates over other processes at $z \gtrsim 850$, due to the large Lyman-α background generated during the out-of-equilibrium cosmological recombination process [32, 33]. More interestingly, FL13 found a significant departure of the Lyman-α color temperature from the gas temperature at $z \gtrsim 500$; as a result, they found a much larger 21-cm signal at these redshifts than previously estimated, though the precise value depends on the recombination code they used [34, 35].

FL13 used the standard output of precision recombination codes (CosmoRec [35] and HyRec [34, 36]) to compute the photon spectrum near the Lyman-α line. However, the accuracy requirement for the recombination problem and high-redshift 21-cm calculation of interest here are vastly different. An accurate net recombination rate is sensitive to the details of Lyman-α radiative transfer mostly at $z \gtrsim 900$, for which decays from $n = 2$ to the ground state are the bottleneck of the recombination process (see e.g. [37]). At these redshifts, the Lyman-alpha line is optically thick to absorption far into the damping wings [38, 39]. As a consequence, the Doppler core contributes negligibly to the net $2p \rightarrow 1s$ decay rate, which is most sensitive to the Lyman-α spectrum in the *damping wings*. At lower redshifts, the net decay rate in the Lyman-α transition is accurately captured by the Sobolev escape formalism [32, 40]. Hence, precision recombination codes need not follow the radiative transfer in the Doppler core in detail – in fact, HyRec only solves for radiative transfer in the Lorentzian wings, starting 2–3 Doppler widths away from line center.

On the other hand, the color temperature relevant for the high-redshift 21-cm problem (and especially its deviation from the matter temperature) is sensitive to the detailed shape of the spectrum *within a few Doppler widths* of line center (see Figure 5). This problem therefore requires a custom-made, high-resolution radiative transfer calculation.

The goal of this work is to thoroughly examine the little-explored 21-cm signal from $z \gtrsim 500$ by performing just such a high-resolution computation. We develop a very-high-resolution Lyα radiative transfer code to ensure that the physics of the WF effect is captured precisely. This code is coupled to the recombination code HyRec [34, 36], which very accurately predicts the high-redshift recombination and thermal history, as well as the damping wings of the Lyα line. We find that the fractional difference between color and matter temperatures at $z \gtrsim 500$ is at most a few times 10^{-5}, up to three orders of magnitude less than the deviation found by FL13.

Secondly, we explicitly account for free-free emission and absorption, which become substantial as the free-electron abundance rises rapidly at $z \gtrsim 10^3$. They lead to

a distortion to the global CMB spectrum proportional to a weighted average of the difference between matter and radiation temperature, of order ~ 1 mK. This distortion overwhelms that due to resonant interaction in the 21-cm line for frequencies $\nu \lesssim 2$ MHz. In other words, the *global* distortion to the CMB spectrum at frequencies $\nu \lesssim 2$ MHz has little to do with the 21-cm transition, and is mostly determined by free-free processes.

We also consider the *fluctuating* part of the signal. Given that the free-free contribution results from a broad integral along the line of sight, we expect it to be mostly smooth, and the fluctuations should indeed be dominated by the 21-cm signal. We find perturbations on the order of $\sim 10^{-7}$ mK at $z \sim 1000$. Beyond that redshift, the high free-free optical depth exponentially suppresses any fluctuations, implying that this is the most distant redshift at which a 21-cm signal can be observed, as anticipated in FL13. We also make a rough estimate of the sensitivity necessary to detect these fluctuations, and find that even a futuristic experiment densely covering the far side of the Moon will be unable to make a detection.

This paper is organized as follows. In Section II, we briefly review the physics underlying the 21 cm signal. We discuss the global signal around recombination in Section III, starting with a description of our high-precision Lyα radiative transfer computation and going on to add in the effects of free-free absorption. Section IV studies the amplitude of 21 cm fluctuations at these redshifts. We discuss our results in Section V and conclude in Section VI. Appendix A provides the details of our numerical implementation, followed by a discussion of convergence tests in Appendix B. We assume a flat ΛCDM cosmology throughout with parameters consistent with the latest measurements. For simplicity, we take $k_B = 1$ throughout.

II. BACKGROUND

A. General considerations

The observable for 21-cm surveys is the difference $T_{21} \equiv T_b - T_r$ between the brightness temperature T_b at a given frequency and the background temperature T_r. An observed frequency $\nu_{\rm obs}$ maps to redshift $z \equiv \nu_{21}/\nu_{\rm obs} - 1$, where $\nu_{21} \approx 1.4$ GHz is the frequency of the spin-flip transition. The brightness temperature contrast is given by

$$T_{21} = \frac{T_S(z) - T_r(z)}{1+z} \left(1 - e^{-\tau_\nu} \right) \approx \frac{T_S(z) - T_r(z)}{1+z} \tau_\nu, \tag{1}$$

The observed intensity depends on the difference between the radiation temperature T_r and the spin temperature T_S, defined by

$$\frac{n_1}{n_0} = 3e^{-T_*/T_S} \approx 3 \left(1 - \frac{T_*}{T_S} \right), \tag{2}$$

where n_1 and n_0 are the densities of atoms in the upper ($F = 1$) and lower ($F = 0$) hyperfine states, $T_\star \equiv h\nu_{21} = 0.068$ K is the temperature corresponding to the energy difference between these levels, and the second approximation is always accurate since all relevant temperatures are significantly higher than T_*.

For our purposes, we will assume that the radiation background at these frequencies is dominated by the CMB blackbody, i.e. $T_r = T_{\rm CMB,0}(1 + z)$. This is the usual assumption made for these types of computations. However, it has been proposed that the recent claimed detection of a stronger-than-expected cosmic dawn 21 cm signal by the EDGES collaboration [41] (though see [42]) may be evidence for an extra low-frequency radio background [43]. An additional radiation component would increase T_r above the CMB value, raising the amplitude of any 21-cm absorption signal. There is no reason as of yet to suspect this additional background at the redshifts we consider, but the possibility is worth noting.

The optical depth τ_ν of the 21-cm transition is given by

$$\tau_\nu = \frac{3}{32\pi} \frac{hc^3 A_{10}}{T_S \nu^2} \frac{x_{\rm HI} n_{\rm H}}{(1+z)(dv_\parallel/dr_\parallel)}, \tag{3}$$

where h is Planck's constant, c is the speed of light, $n_{\rm H}$ is the number density of hydrogen atoms, $x_{\rm HI}$ is the neutral fraction of those atoms, and A_{10} is the Einstein-A coefficient for the 21-cm line. The quantity $(dv_\parallel/dr_\parallel)$ is the derivative of the line-of-sight gas velocity along the photon's path, including both the bulk Hubble flow and peculiar velocities [2]. Combining Equations (1) and (3) yields, in the matter-domination era [1],

$$T_{21} \approx 9 \text{ mK } x_{\rm HI}(1+\delta)(1+z)^{1/2} \left[1 - \frac{T_r}{T_S}\right]$$
$$\times \left[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\right], \tag{4}$$

where δ is the density contrast at the location observed and $H(z)$ is the Hubble parameter.

The spin temperature is determined by three competing processes. Interactions with CMB photons tend to bring the spin temperature in equilibrium with the radiation temperature at a given redshift. Collisions between hydrogen atoms pull the spin temperature instead towards the matter temperature T_m. The third process, the WF effect, alters the spin temperature through interactions with Lyα photons. Indeed, electric dipole selection rules allow for inelastic resonant scattering of Lyα photons, in which the final hyperfine state differs from the initial one [29, 30]. This process acts to bring the spin temperature towards the color temperature T_C of the Lyα line. In most situations, T_C is strongly coupled to T_m, so the WF effect strengthens the coupling between the matter and spin temperatures. However, as we will show later, T_C diverges somewhat from T_m at very high redshifts, which alters the predicted value of T_S.

Under these three competing effects, the spin temperature is

$$T_S^{-1} = \frac{T_r^{-1} + x_c T_m^{-1} + x_\alpha T_C^{-1}}{1 + x_c + x_\alpha}, \tag{5}$$

where x_c and x_α are the dimensionless coupling strengths of the collisional and WF effects, respectively. The coefficient for collisions with species i is given by

$$x_c^i = \frac{n_i \kappa_{10}^i}{A_{10}} \frac{T_\star}{T_r}, \tag{6}$$

where κ_{10}^i is the rate coefficient for species i, which depends on the kinetic temperature T_m. We consider two types of collisions: those between pairs of hydrogen atoms and those between hydrogen atoms and electrons. Values for the rate coefficients for these two cases can be found in Tables 3 and 4 of Ref. [1], using calculations from Refs. [44–46].

B. The Wouthuysen-Field effect

The WF coupling coefficient can be written as

$$x_\alpha = \frac{4P_\alpha}{27A_{10}} \frac{T_\star}{T_r}, \tag{7}$$

where

$$P_\alpha \equiv \frac{8\pi}{c^2} \int f_\nu \sigma_{\rm Ly\alpha} \nu^2 d\nu \tag{8}$$

is the total Lyα scattering rate per atom, $\sigma_{\rm Ly\alpha} = (\pi e^2/m_e c) f_\alpha \phi_\alpha(\nu)$ is the absorption cross section for the Lyα line with oscillator strength $f_\alpha = 0.4162$, $\phi_\alpha(\nu)$ is the line profile, and f_ν is the dimensionless photon occupation number.

In Figure 1, we plot the values of the coupling coefficients x_c and x_α as a function of redshift. As noted by FL13, we find that the WF coupling dominates at $z \gtrsim 800$, which motivates an accurate computation of the Lyα color temperature at these redshifts. The values of f_ν used to determine x_α in this plot were computed using the Lyα radiative transfer code described below in Section III A.

