Evolution of the 21 cm signal throughout cosmic history

Jonathan R. Pritchard* and Abraham Loeb

Institute for Theory & Computation, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street,

Cambridge, Massachusetts 02138, USA

(Received 14 February 2008; published 13 November 2008)

The potential use of the redshifted 21 cm line from neutral hydrogen for probing the epoch of reionization is motivating the construction of several low-frequency interferometers. There is also much interest in the possibility of constraining the initial conditions from inflation and the nature of the dark matter and dark energy by probing the power spectrum of density perturbations in three dimensions and on smaller scales than probed by the microwave background anisotropies. Theoretical understanding of the 21 cm signal has been fragmented into different regimes of physical interest. In this paper, we make the first attempt to describe the full redshift evolution of the 21 cm signal between 0 < z < 300. We include contributions to the 21 cm signal from fluctuations in the gas density, temperature, and neutral fraction, as well as the Ly α flux, and allow for a post-reionization signal from damped Ly α systems. Our comprehensive analysis provides a useful foundation for optimizing the design of future arrays whose goal is to separate the particle physics from the astrophysics, either by probing the peculiar velocity distortion of the 21 cm power spectrum, or by extending the 21 cm horizon to $z \ge 25$ before the first galaxies had formed, or to $z \le 6$ when the residual pockets of hydrogen trace large-scale structure.

DOI: 10.1103/PhysRevD.78.103511

PACS numbers: 98.80.-k, 95.35.+d, 98.62.-g

I. INTRODUCTION

Modern cosmology has advanced considerably in recent years due to precise measurements of the cosmic microwave background (CMB) and large-scale structure [1]. These data sets probe the distribution of matter at redshifts $z \sim 10^3$ and $z \leq 0.3$, respectively, yet only cover $\sim 0.1\%$ of the entire comoving volume of the observable Universe [2]. In addition, significant information about individual galaxies and quasars out to z = 7 has been obtained through the efforts of many telescopes at multiple wavelengths [3]. Nonetheless, our knowledge and understanding of the evolution of most of the baryonic matter, which resided in the intergalactic medium (IGM), is limited. Constraints from the 3rd year Wilkinson Microwave Anisotropy Probe (WMAP) [4] on the optical depth to the CMB ($\tau = 0.093 \pm 0.029$) only weakly constrain the redshift of reionization. The presence of a Gunn-Peterson trough [5] in quasar spectra at z > 6.5 provides a loose limit on the end of reionization [6]. Absorption by the Ly α forest in quasar spectra constrain the IGM temperature to be $T_K = 2 \times 10^4$ K at $z \sim 4$ [7]. Beyond these fragmentary data points lies a complicated story of ionization and heating of the IGM going back to the formation of the first stars [8,9].

Low-frequency observations of the redshifted 21 cm line of neutral hydrogen present one of the most promising avenues for exploring the entirety of the poorly constrained period between recombination and reionization [10,11]. Additionally, mapping hydrogen throughout this period and at lower redshifts raises the prospect of measuring primordial density fluctuations over a much larger volume than that probed by the CMB or galaxy redshift surveys [12,13]. Several experiments are currently being constructed (such as MWA [14], LOFAR [15], PAPER [16], 21CMA [17]), and more ambitious designs are being planned (SKA [18]) to detect the theoretically-predicted 21 cm signal.

The volume to be mapped can be loosely divided into three parts: (i) the "Dark Ages" between the epoch of cosmic recombination and the appearance of the first galaxies $(25 \le z \le 10^3)$; (ii) the epoch of reionization (EoR) during which hydrogen was ionized¹ by the UV radiation from galaxies ($6 \le z \le 25$); and (iii) the post-reionization epoch ($z \leq 6$). The transition between the first two epochs, characterized by the appearance of the first luminous sources, is a period of "cosmic twilight," during which the radiation fields of the sources begins to affect the gas, heating and exciting the largely neutral intergalactic medium (IGM). Eventually, the small ionized regions surrounding groups of sources expand and percolate beginning the process of reionization. After reionization, the residual neutral hydrogen is found in dense clumps (such as damped Ly α systems [21]), self-shielded from ionizing radiation.