The color temperature is by definition the steady-state spin temperature if the WF field effect was the only spin-changing process, i.e. [47]

$$\left(\frac{n_1}{n_0}\right)_{\rm ss} \equiv \frac{\Gamma_{01}}{\Gamma_{10}} \equiv 3\mathrm{e}^{-T_*/T_C} \approx 3\left(1 - \frac{T_*}{T_C}\right), \tag{9}$$

implying

$$\frac{T_*}{T_C} = \frac{\Gamma_{10} - \frac{1}{3}\Gamma_{01}}{\Gamma_{10}}, \tag{10}$$

where Γ_{01} and Γ_{10} are the rates of Lyα scattering events per hydrogen atom changing the hyperfine state $F = 0 \rightarrow$
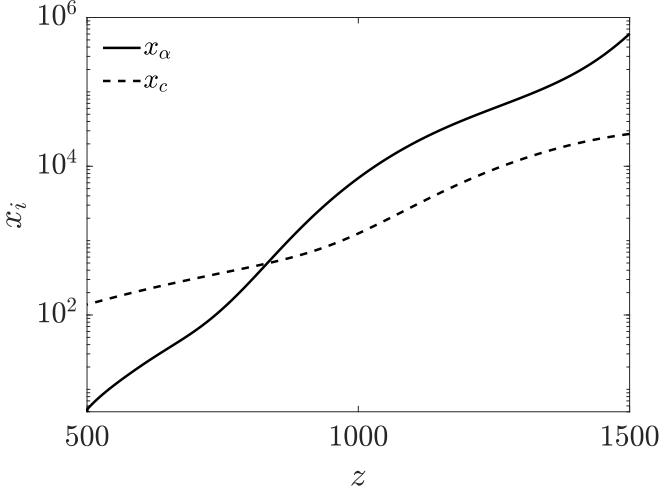
FIG. 1. Coupling strength as a function of redshift for the WF effect x_α (solid) and collisions between hydrogen atoms x_c (dashed).

1, and $1 \rightarrow 0$, respectively. Explicitly, this ratio is

$$\frac{T_*}{T_C} = \frac{\int f_\nu \left[\phi_{10}(\nu) - \phi_{01}(\nu)/3\right] d\nu}{\int f_\nu \phi_{10}(\nu) d\nu} \qquad (11)$$

where ϕ_{10} and ϕ_{01} are the relevant thermally-averaged line profiles, which are computed explicitly in Appendix B of Hirata [47]. They are related through $\phi_{01}(\nu) = 3\phi_{10}(\nu - \nu_{21})$, so that

$$\frac{T_*}{T_C} = \frac{\int f_\nu \left[\phi_{10}(\nu) - \phi_{10}(\nu - \nu_{21})\right] d\nu}{\int f_\nu \phi_{10}(\nu) d\nu}. \qquad (12)$$

Because ν_{21} is small compared to the Doppler width of the Lyα line, we can make the approximation $\phi_{10}(\nu - \nu_{21}) \approx \phi_{10}(\nu) - \nu_{21}d\phi_{10}/d\nu$. We then have

$$T_C \approx \frac{T_\star}{\nu_{21}} \frac{\int f_\nu \phi_{10}(\nu) d\nu}{\int f_\nu \frac{d\phi_{10}}{d\nu} d\nu} = -h \frac{\int f_\nu \phi_{10}(\nu) d\nu}{\int \frac{df_\nu}{d\nu} \phi_{10}(\nu) d\nu}, \qquad (13)$$

where the second equality is obtained by integrating by parts.

Let us point out that if we make the further approximation that the line is infinitely narrow, $\phi_{10}(\nu) = \delta_D(\nu - \nu_{\mathrm{Ly}\alpha})$, this reduces to

$$T_C \approx -h \left(\frac{d \ln f_\nu}{d\nu}\right)^{-1}, \qquad (14)$$

which is the common expression for the color temperature used in many computations, including those of FL13. Except when explicitly stated, we will *not* use Eq. (14) to compute the color temperature in what follows, but rather the more accurate Eq. (13).

Let us now turn our attention to the line profile. A spin-flipping scattering event can proceed through either the $2\mathrm{p}_{1/2}(\mathrm{F} = 1)$ or $2\mathrm{p}_{3/2}(\mathrm{F} = 1)$ intermediate levels.

Far from the Lyα resonance, the amplitudes for these two transitions interfere destructively, causing the profiles to differ significantly from the Voigt profile ϕ_α which characterizes the overall Lyα transition [47]. Exact forms for these profiles can be found in Equation (B18) of Hirata [47]. The unconvolved line profile is explicitly given by

$$\phi_{10} = \frac{\gamma \nu_{\mathrm{FS}}^2}{2\pi \left[(\Delta\nu - \frac{1}{2}\nu_{\mathrm{FS}})^2 + \gamma^2\right] \left[(\Delta\nu + \frac{1}{2}\nu_{\mathrm{FS}})^2 + \gamma^2\right]}, \qquad (15)$$

where $\gamma = 50$ MHz is the natural width of the Lyα transition and $\nu_{\mathrm{FS}} = 10.9$ GHz is the fine-structure frequency separation. Compared to the Voigt profile, this interference profile falls off substantially faster in the wings of the Lyα line, as $\Delta\nu^{-4}$ instead of $\Delta\nu^{-2}$. This behavior is illustrated in Figure 2, both for unconvolved and convolved profiles. The two separate fine-structure lines are visible in the unconvolved interference profile, but as ν_{FS} is smaller than a doppler width, the convolved profile is single-peaked. Just outside the doppler core, the amplitude of interference profile is ~ 4 orders of magnitude lower than that of the Voigt profile, dramatically reducing the importance of the damping wings. In fact, we find that using a Gaussian profile alone instead of the interference profile to compute the color temperature obtains the correct result to within $\sim$a few percent.

III. GLOBAL 21CM SIGNAL

In this section, we will compute the globally averaged 21 cm signal at $z > 500$. Understanding the spin temperature and ultimately the brightness temperature at high redshifts requires accurate computations of several quantities. We obtain the matter temperature and level populations from HYREC. This code does solve for the Lyα radiation field, but with a frequency resolution of a few Doppler widths, and only in the damping wings. While this is sufficient for high-accuracy recombination calculations, it may not be enough for our purposes, so we write our own supplemental high-resolution radiative transfer code to handle the region around $\nu_{\mathrm{Ly}\alpha}$. We will also use the free-election fraction from HYREC to compute the effect of free-free processes on our results and produce a complete estimate of observed brightness temperature contrast at a given observing frequency.

A. Lyman-α Radiative Transfer

To compute the color temperature, we need a high-precision prediction for the spectrum f_ν in the Lyα region. The radiative transfer equation we need to solve for this spectrum is [38, 39, 48]

$$\frac{\partial f_\nu}{\partial t} = H\nu \frac{\partial f_\nu}{\partial \nu} + \dot{f}_\nu|_{\mathrm{em}} + \dot{f}_\nu|_{\mathrm{ab}} + \dot{f}_\nu|_{\mathrm{sc}}. \qquad (16)$$
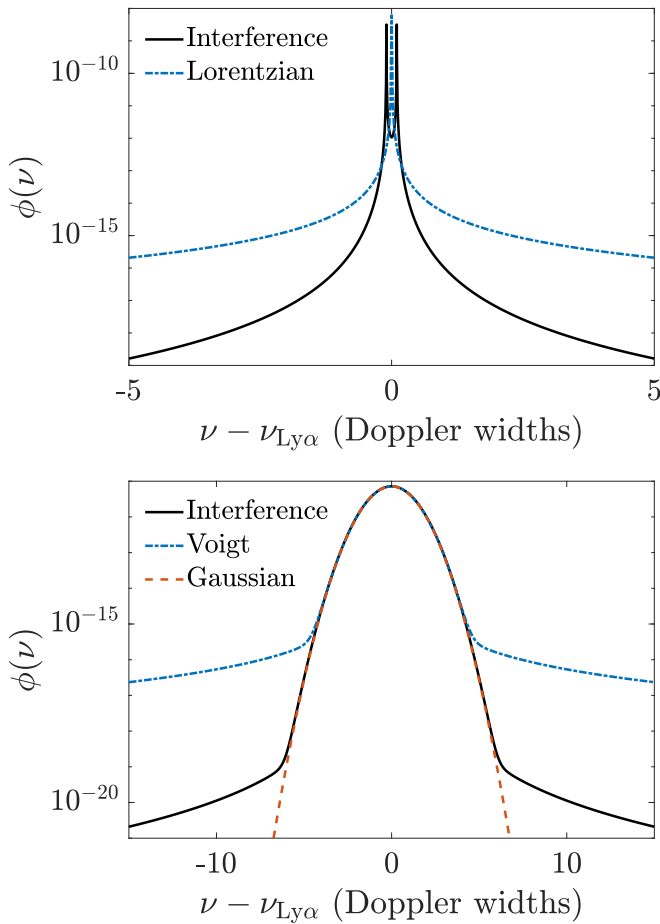
FIG. 2. Lyα line profiles, both unconvolved (top) and convolved (bottom) with the thermal Gaussian profile. Shown are both the usual Lorentzian/Voigt profile (blue dot-dashed) and the WF interference profile (black solid), along with the Doppler profile (red dashed). All profiles are computed at $z = 1000$ using the matter temperature T_m at that redshift.

In the order written above, the evolution of the spectrum is determined by Hubble drift, true emissions and absorptions of Lyα photons, and scatterings. Absorptions and emissions change the total number of photons, while scatterings and the Hubble drift conserve the photon number but shift their frequencies.

We evolve the spectrum using the method of Ref. [48], hereafter HF09. We briefly summarize the computation here, with further details presented for the convenience of the reader in Appendix A. We divide the frequency range around the Lyα line into bins of constant width $\Delta \ln \nu$. In a bin centered at frequency ν_i we evolve the quantity

$$N_i = \frac{8\pi\nu_i^3 \Delta \ln \nu}{c^3 n_H} f(\nu_i), \qquad (17)$$

which gives the number of photons per hydrogen atom in a given bin. We use $10^5 + 1$ bins centered at $\nu_{\mathrm{Ly}\alpha}$ with widths $\Delta \ln \nu = 10^{-7}$. These bins are considerably smaller than those used by HF09, but this fine resolution

is necessary to produce a converged final spectrum (for a discussion of convergence tests, see Appendix B). We use a logarithmic step in scale factor equal to our logarithmic frequency interval, $\Delta \ln a = \Delta \ln \nu$. This allows for a simple implementation of Hubble drift: at each timestep we simply redshift the photon numbers by one frequency bin. After each step, we fill in the boundary bin on the blue side from the output of HyRec. We start our radiative transfer with a blackbody initial condition at $z = 1600$ with temperature equal to the radiation temperature.