Precise measurements of the power-spectrum of 21 cm brightness fluctuations could constrain the initial conditions from inflation [12] (including deviations from scaleinvariance or Gaussianity), as well as the nature of the dark

^{*}Hubble Fellow;

jpritchard@cfa.harvard.edu

¹However, exotic physics such as the decay of dark matter [19] or evaporation of primordial black holes [20] might alter this picture dramatically.

matter (including a measurement of the neutrino mass in the range expected from atmospheric neutrino experiments [2]) and dark energy [22]. The 21 cm fluctuations are expected to simply trace the primordial power spectrum of matter density perturbations either before the first galaxies had formed ($z \ge 25$) [12,23] or after reionization ($z \le 6$)—when only dense pockets of self-shielded hydrogen survive [24,25]. During the epoch of reionization, the fluctuations are mainly shaped by the topology of ionized regions [26–28], and thus depend on astrophysical details. Nevertheless, even during this epoch, the line-of-sight anisotropy of the 21 cm power spectrum due to peculiar velocities, can in principle be used to separate the implications for fundamental physics from the unknown details of the astrophysics [26,29].

Recent work has explored the 21 cm signal that may arise during reionization [30,31] from variations in the Ly α flux [32–34], x-ray heating [35], gas density [12,23], or peculiar velocities [26,29,36]. In addition, the post-reionization signal has been explored by [24]. In general, previous studies have focused on individual regions of parameter space where only one of two of these fluctuations needs to be considered, and the rest are assumed to be negligible. Here, we attempt for the first time to present a global picture of the redshift evolution of the 21 cm signal, including all relevant fluctuations. The full history can be used to select particular redshift (frequency) windows and observing strategies within which fundamental physics would be optimally constrained.

In analyzing the full history of 21 cm fluctuations, we seek to address two main issues. First, we would like to identify the redshift at which the presence of astrophysical sources begin to modify the pristine cosmological signal. This defines the minimum redshift that must be accessible for a future observatory aiming to detect directly the pristine signal (in analogy with CMB experiments). Although the answer is model dependent, we will show that the rapidity with which structures form once nonlinear collapse gets underway means that the model dependence is weak. The weak dependence we find is generic to a Gaussian random field of initial density perturbations, as expected from inflation, for which the fraction of mass incorporated in collapsed objects grows exponentially (on the high-density tail of the Gaussian) at early times. Second, we wish to address the possibility of using the angular separation proposed by [29] to recover cosmological information from 21 cm observations after astrophysical effects become large. We will show that there exists a window in redshift where this technique might be used to extract the dark-matter power spectrum despite the complication of other sources of fluctuations. Finally, we would like to identify the primary astrophysical parameters that control the 21 cm signal at subsequent epochs.

The paper is organized as follows: In Sec. II, we discuss the calculation of the 21 cm signal, attempting to bring together a variety of different sources of fluctuations. The results of these calculations are illustrated in Sec. III. In Sec. IV, we discuss the experimental requirements for detecting the 21 cm signal and explore the performance of three fiducial experiments. Finally, we conclude in Sec. V and discuss future prospects.

Throughout this paper, we assume a cosmology with $\Omega_m = 0.26$, $\Omega_{\Lambda} = 0.74$, $\Omega_b = 0.044$, $H_0 = 100h \,\mathrm{km \, s^{-1} \, Mpc^{-1}}$, with h = 0.74, $n_S = 0.95$, and $\sigma_8 = 0.8$, consistent with the latest measurements [4].

II. THEORETICAL FRAMEWORK

A. 21 cm brightness temperature

We begin by briefly summarizing the physics of the 21 cm signal and refer the interested reader to Furlanetto, Oh, and Briggs [10] for further details. The 21 cm line of the hydrogen atom results from hyperfine splitting of the 1*S* ground state due to the interaction of the magnetic moments of the proton and the electron. The HI spin temperature T_S is defined through the ratio between the number densities of hydrogen atoms in the 1*S* triplet and 1*S* singlet levels, $n_1/n_0 = (g_1/g_0) \exp(-T_*/T_S)$, where $(g_1/g_0) = 3$ is the ratio of the spin degeneracy factors of the two levels, and $T_* \equiv hc/k\lambda_{21 \text{ cm}} = 0.068 \text{ K}$. The optical depth of this transition is small at all relevant redshifts, yielding brightness temperature fluctuations

$$T_b = 27 x_{\rm HI} \left(1 + \frac{4}{3} \delta_b \right) \left(\frac{\Omega_b h^2}{0.023} \right) \left(\frac{0.15}{\Omega_m h^2} \frac{1+z}{10} \right)^{1/2} \times \left(\frac{T_s - T_\gamma}{T_s} \right) \, \text{mK}, \tag{1}$$