The true emission rate is computed based on the population of the various excited states of a hydrogen atoms along with their decay rates. Level populations are obtained from HyRec, and the emission rate is corrected[1] to take into account phase space factors and other nonresonant effects. The absorption rate is obtained from the emission rate by detailed balance, assuming the low-lying ns and nd states are in Boltzmann equilibrium with $2p$ at temperature T_r.

The spectrum of photons near the Lyα transition is smooth on the scales of a doppler width (see e.g. the discussion in [39]), so we can evolve the spectrum using the Fokker-Planck diffusion approximation [49]. In principle, we should consider all six of the different transitions described in Ref. [47] for different hyperfine levels. However, as stated above, all of these different lines fall well within the Lyα Doppler width at $z \gtrsim 500$, so we will again treat the Lyα line as a single transition with a single Voigt profile (the interference term in ϕ_{10} does not appear in the full Lyα profile). We only account for kinetic diffusion, and neglect the spin diffusion term considered in Ref. [47]. This is justified because the hyperfine splitting is much smaller than a Doppler width at the redshifts of interest. We use a tri-diagonal discretized diffusion operator, for which we enforce detailed balance at the matter temperature T_m. This translates the fact that Lyα photons mostly exchange energy with the kinetic degrees of freedom of the gas during scattering events (we will get back to the exchange of energy with the spin degrees of freedom below). We evolve the photon field at each timestep using an implicit Euler method, after which we shift each value redward to account for Hubble redshifting.

The result of our computation at $z = 1000$ is plotted in Figure 3. The solid curve shows the spectrum including all of the effects discussed above, while the dot-dashed curve shows the effect of neglecting scattering. As expected, both spectra approach the equilibrium spectrum $f_\nu^{\mathrm{eq}} \approx (x_{2p}/3x_{1s}) \exp\left[-h(\nu - \nu_{\mathrm{Ly}\alpha})/T_r\right]$ (shown by the dashed curve) within a few doppler widths of the line center. Note that, though the difference cannot be seen by eye, scattering actually drives the spectrum towards equilibrium with the kinetic degrees of freedom of matter at temperature T_m rather than T_r. See Figure 5 for more details.
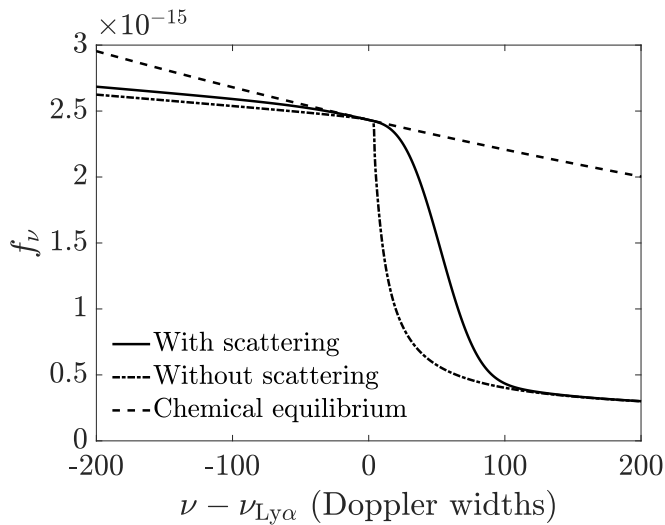
---

[1] See Eq. (A2) in Appendix A

FIG. 3. Spectra near the Lyα transition at $z = 1000$, calculated assuming the full diffusion behavior (solid), neglecting scattering effects (dot-dashed), and assuming equilibrium (dashed). Values on the x-axis give the number of doppler widths from line center.

B. Lyα Color temperature

Now that we have our Lyα spectrum, we can use the results from Section II to compute the color temperature. In Figure 4 we plot the deviation of the color temperature from the radiation temperature as a function of redshift. The black solid curve shows the result evaluated by applying Equation 13 to the spectrum computed in the previous section. We have also plotted for comparison the matter temperature output from HyRec (black dot-dashed curve). From this it can clearly be seen that at $z \sim 500$, the usual assumption that $T_C \approx T_m$ is quite accurate. However, at $z \gtrsim 1000$, the color temperature difference approaches a plateau, differing from T_r by a few parts in 10^5. At the highest redshifts we consider this effect has become quite significant, as the color temperature contrast is about an order of magnitude higher than that of the matter temperature.

We also plot for the sake of comparison the color temperature obtained using the narrow-line approximation from Equation 14 (blue dot-dashed curve). The color temperature calculated in this limit tracks the matter temperature to much higher redshifts, plateauing to a value a few times lower than that of the full computation. This suggests that photons outside of the very center of the line contribute significantly to the deviation of the color temperature from the gas temperature.

We can confirm this conclusion by looking closer at our computed spectrum. We start by using Eq. (13) to write

$$1 - \frac{T_C}{T_m} \approx \frac{T_m}{T_C} - 1 \approx \int \mathcal{F}(\nu) d\nu \qquad (18)$$
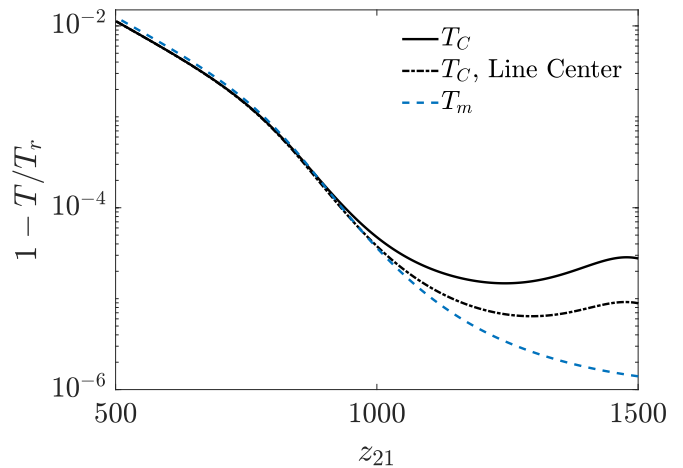

FIG. 4. Deviation of color temperature T_C from radiation temperature T_r as a function of redshift, computed both using the full Lyα spectrum as in Eq. (13) (black solid) and only the line center as in Eq. (14) (black dot-dashed). Shown for comparison (blue dashed) is the deviation between the matter temperature T_m and the radiation temperature T_r.

where we have defined

$$\mathcal{F}(\nu) \equiv -\frac{1}{f_{\nu_\alpha}} \left[f_\nu + \frac{T_m}{h} \frac{df_\nu}{d\nu} \right] \phi_{10}(\nu). \qquad (19)$$

It can easily be seen that $\mathcal{F}(\nu) = 0$ and $T_C = T_m$ if the spectrum is a blackbody at temperature T_m (in the Wien-tail approximation).

Figure 5 uses the quantity $\mathcal{F}(\nu)$ to illustrate the source in frequency space of the deviation between color and matter temperatures. The blue curve shows $\mathcal{F}(\nu)$ (multiplied by a Doppler width) for the same spectrum used in Figure 4 at $z = 1000$, roughly the point where T_C begins to diverge from T_m. Results are shown using both the full interference profile and the doppler profile alone. A sizable excess can be seen in the five or so doppler widths blueward of the line center, translating to the excess over the line center value seen in Figure 4. Note that the values of $\mathcal{F}$ for both profiles are quite similar. They yield the same value of $1 - T_C/T_r$ to within 3%, justifying our use of a Gaussian profile in our full calculation.

C. Cosmological free-free effects

At the redshifts we consider here, the large free-electron abundance means that we must consider the effects of free-free processes on our spectrum. We will quantify the strength of these effects here.

We denote by $\alpha_{ff}(\nu, z)$ the net free-free absorption coefficient (including stimulated emission) for photons with frequency ν at redshift z. For thermal Bremsstrahlung where the electron temperature is equal to the matter
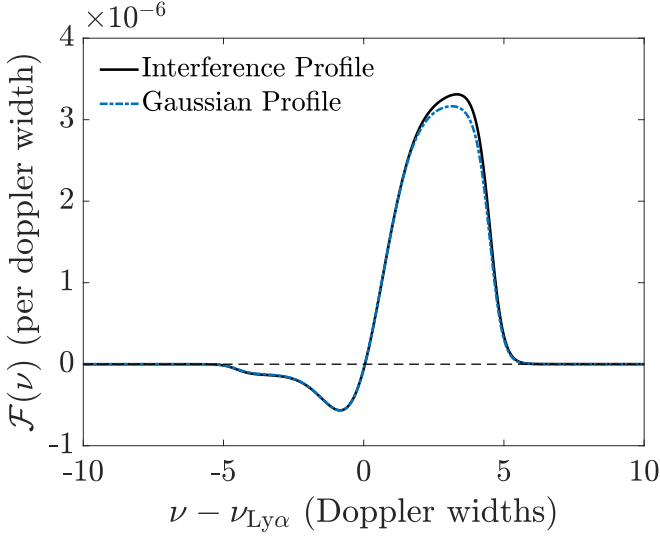
FIG. 5. The relative contribution from different frequencies to the difference between T_C and T_m at $z = 1000$, illustrated using the quantity $\mathcal{F}(\nu)$ defined in Equation (19). The black solid curve shows the result using the full convolved interference profile, while the blue dot-dashed curve shows the result obtained from a Gaussian profile alone. Both curves are normalized such that their integral yields the correct value of $1 - T_C/T_m$.

temperature, we have [50]

$$\alpha_{\mathrm{ff}}(\nu, z) = \frac{4e^6}{3m_e c\nu^2}\left(\frac{2\pi}{3m_e T_m^3}\right)^{1/2} n_e^2 \, g_{\mathrm{ff}}(\nu, T_m), \quad (20)$$

where n_e is the number density of free electrons (or protons, assuming that Helium is fully recombined). We obtain values for the free-free Gaunt factor g_{ff} from [51]. Note that, as in our previous calculations, we are working in the Rayleigh-Jeans regime.