Here, $x_{\rm HI}$ is the neutral fraction of hydrogen, δ_b is the fractional overdensity in baryons, and the factor of 4/3 arises from including the effect of peculiar velocities and averaging the signal over angles. The spin temperature is given by

$$T_{S}^{-1} = \frac{T_{\gamma}^{-1} + x_{\alpha}T_{\alpha}^{-1} + x_{c}T_{K}^{-1}}{1 + x_{\alpha} + x_{c}},$$
 (2)

where T_{α} is the color temperature of the Ly α radiation field at the Ly α frequency and is closely coupled to T_K by recoil during repeated scattering, and x_c , x_{α} are the coupling coefficients due to atomic collisions and scattering of Ly α photons, respectively. The spin temperature becomes strongly coupled to the gas temperature when $x_{\text{tot}} \equiv x_c + x_{\alpha} \gtrsim 1$.

The collisional coupling coefficient is given by

$$x_{c} = x_{c}^{HH} + x_{c}^{eH} = \frac{4T_{\star}}{3A_{10}T_{\gamma}} [\kappa_{1-0}^{HH}(T_{k})n_{H} + \kappa_{1-0}^{eH}(T_{k})n_{e}],$$
(3)

where $A_{10} = 2.85 \times 10^{-15} \text{ s}^{-1}$ is the spontaneous emission coefficient, the scattering rate between hydrogen

atoms κ_{1-0}^{HH} is tabulated as a function of T_k [37,38], and the scattering rate between electrons and hydrogen atoms κ_{1-0}^{eH} is taken from [39]. For a more detailed analysis of the collisional coupling, see [40].

The Wouthysen-Field effect [41,42] coupling is given by

$$x_{\alpha} = \frac{16\pi^2 T_{\star} e^2 f_{\alpha}}{27A_{10} T_{\chi} m_{e} c} S_{\alpha} J_{\alpha},\tag{4}$$

where $f_{\alpha} = 0.4162$ is the oscillator strength of the Ly α transition. S_{α} is a correction factor of order unity, which describes the detailed structure of the photon distribution in the neighborhood of the Ly α resonance [43–46]. We make use of the approximation for S_{α} outlined in [46]. For the models considered in this paper, Ly α coupling dominates over collisional coupling at redshifts $z \leq 25$, although we consider this more carefully later.

Fluctuations in the 21 cm signal may be expanded to linear order [10]

$$\delta_{T_b} = \beta_b \delta_b + \beta_x \delta_x + \beta_\alpha \delta_\alpha + \beta_T \delta_T - \delta_v, \quad (5)$$

where each δ_i describes the fractional variation in the quantity *i*, and we include fluctuations in the baryon density (b), neutral fraction (*x*), Ly α coupling coefficient (α), gas temperature (*T*), and line-of-sight peculiar velocity gradient (ν). The expansion coefficients are given by

$$\beta_{b} = 1 + \frac{x_{c}}{x_{tot}(1 + x_{tot})},$$

$$\beta_{x} = 1 + \frac{x_{c}^{HH} - x_{c}^{eH}}{x_{tot}(1 + x_{tot})},$$

$$\beta_{\alpha} = \frac{x_{\alpha}}{x_{tot}(1 + x_{tot})},$$

$$\beta_{T} = \frac{T_{\gamma}}{T_{K} - T\gamma} + \frac{1}{x_{tot}(1 + x_{tot})} \left(x_{c}^{eH} \frac{d\log \kappa_{10}^{eH}}{d\log T_{K}} + x_{c}^{HH} \frac{d\log \kappa_{10}^{HH}}{d\log T_{K}} \right).$$
(6)

It is important to realize that fluctuations in x_H can be of order unity leading to quartic terms involving the ionization field contributing to the brightness temperature power spectrum.