We denote by $T_b(\nu_0, z)$ the brightness temperature at frequency $(1 + z)\nu_0$ and redshift z, and $t_b(\nu_0, z) \equiv [T_b(\nu_0, z) - T_r(z)]/(1 + z)$. We also define $t_m(z) \equiv [T_m(z) - T_r(z)]/(1+z)$ and $t_s(z) \equiv [T_s(z) - T_r(z)]/(1+z)$. In the Rayleigh-Jeans regime of interest, the radiative transfer equation can be rewritten as the following differential equation for t_b, at constant ν_0:

$$\frac{dt_b}{dz} = \frac{d\tau_{\mathrm{ff}}}{dz}(t_b - t_m) + \frac{d\tau_{21}}{dz}(t_b - t_s), \quad (21)$$

where the differential free-free optical depth is

$$\frac{d\tau_{\mathrm{ff}}}{dz}(\nu_0, z) \equiv \frac{\alpha_{\mathrm{ff}}(\nu_0(1+z), z)}{(1+z)H(z)}, \quad (22)$$

and, approximating the 21-cm line as narrow,

$$\frac{d\tau_{21}}{dz}(\nu_0, z) \equiv \tau_{21}(z_{21})\delta_{\mathrm{D}}(z - z_{21}), \quad 1 + z_{21} \equiv \frac{\nu_{21}}{\nu_0}. \quad (23)$$

In this equation, z_{21} is the redshift at which a photon of frequency ν_0 today was resonant with the 21-cm hyperfine transition.

This equation can be solved in three steps: (i) Assuming $t_b(z = \infty) = 0$, the brightness temperature just prior to resonance is

$$t_b(z_{21}^+) = t_{\mathrm{ff}}(\nu_0, z_{21}), \quad (24)$$

where

$$t_{\mathrm{ff}}(\nu_0, z) \equiv \int_z^\infty dz' \, \exp\left[-\tau_{\mathrm{ff}}(\nu_0; z, z')\right]\frac{d\tau_{\mathrm{ff}}}{dz'} t_m(z'), (25)$$

$$\tau_{\mathrm{ff}}(\nu_0; z, z') \equiv \int_z^{z'} dz'' \frac{d\tau_{\mathrm{ff}}}{dz''}(\nu_0, z''). \quad (26)$$

(ii) Assuming $\tau_{21} \ll 1$, the change in brightness temperature after resonant interaction with the 21-cm line is

$$t_b(z_{21}^-) \approx t_b(z_{21}^+)(1 - \tau_{21}) + \tau_{21}t_s \approx t_b(z_{21}^+) + \tau_{21}t_s. \quad (27)$$

(iii) After resonant interaction with the 21-cm line, the radiation changes through free-free processes alone:

$$T_{21}(\nu_0) = t_b(\nu_0, 0) = t_b(z_{21}^-) \exp\left[-\tau_{\mathrm{ff}}(\nu_0; 0, z_{21})\right]$$
$$+ \int_0^{z_{21}} dz \, \exp\left[-\tau_{\mathrm{ff}}(\nu_0; 0, z)\right]\frac{d\tau_{\mathrm{ff}}}{dz} t_m(z) (28)$$

Sewing everything together, we arrive at the following brightness temperature contrast at redshift zero,

$$T_{21}(\nu_0) = \tau_{21}t_s(z_{21}) \exp\left[-\tau_{\mathrm{ff}}(z_{21})\right] + t_{\mathrm{ff}}(\nu_0, 0), \quad (29)$$

where, for short, we have defined

$$\tau_{\mathrm{ff}}(z_{21}) \equiv \tau_{\mathrm{ff}}(\nu_0; 0, z_{21}). \quad (30)$$

So the effect of free-free absorption on the global signal is twofold: (i) it exponentially damps the standard contribution by $e^{-\tau_{\mathrm{ff}}}$, and (ii), it adds a frequency-dependent offset $t_{\mathrm{ff}}(\nu_0, 0)$. We can rewrite the latter as the following weighted integral of t_m:

$$t_{\mathrm{ff}}(\nu_0, 0) \equiv \int_0^\infty dz \, V_{\mathrm{ff}}(\nu_0, z)t_m(z), \quad (31)$$

where

$$V_{\mathrm{ff}}(\nu_0, z) \equiv \exp\left[-\tau_{\mathrm{ff}}(\nu_0; 0, z)\right]\frac{d\tau_{\mathrm{ff}}}{dz}(\nu_0, z). \quad (32)$$

The function V_{ff} plays the role of a visibility function similar to that of CMB anisotropies: at fixed ν_0 its integral over redshift is unity, and it weighs the contribution of the source function $t_m(z)$ in Eq. (31).

Figure 6 shows the value of τ_{ff} between redshifts 0 and z_{21} for photons emitted at $\nu_0 = \nu_{21}/(1 + z_{21})$, computed using the output of HYREC. For photons emitted well after recombination, the optical depth is small because the ionization fraction is small. The optical depth then rises rapidly following the rising ionization fraction, reaching unity around $z \sim 1000$ and several thousand by $z \sim 1500$.

Figure 7 shows the visibility function V_{ff} defined in Eq. (32) for instruments targeting $z_{21} = 500, 1000,$ and 1500.
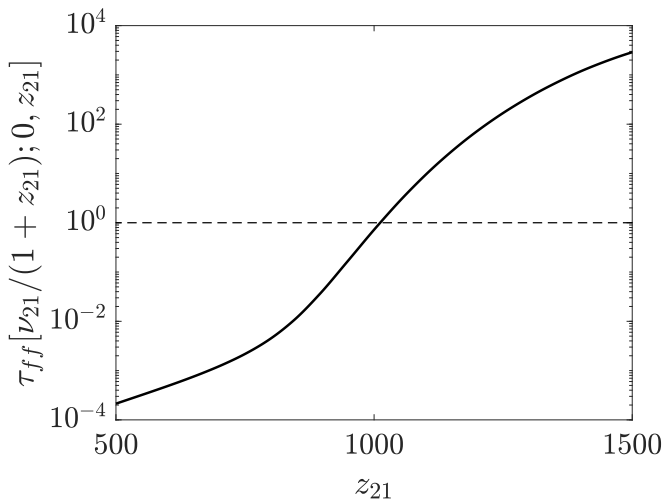
FIG. 6. Free-free optical depth τ_{ff} between redshifts 0 and z_{21} for photons with frequency ν_{21} at z_{21}. The free-free process becomes optically thick (marked with dashed line) at roughly $z \sim 1000$.
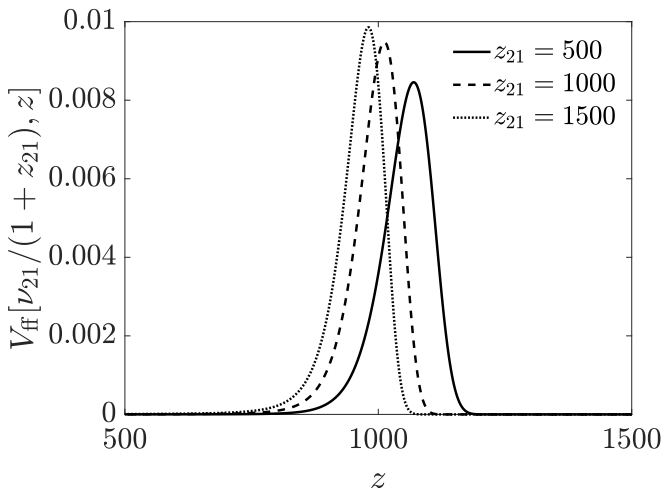

FIG. 7. Free-free visibility functions $V_{\rm ff}$ for instruments targeting redshifts z_{21}=500 (solid), 1000 (dashed), and 1500 (dotted).

One can easily show [e.g. for a power-law $\tau_{\rm ff}(z)$] that the visibility function peaks near $\tau_{\rm ff} \sim 1$, i.e a surface of last free-free interaction, by analogy to the CMB surface of last scattering. As expected, our plotted $V_{\rm ff}(z)$'s all peak near $z \sim 1000$, roughly where we expect $\tau_{\rm ff}$ to be unity. The peaks shift slightly because each curve deals with photons at different frequencies. As one increases the redshift z_{21} of resonance (hence decreasing ν_0), the free-free optical depth increases, and the free-free visibility function peaks at lower redshifts, where $|t_m|$ is larger. We therefore expect $|t_{\rm ff}(\nu_0, 0)|$ to increase with z_{21}. In other words, the impact of free-free effects on our signal should be higher for instruments targeting higher redshifts.

D. Galactic free-free effects

The plasma in the Milky Way ISM is also a source of free-free opacity at low frequencies, which can obscure the cosmological signal. The free-free opacity can be modeled as

$$\tau_{\rm ff, Gal}(\nu) = \left(\frac{\nu_\star}{\nu}\right)^{2.1}, \tag{33}$$

where $\nu_\star$ is a transition frequency where the optical depth is unity, which depends on the line of sight (lines of sight with longer path length, or through denser or colder gas, will have more absorption and a larger $\nu_\star$). The power law deviates slightly from 2 due to the Coulomb logarithmic factor in the Gaunt factor. Modeling of the radio spectrum of the Galactic poles – which is dominated by synchrotron emission and free-free absorption – suggests $\nu_\star \approx 2.2$ MHz (e.g., Ref. [52]; this is $F = 5$ in their notation). This means that at $z \gtrsim 650$, the Milky Way ISM is optically thick to the cosmological 21 cm signal in the sense that $\tau_{\rm ff, Gal} > 1$; however, in the range considered in this paper the optical depth is only a few, and so some signal survives and reaches an observer inside the Milky Way. Furthermore, it seems likely that the free-free absorption will turn out to be patchy when the MHz sky is observed at high resolution instead of with single antennas, with dense structures corresponding to peaks in $\tau_{\rm ff, Gal}$. (This is seen in the Galactic Plane in the 36–73 MHz band, e.g. [53].) This means that the optical depth inferred from wide-angle temperature measurements may be an underestimate on some lines of sight, and an overestimate on others.

E. Results

We plot in Figure 8 the predicted global observed brightness temperature T_{21} as a function of redshift of resonance z_{21}. Note that for ease of viewing we have plotted $-T_{21}$ so that, though $T_b < T_r$, a stronger signal still appears higher on the y-axis. The dashed red curve shows the results which would be obtained neglecting both the WF effect (by setting $x_\alpha = 0$ in Eq. (5)) and free-free effects computed above. The blue dot-dashed curve shows the effect of adding in the WF coupling with our computed $T_C(z_{21})$. As expected based on Figure 4, this has the effect of substantially increasing the amplitude of the observed signal beyond $z_{21} \sim 1000$. The solid black curve shows the full result of evaluating Equation (29) including free-free. At $z_{21} = 500$, the 21 cm signal still dominates over the free-free contribution. However, as predicted in the previous section, the amplitude of the free-free effect increases with redshift, overwhelming the 21 cm component for $z_{21} \gtrsim 800$.

It should also be noted that all three curves in Figure 8 converge at $z \sim 500$. Below this redshift, the WF and free-free effects we consider here become subdominant to collisional coupling to the cooling hydrogen atoms.
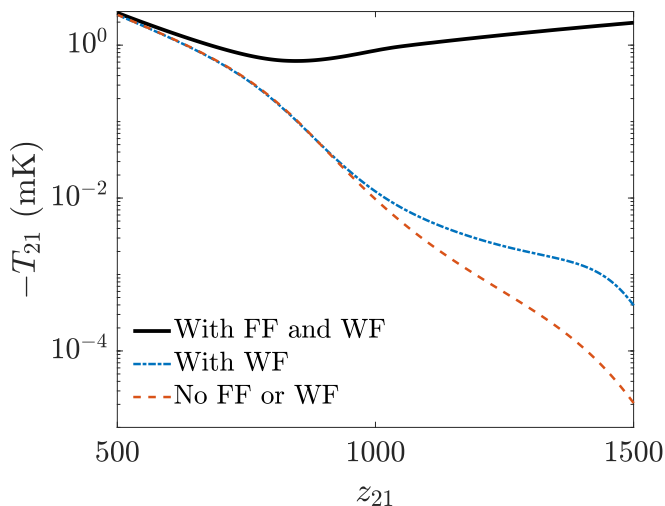
FIG. 8. Predicted brightness temperature in the 21 cm line as a function of redshift of resonance. Results are plotted using our full computation (solid black), as well as in the limits where free-free (dot-dashed blue) and both free-free and WF (dashed red) are neglected. Note that the y-axis shows the negative of the true brightness temperature contrast. Note that, though we have plotted all three curves as a function of z_{21}, unlike the other two the black curve includes free-free contributions from the entire line of sight, not just a narrow band around z_{21}.

Therefore, past computations of the 21 cm signal for $z \lesssim 500$ should not be significantly affected by the processes we consider here.

In conclusion of this section, we found that the *global* (monopole) distortion to the CMB at frequencies $\nu \lesssim 2$ MHz is mostly determined by free-free interactions, and has little dependence on the 21-cm transition. We now turn on to the study of fluctuations.

IV. 21 CM FLUCTUATIONS

We now turn our attention to the fluctuations in the 21 cm signal from the recombination epoch, which could in principle be detectable with interferometers.

A. Accessible scales

The derivation above still applies to the case where the optical depths and temperature are spatially-dependent functions. Our first consideration is the range of radial ($k_\parallel$) and transverse ($k_\perp$) wavenumbers that are *in principle* observable. In this section, we define the re-scaled redshift variable

$$z_3 \equiv \frac{1+z}{1000} \qquad (34)$$

for simplicity of notation.

1. Radial wavenumbers

In the radial direction, we recall that a sinusoidal fluctuation in brightness temperature consisting of N oscillations per e-fold in frequency has a radial wavenumber

$$k_\parallel = 2\pi N \frac{H(z)}{c(1+z)} = 0.024 N z_3^{1/2} \mathrm{Mpc}^{-1}. \qquad (35)$$

If one fits and removes a spectrally smooth foreground model, then modes with low N are necessarily lost. The minimum N that survives depends on the complexity of the foreground model required and the bandwidth over which it can be fit. Even in the extraordinarily over-optimistic case that over an on octave-wide bandwidth ($\Delta \ln \nu = 0.7$) a quartic polynomial (2 oscillations) would suffice, we would have $N_{\min} \approx 2/0.7$ and $k_{\parallel,\min} \sim 0.07 z_3^{1/2} \mathrm{Mpc}^{-1}$. Realistic foregrounds, and consideration of spectral structure imprinted by the instrument, may set a much larger value of $k_{\parallel,\min}$.

2. Transverse wavenumbers

In the transverse direction, scattering in the Milky Way's ISM provides a serious obstacle to observations at low frequencies. This scattering smears any cosmological signal into a seeing disk, or equivalently suppresses the electric field correlation function at long baselines. The NE2001 model [54] gives a scattering measure of 8.5×10^{-5} (1.8×10^{-4}) kpc m$^{-20/3}$ at the North (South) Galactic Pole. This gives an angular broadening of $\theta = 0.0041 z_3^{2.2}$ radians at the NGP (the coefficient is 0.0064 at the SGP). This implies a cutoff at the scale

$$k_{\perp,\max} = \frac{2.355}{\theta D(z)} = 0.041 z_3^{-2.2} \mathrm{Mpc}^{-1} \qquad (36)$$

at the NGP (the coefficient is 0.026 at the SGP). Fourier modes with larger transverse wavenumbers than this are washed out. There is quite a bit of uncertanties in these numbers, since the scattering model has not been validated as such low frequencies; nevertheless, the qualitative picture that angular broadening is enormous in the MHz regime seems inescapable.

Regardless of the direction of the Fourier mode, there is a maximum wavenumber – the baryon Jeans scale k_{Jeans} – beyond which the cosmological signal decreases rapidly. Given Eq. (36), this is only relevant in the radial direction, and sets $k_{\parallel,\max} \sim k_{\mathrm{Jeans}}$.

B. Estimation of the signal strength

We estimate the characteristic fluctuation δT_{21} of these fluctuations as the square root of the dimensionless 21-cm power spectrum, using a method similar to that of FL13. To facilitate comparison with their results, we neglect perturbations to the ionization fraction and to the

gas temperature. The fluctuation amplitude is then determined by fluctuations in the optical depth, which only depend on the baryon density and velocity perturbations, δ_b and $\delta_v \equiv (1+z) \, dv_\parallel/dr_\parallel/H(z) - 1$, respectively, as well as perturbations in the spin temperature. In Fourier space, and assuming linear perturbations, δ_v is given by the linearized continuity equation, implying

$$\delta_v = -\mu^2 \kappa \delta_b, \quad \mu \equiv \hat{k} \cdot \hat{n} = k_\parallel/k, \quad (37)$$

where $\hat{n}$ is the line of sight, $\hat{k}$ is the direction of the Fourier wavenumber and $\kappa \equiv |d\ln\delta_b/d\ln(1+z)|$ is the growth rate of baryonic perturbations. Typically, we find $\kappa \sim 5$ for the wavelengths and redshifts of interest. This is larger than unity (as assumed in FL13) because near recombination, baryons start nearly unclustered and quickly catch up to the dark matter. Using Eq. (4), we see that the 21-cm brightness contrast is proportional to $(1+\delta_b-\delta_v)(1-T_r/T_s) = \big(1+(1+\mu^2\kappa)\delta_b\big)\,(1-T_r/T_s)$.

Because the coupling x_c between T_m and T_s increases with baryon density, we can write $x_c = \overline{x}_c(1+\delta_b)$. After some algebra, we can then say that

$$1 - T_r/T_s = \overline{1 - T_r/T_s}\,(1+X\delta_b), \quad (38)$$

where

$$X \equiv \frac{\overline{x}_c(T_m-T_r)}{\overline{x}_c(T_m-T_r)+x_\alpha(T_c-T_r)} - \frac{\overline{x}_c}{1+\overline{x}_c+x_\alpha}, \quad (39)$$

and we have made the approximation $1 - T_r/T_m \approx T_m/T_r - 1$. The intensity fluctuation amplitude is then

$$\delta T_{21} = T_{21}(1+\delta_b-\delta_v+X\delta_b)$$
$$= (1+X+\mu^2\kappa)\delta_b. \quad (40)$$

We need now only one more component for our fluctuation amplitude: the effect of Bremsstrahlung. The line-of-sight integral in Eq. (31) implies that the second term in Eq. (29) ought to have little fluctuations on scales smaller than the width of the visibility function, corresponding to wavenumbers $k_\parallel \gtrsim 0.05$ Mpc^{-1}. Given the lower limit on observable radial wavenumbers discussed in Section IV A 1, we conclude that foreground cleaning will remove any observable fluctuations of the second term in Eq. (29). Fluctuations in the first term do survive, albeit supressed by free-free opacity:

$$\delta T_{21}^{\rm obs} = \delta T_{21}^0 e^{-\tau_{\rm ff}}. \quad (41)$$

This will exponentially suppress any fluctuations beyond $z \sim 1000$.