Noting that in Fourier space $\delta_{\partial v} = -\mu^2 \delta$ [36], where μ is the angle between the line of sight and the wavevector **k** of the Fourier mode, we may use Eq. (5) to form the power spectrum

$$P_{T_b}(k, \mu) = P_{bb} + P_{xx} + P_{\alpha\alpha} + P_{TT} + 2P_{bx} + 2P_{b\alpha}$$
$$+ 2P_{bT} + 2P_{x\alpha} + 2P_{xT} + 2P_{\alpha T} + P_{x\delta x\delta}$$
$$+ \text{ other quartic terms} + 2\mu^2(P_{b\delta} + P_{x\delta}$$
$$+ P_{\alpha\delta} + P_{T\delta}) + \mu^4 P_{\delta\delta} + 2P_{x\delta\delta_v x}$$
$$+ P_{x\delta_v\delta_v x} + \text{ other quartic terms with } \delta_v.$$
(7)

Here, we note that all quartic terms must be quadratic in x_H and separate them depending upon whether they contain powers of δ_v or not. Those that contain powers of δ_v will not be isotropic and so contribute to the angular dependence of P_{T_b} (see [26] for further discussion).

We may rewrite Eq. (7) in more compact form

$$P_{T_b}(k,\mu) = P_{\mu^0}(k) + \mu^2 P_{\mu^2}(k) + \mu^4 P_{\mu^4}(k) + P_{f(k,\mu)}(k,\mu),$$
(8)

where we have grouped those quartic terms with anomalous μ dependence into the term $P_{f(k,\mu)}(k, \mu)$. In principle, high-precision measurements of the 3D power spectrum will allow the separation of $P_{T_b}(k, \mu)$ into these four terms by their angular dependence on powers of μ^2 [29]. The contribution of the $P_{f(k,\mu)}(k, \mu)$ term, with its more complicated angular dependence, threatens this decomposition [26]. Since this term is only important during the final stages of reionization, we will ignore it in this paper noting only that the angular decomposition by powers of μ^2 may not be possible when ionization fluctuations are important.

It is unclear whether the first generation of 21 cm experiments will be able to achieve the high signal to noise required for this separation [26]. Instead, they might measure the angle-averaged quantity

$$\bar{P}_{T_b}(k) = P_{\mu^0}(k) + P_{\mu^2}(k)/3 + P_{\mu^4}(k)/5 \qquad (9)$$

(where we neglect the $P_{f(k,\mu)}(k, \mu)$ term). In presenting our results, we will concentrate on $\bar{P}_{T_b}(k)$, which is easiest to observe and discuss the angular separation separately in Sec. III C. We will typically plot the power per logarithmic interval $\Delta = [k^3 P(k)/(2\pi^2)]^{1/2}$.

B. Astrophysical modeling

Before calculating the evolution of the fluctuations, we must first provide an astrophysical model for the sources of radiation that modify the 21 cm signal. We follow the model of [47] with some minor modifications.

We must specify a luminosity and spectrum for sources of ionizing photons, Lyman series photons, and x-rays. The variation in spectra has relatively little effect on the fluctuations (but see [33,35] for further discussion on this point), so that for our purposes we will fix the spectrum and consider models with different luminosities. We specify three key parameters for our astrophysical sources: the number of ionizing photons produced per baryon in stars that contribute to ionizing the IGM $N_{\text{ion,IGM}} = f_{\text{esc}}N_{\text{ion}}$, the number of Ly α photons produced per baryon in stars N_{α} , and the energy in x-rays produced per baryon in stars ϵ_X . With this parametrization, we are separating the physics of the sources from the star-formation history, as the quantity of physical relevance will be the above numbers multiplied by the efficiency of star formation f_* . For example, the conventional definition of the ionizing efficiency $\zeta =$