Figure 9 shows the amplitude of the temperature fluctuations integrated over the range of accessible scales defined in the previous section. Specifically, we plot the square root of the variance of the fluctuations,

$$(\mathrm{Var}\,T_{21})^{1/2}$$
$$= \left[\int P_{21}(k)\frac{d^3k}{(2\pi)^3}\right]^{1/2}$$
$$= \left[\frac{T_{21}^2}{(2\pi)^2}\int k_\perp(1+X+\mu^2\kappa)^2 P_b dk_\perp dk_\parallel\right]^{1/2}, \quad (42)$$
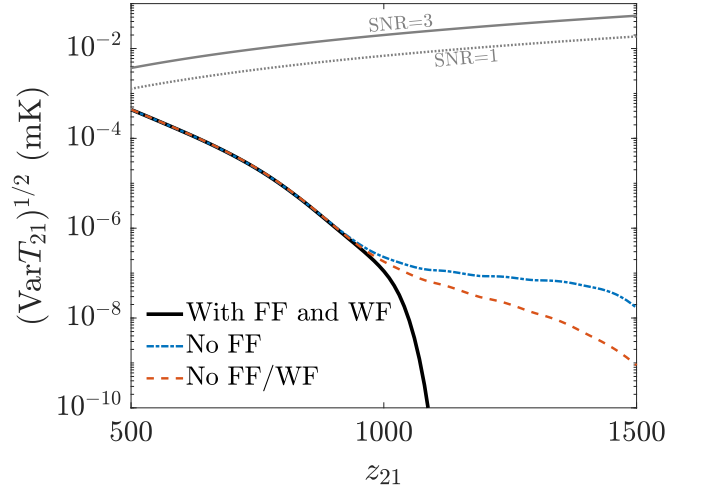

FIG. 9. Amplitude of the 21-cm brightness fluctuations integrated over the observable range of scales under our full calculation (solid black curve), neglecting free-free effects (blue dot-dashed), and neglecting both free-free and WF effects (red dashed). Also plotted are the minimum amplitudes detectable with an idealized instrument with $S/N = 3$ (grey solid) and $S/N = 1$ (grey dotted).

where we have computed the baryon power spectrum $P_b(k)$ using CAMB [55]. In addition to the results of our full computation, we have also plotted amplitudes with the free-free and WF effects removed. Though the WF effect acts to increase the amplitude over the collision-only prediction, this contribution only becomes significant above $z \sim 1000$, in the free-free-suppressed regime. Thus we see that 21-cm brightness temperature fluctuations cannot be observed even in principle beyond recombination, despite the fact that there is a non-vanishing (albeit small) global signal.

C. Idealized signal-to-noise ratio

We may now consider whether the recombination-era signal is detectable by a plausible future experiment. Due to the frequency range, this would have to be in space, either free-floating or possibly on the Moon.

For an ideal densely filled array of collecting area A, observing a region of solid angle $\Delta\Omega$ large compared to the beam size ($\lambda_{\rm obs}^2/A$), and averaging over a range of frequencies $\Delta\nu$ for an observation time $t_{\rm obs}$, the idealized uncertainty in the brightness temperature is

$$\sigma_T = \sqrt{\frac{2}{N_{\rm modes}}}\,T_{\rm sky} = \frac{\lambda_{\rm obs}}{\sqrt{2A\,\Delta\Omega\,\Delta\nu\,t_{\rm obs}}}\,T_{\rm sky}, \quad (43)$$

where $T_{\rm sky}$ is the sky temperature, $\lambda_{\rm obs}$ is the observed wavelength, and $N_{\rm modes}$ is the total number of modes of the electromagnetic field collected by the experiment within that range of frequencies and observing directions (and counting both polarizations). Based on the *Radio*

Astronomy Explorer 2 maps [56], we obtain a sky temperature near the Galactic poles of 8 MK (2.2 MHz) or 13 MK (1.31 MHz), which treated as a power law would give a brightness temperature of $T_{\rm sky} = 12z_3^{0.94}$ MK. The observed wavelength is $\lambda_{\rm obs} = 0.21z_3$ km. Then:

$$\sigma_T = 6.9 \times 10^{-3} \left(\frac{A}{A_{\rm L}} \frac{\Delta\Omega}{\pi\,{\rm sr}} \frac{\Delta\nu/\nu}{0.5} \frac{t_{\rm obs}}{100\,{\rm yr}} \right)^{-1/2} z_3^{2.44}\,{\rm mK}, \tag{44}$$

where we have scaled the area relative to $A_{\rm L} = \pi R_{\rm L}^2 = 9.5 \times 10^6\,{\rm km}^2$, the cross-sectional area of the Moon.

Let us now consider a signal on the sky with variance $\operatorname{Var} T_{21}$. Suppose that we divide a survey into $N_{\rm 3D}$ 3D pixels in frequency-angular position space of volume $\Delta\nu_1\,\Delta\Omega_1$ (units: Hz sr), each with a signal $T_{21}(\nu,\Omega)$. To obtain the most optimistic S/N calculation, we suppose that (somehow) the relevant modes of the cosmological density field were already measured and we attempted a 21 cm detection by cross-correlation against a perfect template. (Given that the only other tracer at these redshifts is the CMB itself, which provides only 2D information, such a perfect template may not exist; we therefore view this calculation as a way to obtain an optimistic bound to the S/N ratio.) In this case, the total S/N ratio would be

$$\begin{aligned} \left(\frac{S}{N}\right)^2 &= \sum_{\rm 3D\,pix} \left[\frac{T_{21}(\nu,\Omega)}{\sigma_T(\Delta\nu_1,\Delta\Omega_1)} \right]^2 \\ &= N_{3D} \frac{\operatorname{Var} T_{21}}{\sigma_T^2(\Delta\nu_1,\Delta\Omega_1)} \\ &= \frac{\operatorname{Var} T_{21}}{\sigma_T^2(\Delta\nu_{\rm tot},\Delta\Omega_{\rm tot})}, \end{aligned} \tag{45}$$

where the temperature uncertainty in the last line σ_T^2 is computed according to Eq. (44) using the bandwidth $\Delta\nu_{\rm tot}$ and the solid angle $\Delta\Omega_{\rm tot}$ of the whole survey volume instead of an individual 3D pixel.

We can see that for a "full survey" that consists of one Galactic cap (π sr), and taking a wide bin width in redshift or frequency $\Delta\nu/\nu = \Delta z/(1+z) = 0.5$, assuming the existence of a perfect cross-correlation template, assuming that we can densely fill the entire Moon with receivers and acquire 100 years of good data, and assuming perfect foreground subtraction, a signal with $(\operatorname{Var} T_{21})^{1/2} = 0.0069z_3^{2.44}$ mK could be measured with $S/N = 1$ and a signal at $0.02z_3^{2.44}$ mK could be measured with $S/N = 3$. As shown in Figure 9, even this very optimistic scenario is not sufficient for a detection of the recombination-era 21 cm signal.

V. DISCUSSION

Interactions between hydrogen atoms and Lyα photons in the very high-redshift Universe create a deviation between the spin temperature and the CMB brightness temperature, leading to a small but in principle measurable 21-cm signal at $z \gtrsim 500$. When we take into account an accurate estimate of the WF effect, we find that it along with coupling with CMB photons and collisions between hydrogen atoms produce a global 21-cm amplitude of roughly ~ 2.5 mK at $z \sim 500$, falling to ~ 0.01 mK at $z \sim 1000$. Free-free effects along the line of sight to the target redshift then increase the amplitude of the global brightness temperature deviation to $\sim 1 - 2$ mK at $z \sim 1000$ and beyond as photons redshift through successive shells of cooling free electrons.

This may make it appear at first glance that there is an observable 21-cm global signal beyond recombination. Unfortunately, once a continuum process such as Bremsstrahlung begins to dominate, it is no longer possible to isolate emission from a single redshift slice, as the Bremsstrahlung contribution comes from integrating conditions along the entire line of sight.

However, this same continuum behavior which makes it impossible to obtain global tomography at these high redshifts means that the contribution from Bremsstrahlung to 21 cm fluctuations can be subtracted along with usual 21 cm foregrounds. At $z \sim 1000$ we predict small and likely unobservable perturbations with an amplitude of $\sim 10^{-7}$ mK. Even this small signal is quickly drowned out by free-free absorption beyond $z \sim 1100$, as the ionization fraction and as a consequence the free-free optical depth steeply increase. This therefore is the ultimate redshift limit for 21 cm cosmology, beyond which no signal can be recovered.

Computing this 21 cm signal requires an extremely precise calculation of the color temperature. Because the fractional difference between the color and radiation temperatures is $\sim 10^{-4}$ at $z \sim 1000$, a 0.1% error on T_C will lead to order unity errors on the observed brightness temperature. Such an error is likely the cause of the discrepancy between our results and those of FL13, who obtained a ~ 1 mK global signal at $z \sim 1000$ from the WF effect alone, along with a fluctuation amplitude of $\sim 10^{-4}$ mK. These are ~ 100 times larger than our estimates in this redshift range.

Our computation also makes other improvements to the spin temperature calculation. We use the full frequency range of the Lyα profile when computing the color temperature, rather than just the slope at line center. As shown in Figure 4, this increases our signal amplitude by a factor of ~ 2. Fortunately, the extreme $< 0.1\%$ precision is only required in the computation of T_C, so that the small difference $T_C - T_r$, hence the brightness temperature contrast, are accurately estimated. We do not require sub-percent precision on overall multiplicative or additive factors, such as the free-free contributions. More substantial approximations can still affect our result though. For example, our fluctuation amplitude is enhanced by a factor of a few because we use the numerically-computed growth frate for the baryon perturbations, rather than the dark matter growth rate as in FL13.

There are still several effects which we have not included in our calculation, though we do not expect any

of them to alter our results significantly. In our radiative transfer calculation, we treated the line profile around Lyα as a single Voigt profile, rather than using the full form for the profile computed in [47] including all of the different hyperfine transitions. This should in principle alter both our computation of the Lyα spectrum and that of the color temperature. In addition, as we mentioned above, the WF coupling induces an additional diffusion term into the Fokker-Planck equation which we neglect. This spin-diffusion term should act to bring the color temperature into equilibrium with the spin temperature. Including this effect accurately would require recursively solving for the spin and color temperatures together. However, since we are working at high redshift, the Lyα line is fairly wide and both of these effects should be weaker. Comparison between the results of Refs. [47] and [57] suggests that both of these corrections should come in a the sub-percent level at our redshifts of interest. We also neglected the impact of ionization fluctuations on our perturbation amplitudes, which would alter the fluctuation size somewhat. While these effects, if included, may have a small effect on our computed signals, we do not expect them to be large enough to change either the dominance of free-free on the global signal or the undetectability of $z \sim 1000$ 21-cm fluctuations.

Though our computations are limited to redshifts above 500, our results have an interesting implication for current measurements of lower redshift 21 cm signals. In particular, the Experiment to Detect the Global Epoch of Reionization Signature (EDGES) experiment recently reported a deeper-than-expected 21 cm absorption trough at a frequency corresponding to the expected redshift of cosmic dawn [41]. In standard models of 21 cm evolution, it is assumed that Lyα photons from the first stars would couple T_s to the adiabatically-decreasing matter temperature at this redshift, but the spin temperature inferred from the EDGES result is significantly lower than the minimum plausible matter temperature. Barkana [58] proposes that a spin temperature below the matter temperature is evidence for baryon-dark matter interactions cooling the matter faster than expected [27, 28]. However, in our work, albeit at very different redshifts, we find a spin temperature outside the range bounded by T_m and T_r driven only by Lyα scatterings. Though our signal is quite small, it motivates further investigation of other subtle radiative transfer effects to determine whether unexpected results could be produced through more conventional means.

VI. CONCLUSION

In the wake of the first possible detection of a 21-cm signal at $z \sim 17$, it is important to determine the range of redshifts and scales that future 21-cm experiments may in principle be able to access. In this paper we have addressed the following question in detail: how close to the CMB last scattering surface could 21-cm observations reach? To do so, we presented the most accurate calculation to date of the 21-cm global intensity and fluctuation amplitudes at redshifts beyond $z \sim 500$. We performed a high-precision radiative transfer calculation of the spectrum near Lyα, in order to quantify the Wouthuysen-Field coupling at high redshifts. Despite a large coupling strength due to the intense Lyα background resulting from cosmological recombination, we found that the departure of the Lyα color temperature from the gas temperature is substantially smaller than what was found in FL13. This is due to the very efficient frequency diffusion of photons in the core of the line, combined with a suppressed contribution of the damping wings to spin-flip scatterings. We moreover accounted for the effects of free-free processes for the first time. We found that the global 21-cm signal at $z \gtrsim 800$ is, in fact, determined mostly by these processes rather than the resonant interactions with the 21-cm transition. Nonetheless, 21-cm fluctuations do retain tomographic information, in principle up to $z \sim 1100$ or so before being exponentially damped by bremsstrahlung absorption. Finally, we estimated the detectability of the signal, accounting for smooth foreground subtraction and angular smearing due to scattering in the Milky Way's interstellar medium. We found that even a highly optimistic instrumental setup, with collecting area of the order of the Moon's surface area and a century-long observation time, could detect 21-cm fluctuations only up to redshifts of a few hundreds.

On the bright side, there still remains a spectacular trove of information to be collected with the 21-cm signal at $z \lesssim 200$ or so. As instrumentation progresses and observations push their horizon deeper in the Universe's early stages, it will be important to carefully model the expected signal. As illustrated in this work, the physics of the 21-cm signal is rich and subtle, and careful modeling could very well lead to interesting surprises.

ACKNOWLEDGMENTS

[1] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Phys. Rep. **433**, 181 (2006), astro-ph/0608032.

[2] M. F. Morales and J. S. B. Wyithe, Annual Rev. Astron. Astrophys. **48**, 127 (2010), arXiv:0910.3010.

[3] J. R. Pritchard and A. Loeb, Reports on Progress in Physics **75**, 086901 (2012), arXiv:1109.6012.

[4] M. McQuinn, Annual Rev. Astron. Astrophys. **54**, 313 (2016), arXiv:1512.00086.

[5] P. Madau, A. Meiksin, and M. J. Rees, Astrophys. J. **475**, 429 (1997), astro-ph/9608010.

[6] M. P. van Haarlem *et al.*, Astron. Astrophys. **556**, A2 (2013), arXiv:1305.3550 [astro-ph.IM].

[7] Z. S. Ali *et al.*, Astrophys. J. **809**, 61 (2015), arXiv:1502.06016.

[8] S. J. Tingay *et al.*, PASA **30**, e007 (2013), arXiv:1206.6945 [astro-ph.IM].

[9] D. R. DeBoer *et al.*, ArXiv e-prints (2016), arXiv:1606.07473 [astro-ph.IM].

[10] M. Santos, P. Bull, D. Alonso, S. Camera, P. Ferreira, G. Bernardi, R. Maartens, M. Viel, F. Villaescusa-Navarro, F. B. Abdalla, M. Jarvis, R. B. Metcalf, A. Pourtsidou, and L. Wolz, Advancing Astrophysics with the Square Kilometre Array (AASKA14) , 19 (2015), arXiv:1501.03989.

[11] J. O. Burns, J. Lazio, S. Bale, J. Bowman, R. Bradley, C. Carilli, S. Furlanetto, G. Harker, A. Loeb, and J. Pritchard, Advances in Space Research **49**, 433 (2012), arXiv:1106.5194.

[12] K. Bandura *et al.*, in *Ground-based and Airborne Telescopes V*, Proc. SPIE, Vol. 9145 (2014) p. 914522, arXiv:1406.2288 [astro-ph.IM].

[13] L. B. Newburgh *et al.*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Proc. SPIE, Vol. 9906 (2016) p. 99065X, arXiv:1607.02059 [astro-ph.IM].

[14] A. Cohen, A. Fialkov, R. Barkana, and M. Lotem, ArXiv e-prints (2016), arXiv:1609.02312.

[15] A. Loeb and M. Zaldarriaga, Physical Review Letters **92**, 211301 (2004), astro-ph/0312134.

[16] A. Lewis and A. Challinor, Phys. Rev. D **76**, 083005 (2007), astro-ph/0702600.

[17] L. W. Brown, Astrophys. J. **180**, 359 (1973).

[18] R. Manning and G. A. Dulk, Astron. Astrophys. **372**, 663 (2001).

[19] R. Barkana, Phys. Rep. **645**, 1 (2016), arXiv:1605.04357.

[20] Y. Ali-Haïmoud, P. D. Meerburg, and S. Yuan, Phys. Rev. D **89**, 083506 (2014), arXiv:1312.4948.

[21] M. Shiraishi, J. B. Muñoz, M. Kamionkowski, and A. Raccanelli, Phys. Rev. D **93**, 103506 (2016), arXiv:1603.01206.

[22] A. Cooray, Physical Review Letters **97**, 261301 (2006), astro-ph/0610257.

[23] A. Pillepich, C. Porciani, and S. Matarrese, Astrophys. J. **662**, 1 (2007), astro-ph/0611126.

[24] J. B. Muñoz, Y. Ali-Haïmoud, and M. Kamionkowski, Phys. Rev. D **92**, 083508 (2015), arXiv:1506.04152.

[25] S. R. Furlanetto, S. P. Oh, and E. Pierpaoli, Phys. Rev. D **74**, 103502 (2006), astro-ph/0608385.

[26] A. Natarajan and D. J. Schwarz, Phys. Rev. D **80**, 043529 (2009), arXiv:0903.4485 [astro-ph.CO].

[27] H. Tashiro, K. Kadota, and J. Silk, Phys. Rev. D **90**, 083522 (2014), arXiv:1408.2571.

[28] J. B. Muñoz, E. D. Kovetz, and Y. Ali-Haïmoud, Phys. Rev. D **92**, 083528 (2015), arXiv:1509.00029.

[29] S. A. Wouthuysen, Astron. J. **57**, 31 (1952).

[30] G. B. Field, Proceedings of the IRE **46**, 240 (1958).

[31] A. Fialkov and A. Loeb, J. Cosm. Astropart. Phys. **11**, 066 (2013), arXiv:1311.4574.

[32] P. J. E. Peebles, Astrophys. J. **153**, 1 (1968).

[33] Y. B. Zel'dovich, V. G. Kurt, and R. A. Syunyaev, Soviet Journal of Experimental and Theoretical Physics **28**, 146 (1969).

[34] Y. Ali-Haïmoud and C. M. Hirata, Phys. Rev. D **83**, 043513 (2011), arXiv:1011.3758.

[35] J. Chluba and R. M. Thomas, Mon. Not. R. Astron. Soc. **412**, 748 (2011), arXiv:1010.3631.

[36] Y. Ali-Haïmoud and C. M. Hirata, Phys. Rev. D **82**, 063521 (2010), arXiv:1006.1355.

[37] Y. Ali-Haimoud, *A new spin on primordial hydrogen recombination and a refined model for spinning dust radiation*, Ph.D. thesis, California Institute of Technology (2011).

[38] C. M. Hirata, Phys. Rev. D **78**, 023001 (2008), arXiv:0803.0808.

[39] Y. Ali-Haïmoud, D. Grin, and C. M. Hirata, Phys. Rev. D **82**, 123502 (2010), arXiv:1009.4697 [astro-ph.CO].

[40] S. Seager, D. D. Sasselov, and D. Scott, Astrophys. J. Lett. **523**, L1 (1999), astro-ph/9909275.

[41] J. D. Bowman, E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and M. Nivedita, Nature **555**, 67 (2018).

[42] R. Hills, G. Kulkarni, P. D. Meerburg, and E. Puchwein, ArXiv e-prints (2018), arXiv:1805.01421.

[43] C. Feng and G. Holder, Astrophys. J. Lett. **858**, L17 (2018), arXiv:1802.07432.

[44] B. Zygelman, Astrophys. J. **622**, 1356 (2005).

[45] K. Sigurdson and S. R. Furlanetto, Physical Review Letters **97**, 091301 (2006), astro-ph/0505173.