For convenience, we further define f_{α} and f_X via $N_{\alpha} =$ $f_{\alpha}N_{\alpha,\text{ref}}$ and $\epsilon_X = f_X\epsilon_{X,\text{ref}}$, where we take the reference values appropriate for normal (so-called, Population II) stars $N_{\alpha,\text{ref}} = 6590$ and for starburst galaxies with a power-law spectrum of index $\alpha_s = 1.5$ and $\epsilon_{X,ref} =$ 560 eV [48]. We use the spectra appropriate for these sources throughout. For comparison, in this notation, the very massive (*Population III*) stars have [49], $N_{\alpha} = 3030$ $(f_{\alpha} = 0.46)$, when the contribution from higher Lyman series photons is included. We expect the value of f_{α} to be close to unity. In contrast, f_X is relatively unconstrained with values $f_X \leq 10^3$ possible without violating the WMAP CMB optical depth constraints. Constraining this parameter will mark a step forward in our understanding of the thermal history of the IGM and the population of x-ray sources at high redshifts.

Once reionization is complete, we assume that photoionization heats the gas to 3×10^4 K. This is needed only for determining the thermal cutoff scale for the baryonic fluctuations for the post-reionization 21 cm signal.

C. Calculation of fluctuations

The focus of this paper is to combine different sources of 21 cm fluctuation into a complete picture of the evolution of the 21 cm signal from the epoch when the cosmic gas thermally decoupled from the CMB ($z \sim 200$) to the present time. In calculating the various sources of fluctuations, we draw heavily upon the existing literature. Below, we briefly discuss the relevant calculations and then describe how we piece them together.

The 21 cm signal has been most extensively studied within the EoR, where it is most accessible to observations. During this epoch large-scale (Mpc) ionized (HII) regions grow around clusters of sources and delineate a density contrast of order unity in the hydrogen abundance, far greater than the underlying fluctuations in the matter density on the same scales [50]. Analytic models of the HII bubble size and resulting 21 cm correlations were formulated by Furlanetto et al.[31], who considered bubbles to self-ionize once they contained a sufficient mass fraction in galaxies (ionizing sources). This basic formalism has been applied to numerical simulations [51] for more detailed calculations of the ionization power spectrum. Complementary simulations by Trac and Cen [52] have been compared directly with the Furlanetto *et al.* [31] model showing reasonable agreement [27]. For computational ease, we choose to use the fully analytic calculation of [31] for the ionization fluctuations, including a correction for an error in the bubble bias as noted by McQuinn et al. [53]. This prescription allows us to calculate P_{xx} and $P_{x\delta}$. The key parameter in the Furlanetto *et al.* [31] model is the ionizing efficiency of sources ζ ; we extract this redshift-dependent quantity from our astrophysical model by relating the ionized fraction x_i to the total collapse fraction f_{coll} through $x_i = \zeta f_{\text{coll}}$.

Fluctuations in the Ly α coupling originate from spatial variation of the Ly α flux, which arise primarily due to the strong clustering of early sources. The calculation of these fluctuations has been discussed in detail by Barkana and Loeb [32] and Pritchard and Furlanetto [34]. We consider fluctuations from Ly α photons arising from both stellar and x-ray sources [33,54,55]. Since the spectrum of stellar and x-ray sources will in general be different and as the mean free path of x-ray photons is very different from that of Lyman series photons the radial profile of Ly α photons about stellar and x-ray sources is expected to very different. This in turn will lead to different scale dependence in the resulting Ly α fluctuations. It is therefore important to consider both contributions to the Ly α flux.

We follow the model of Barkana and Loeb [32] in calculating δ_{α} by integrating the flux from all sources weighted appropriately by distance and in addition account for higher Lyman series photons [34,44]. Recent investigations have shown that including the scattering of Ly α photons in the IGM near to sources redistributes the Ly α flux leading to greater power on large scales [56–58]. We neglect this effect estimating that this would change our results at the 10–20% level when Ly α fluctuations dominate.

Temperature fluctuations arise initially through the competition between adiabatic cooling and Compton scattering on the CMB (which drives the gas toward being isothermal). After thermal decoupling at $z \sim 200$, the former process dominates and temperature fluctuations grow since the rate of cooling is density dependent. Once starformation begins, x-ray heating drives large temperature variation due to its highly inhomogeneous nature. We calculate the evolution of temperature fluctuations using the formalism of Pritchard and Furlanetto [35], including all three processes. Temperature fluctuations may also arise as a result of exotic heating mechanisms such as the decay of dark matter [19]. In this paper, we are interested in producing a "standard" picture of the 21 cm signal and will leave exotic deviations to future work.

Before star-formation, 21 cm fluctuations are determined by simple atomic physics and so, potentially, provide a window into the physics of inflation, via precise measurements of the spectrum of density fluctuations [12]. Detailed numerical work [23,59] has shown this period to be rich in physics as recombination, thermal coupling, gas pressure, and other physical processes interact to determine the final 21 cm signal. Since Compton drag on the CMB distinguishes between baryons and dark matter, detailed calculation must track the evolution of both components separately [60]. After thermal decoupling, however, baryons fall into the potential wells created by the dark matter and by $z \leq 30$. It is a good approximation to take the baryons as tracking the dark matter, although this ignores the presence of acoustic oscillations in the baryon component and pressure smoothing on small scales. Baryon acoustic oscillations represent a major target for galaxy surveys and 21 cm observations, since they help to constrain dark energy via the angular-diameter-distance redshift relation [61]. 21 cm observations of baryon oscillations complement galaxy surveys well [22,25,60], since they cover a higher redshift range and have a different set of biases than galaxy surveys, which may be affected by inhomogeneous galaxy formation [62].

In our calculations, we assume that baryons trace the underlying density field on large scales but are smoothed by the finite pressure of the gas on small scales [63] so that $\delta_b = \delta (1 + k^2 R_F^2)^{-1}$, where $R_F = 1/k_F$ and k_F is the filtering scale [64]. Thermal broadening introduces a second smoothing scale R_T [32], which we incorporate through a Gaussian cutoff factor $\exp(-k^2 R_T^2)$. Until x-ray heating becomes important $R_F > R_T$, while afterwards $R_T > R_F$ as the filter mass takes a finite period of time to respond to the greatly increased gas temperature. These smoothing scales are small (typically $R_T \sim R_F \sim \text{kpc}$) until photoionization heats the gas to $T_K \sim 10^4 \ {\rm K}$ and smoothing becomes important for wave numbers $k \ge$ 10 Mpc^{-1} . We calculate linear density fluctuations using the fitting function of [65]. Nonlinear density effects become important even at high redshift, and we incorporate these using the halo model [66].

After reionization is complete, a few percent of all neutral hydrogen (by mass) remain in over dense clumps that are self-shielded ($\Omega_{\rm HI} = 10^{-3}$). These clumps are observed as high-column density absorbers in quasar spectra in the form of Lyman-limit or damped $Ly\alpha$ systems [21]. Fluctuations in the 21 cm brightness from this postreionization epoch will trace the density fluctuations with a nearly constant bias [24]. Following the calculations of Wyithe and Loeb [24], we take the post-reionization signal to be $P_{21}(k) = b_{21}^2 P_{\delta\delta}(k)$ with $b_{21} = 0.03$ K. Although the fluctuations here are much weaker than before reionization, the vastly diminished foreground at the appropriate higher frequencies makes observations feasible. Observation of fluctuations in this regime may be used to detect baryon acoustic oscillations in the matter power spectrum and constrain the equation of state of the dark energy at higher redshifts than those accessible to supernova or galaxy surveys [22].

The existence of numerous cross terms in Eq. (7) connecting different fluctuations requires some thought. Following the approach of Barkana and Loeb [32], we relate δ_{α} to the baryon density field using a scaledependent linear function $\delta_{\alpha} = W_{\alpha}(k)\delta_{b}$. Similarly, we set $\delta_{T} = g_{T}(k)\delta_{b}$ [35]. In this way, we are able to relate the cross terms $P_{\alpha T}$, P_{bT} , $P_{b\alpha}$, and equivalent δ terms to $P_{\delta\delta}$. Since the Ly α flux, gas temperature, and baryon density do not interact directly with one another these cross terms should be a good approximation.