[46] S. R. Furlanetto and M. R. Furlanetto, Mon. Not. R. Astron. Soc. **374**, 547 (2007), astro-ph/0608067.

[47] C. M. Hirata, Mon. Not. R. Astron. Soc. **367**, 259 (2006), astro-ph/0507102.

[48] C. M. Hirata and J. Forbes, Phys. Rev. D **80**, 023001 (2009), arXiv:0903.4925.

[49] G. B. Rybicki and I. P. dell'Antonio, Astrophys. J. **427**, 603 (1994), astro-ph/9312006.

[50] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics, by George B. Rybicki, Alan P. Lightman, pp. 400. ISBN 0-471-82759-2. Wiley-VCH , June 1986.* (1986) p. 400.

[51] P. A. M. van Hoof, R. J. R. Williams, K. Volk, M. Chatzikos, G. J. Ferland, M. Lykins, R. L. Porter, and Y. Wang, Mon. Not. R. Astron. Soc. **444**, 420 (2014), arXiv:1407.5048 [astro-ph.HE].

[52] H. V. Cane, Mon. Not. R. Astron. Soc. **189**, 465 (1979).

[53] M. W. Eastwood, M. M. Anderson, R. M. Monroe, G. Hallinan, B. R. Barsdell, S. A. Bourke, M. A. Clark, S. W. Ellingson, J. Dowell, H. Garsden, L. J. Greenhill, J. M. Hartman, J. Kocz, T. J. W. Lazio, D. C. Price, F. K. Schinzel, G. B. Taylor, H. K. Vedantham, Y. Wang, and D. P. Woody, ArXiv e-prints (2017), arXiv:1711.00466 [astro-ph.IM].

[54] J. M. Cordes and T. J. W. Lazio, ArXiv Astrophysics e-prints (2002), astro-ph/0207156.

[55] A. Lewis and S. Bridle, Phys. Rev. D **66**, 103511 (2002), astro-ph/0205436.

[56] J. C. Novaco and L. W. Brown, Astrophys. J. **221**, 114 (1978).

[57] X. Chen and J. Miralda-Escudé, Astrophys. J. **602**, 1 (2004), astro-ph/0303395.

[58] R. Barkana, Nature **555**, 71 (2018).

Appendix A: Numerical computation of the Lyα spectrum

Here we describe in detail the numerical radiative transfer computation carried out in Section III A. We use the method of HF09. No claim of originality is made in this section, these details are included here only for the sake of completeness.

As stated above, the spectrum evolves due to Hubble drift, emissions, absorptions, and scatterings. We consider the Hubble effect first. In order to conserve photons, rather than writing a derivative operator for the Hubble drift, we chose the timestep of our Fokker-Planck solver such that $\Delta \ln \nu = \Delta \ln a$. Under this assumption, at each time step the photons in one bin simply shift to the next lower-frequency bin, i.e.

$$N_i(a + \Delta a) = N_{i+1}(a). \tag{A1}$$

We obtain the value for the bluemost bin from the HyRec spectrum. Because we chose such a large number of bins, this bin is sufficiently far from the Lyα line that the exact choice of its frequency does not affect our final results.

The emission rate of Lyα photons is set by sum of the decay rates $\Gamma_{u \to 2p}$ from higher-level states u to the $2p$ state multiplied by the branching fraction $P_{2p \to 1s}$ for $2p$ to decay to $1s$. The contribution $\dot{N}|_{\mathrm{em}}$ to the time derivative of N_i can be written as

$$\dot{N}_i|_{\mathrm{em}} = \sum_u x_u \Gamma_{u \to 2p} P_{2p \to 1s} \mathcal{E}(\nu_i) \phi_\alpha(\nu_i) \Delta \nu_i, \tag{A2}$$

where x_u is the fraction of hydrogen atoms in state u. Values for x_{1s} and x_{2p} are given by HyRec, populations for higher states are computed assuming thermal equilibrium. The quantity $\mathcal{E}(\nu)$ is a correction factor for the emission rate due to photon phase space factors and the presence of other resonances, its value is given by equations (25-26) of HF09. Once we have the emission rate, we can use detailed balance to write the absorption rate $\dot{N}_i|_{\mathrm{ab}}$ as

$$\frac{\dot{N}_i|_{\mathrm{ab}}}{\dot{N}_i|_{\mathrm{em}}} = -\frac{N_i}{N_{\mathrm{eq}}} \left(\frac{\nu_{\mathrm{Ly}\alpha}}{\nu_i} \right)^3 e^{h(\nu_i - \nu_{\mathrm{Ly}\alpha})/T_r}, \tag{A3}$$

where

$$N_{\mathrm{eq}} = \frac{8\pi \nu_{\mathrm{Ly}\alpha}^3}{c^3 n_{\mathrm{H}}} \frac{x_{2p}}{3 x_{1s}} \Delta \ln \nu \tag{A4}$$

is the equilibrium value of N_i at line center.

True absorptions and emissions change the overall number of photons in our calculations. We now need to add in the final effect, scattering, which preserves photon number but alters their frequency. If F_i is the net flux of photons from bin $i + 1$ to i due to scattering, then

$$\dot{N}_i|_{\mathrm{sc}} = F_i - F_{i-1}. \tag{A5}$$

Because we are working in the Fokker-Planck approximation, these fluxes are linear in N_i and $dN_i/d\nu$, so we can write

$$F_i = -\zeta_i N_i + \eta_i N_{i+1}. \tag{A6}$$

One relation between our new constants ζ_i and η_i can be determined by requiring the F_i be zero in thermal equilibrium, which yields

$$\frac{\zeta_i}{\eta_i} = \frac{\nu_{i+1}^3}{\nu_i^3} e^{-h(\nu_{i+1} - \nu_i)/T_m}. \tag{A7}$$

The matter temperature T_m appears in the exponent because scattering effects tend to bring the color temperature in to equilibrium with the kinetic temperature of the gas.

In terms of ζ_i and η_i, the diffusion coefficient can be written as

$$\mathcal{D}(\nu) = \frac{1}{2}(\eta_i + \zeta_{i-1})(\Delta \nu)^2. \tag{A8}$$

Comparing this to the actual value of $\mathcal{D}(\nu)$ from Hirata [47] and assuming that ζ_i and η_i are slowly varying functions yields the result that

$$\eta_i + \zeta_i = \frac{H \nu_{\mathrm{Ly}\alpha} \sigma_\nu^2 \tau_{\mathrm{Ly}\alpha} \left[\phi_\alpha(\nu_i) + \phi_\alpha(\nu_{i+1}) \right]}{(\nu_{i+1} - \nu_i)^2}, \tag{A9}$$

where $\tau_{\mathrm{Ly}\alpha}$ is the optical depth in the Lyα line. By combining equations (A7) and (A9), we can determine F_i and therefore $\dot{N}_i|_{\mathrm{sc}}$. Note that in this expression we have set the fraction of absorptions which lead to scattering equal to unity, which should be true to high precision. We assume that there is no scattering on either boundary of our calculation.

We solve our Fokker-Planck problem using a backwards Euler solver. We can write the full time derivative of N_i as

$$\dot{N}_i = \dot{N}_i|_{\mathrm{em}} + \dot{N}_i|_{\mathrm{ab}} + \dot{N}_i|_{\mathrm{sc}} = C_{ij} N_j. \tag{A10}$$

The matrix C_{ij} is tridiagonal, since the derivative in one bin depends only on the values in itself and those immediately adjacent. On each timestep, we solve

$$\frac{N_i(t + \Delta t) - N_i(t)}{\Delta t} = C_{ij}(t + \Delta t) N_j(t + \Delta t) \tag{A11}$$

using the standard tridiagonal method to obtain new values of N_i. We then shift each value redward by one bin, filling in the bluemost bin with the value obtained from HyRec.

Appendix B: Convergence Tests

The 21 cm signal we predict here is the result of a subpercent difference between the spin and radiation temperatures. It is therefore crucial that we make certain

that our numerical calculations are as accurate as possible. The most likely place for numerical errors to creep in is the Lyα radiative transfer computation described in Section III A. In this appendix, we will briefly describe the efforts we have made to demonstrate the convergence of this code.

We mentioned above that we use nearly 100 times higher spectral and temporal resolution for our code than was used in the similar work by HF09. This is required to ensure that the spectrum in both the core and wings of the Lyα line converges to the precision we require. If we use the bin sizes specified in HF09, we find that amplitude of the plateau in $1 - T_C/T_r$ from Figure 4 is higher by roughly a factor of two. In the lower resolution calculation, the slope of the blue side spectrum is slightly steeper than that plotted in Figure 3. This drives the color temperature further out of equilibrium than it otherwise should be. It is unclear whether this discrepancy is sourced by insufficient temporal resolution or insufficient spectral resolution, as our prescription for Hubble redshifting requires us to keep the two step sizes equal. We selected the final resolution used in the body of the paper by gradually increasing the resolution until the curves plotted in Figure 4 converged. At our full resolution, halving the time and frequency steps changes the values of those curves by <1%.

Similarly, we use nearly 100 times more frequency bins than HF09 in our calculation. At fixed frequency step, the number of bins sets the total width in frequency space of our high-precision computation. Because the wings of the Lyα transition allow for significant scattering of photons well outside the line center, it is necessary to use quite wide boundary conditions. With the HF09 bin counts, small differences between the HyRec spectrum and the correct value contaminate our color temperature. We selected our final bin counts using the same criterion as for the spectral resolution.

It is also possible that there could be some error introduced by our choice of high-redshift boundary condition. We use a blackbody initial spectrum, but is unlikely that the spectrum at our initial redshift is an exact blackbody. However, because the optical depth to Lyα scatterings is so high, any effects caused by this choice are quickly damped away. For example, if we start our calculation at $z = 1500$ instead of 1600 the value of $1 - T_C/T_r$ converges to that plotted in Figure 4 to within <0.1% by $z = 1490$. We can therefore be fairly certain that the results presented above are independent of our choice of initial condition.