The treatment of the neutral fraction fluctuations is more complicated. Once HII bubbles begin to form, we are faced with a two-phase medium composed of almost fully ionized HII regions and the largely neutral IGM outside. We must distinguish between x_i , the fraction of the cosmic gas incorporated in fully ionized bubbles, and x_e the ionized fraction in the largely neutral IGM (which could be increased beyond its primordial value due to diffuse x-rays). Neutral fraction fluctuations arise from both, although in most models that we consider x_e remains small at all times and so fluctuations in x_e are also small. The calculations of δ_{α} and δ_{T} by Barkana and Loeb [32] and Pritchard and Furlanetto [35] assume that the filling fraction of HII regions is negligible. However, here we must relax that assumption in order to calculate terms like $P_{x\alpha}$. Since there is essentially no 21 cm signal from the ionized HII regions, the HII regions act to mask the fluctuations in α and T. Overdense regions will host sources of $Ly\alpha$ and x-ray photons, as well as ionizing photons, suggesting that cross terms of the form $P_{x\alpha}$ and P_{xT} may be complicated. We show later that these cross terms are subdominant, so that, for a first approximation, we may neglect these complications and model the cross terms as $P_{x\alpha} = W_{\alpha}(k)P_{xb}$, $P_{xT} = g_T(k)P_{xb}$, and so on. Detailed calculation will require numerical simulation of the ionization and radiation fields to determine the cross-terms.

We apply the model of Furlanetto *et al.* [31] to calculate P_{xx} , $P_{x\delta}$, and $P_{x\delta x\delta}$. Note that this model first calculates the correlation functions and Fourier transforms. For the quartic term, we ignore the possible connected part so that $P_{x\delta x\delta} = P_{xx}P_{\delta\delta} + 2P_{x\delta}^2$. Quartic terms involving δ_v contain angular dependence that threatens the possibility of the μ^2 decomposition [26]. The other quartic terms we handle in the same way as terms like $P_{x\alpha}$, first reducing them to a quartic of δ and X multiplied by factors of W_{α} and g_T , then calculating them in the above way. These are generally small corrections, relevant only at the end of reionization.

III. RESULTS

A. Mean history

We choose three astrophysical models chosen to reproduce the 3rd year WMAP central and $\pm 1 - \sigma$ for the cumulative electron-scattering optical depth τ . Model A uses ($N_{\text{ion,IGM}}$, f_{α} , f_{X} , f_{*}) = (200, 1, 1, 0.1) giving z_{reion} = 6.47 and τ = 0.063. Model B uses ($N_{\text{ion,IGM}}$, f_{α} , f_{X} , f_{*}) = (600, 1, 0.1, 0.2) giving z_{reion} = 9.76 and τ = 0.094. Model C uses ($N_{\text{ion,IGM}}$, f_{α} , f_{X} , f_{*}) = (3000, 0.46, 1, 0.15) giving z_{reion} = 11.76 and τ = 0.115. The evolution of the mean temperatures with redshift for these models is shown in Fig. 1. The top axis shows the relevant observing frequency range given by $\nu_{\text{obs}} = \nu_{21 \text{ cm}}/(1 + z)$. Since we are leaving the cosmology unchanged between models, the mean history only varies once star formation becomes important around $z \approx 25$. The most significant difference between these models is the redshift of reionization that cuts off the 21 cm signal at low redshift and is most responsible for the different τ values. Model C has a significantly decreased Ly α flux, which shows itself through a decrease in the strength of the absorption signal.

The global evolution of the 21 cm signal is itself the target of several experiments, e.g. the "Cosmological Reionization Experiment" (CoRE) [67] and the "Experiment to Detect the Global EOR Signature" (EDGES) [68]. Although such an observation is conceptually simple, it is experimentally challenging. Distinguishing between a global signal and other sources of all sky emission, including galactic synchrotron, freefree radiation, and the CMB is difficult. Experimental detection relies upon reionization proceeding rapidly leading to a distinctive steplike feature in the frequency direction, which would not be expected to be produced by the spectrally smooth foregrounds. With the assumption of sharp reionization, EDGES [68] places an initial constraint that $\bar{T}_b < 450$ mK at z = 8. While this is far from the expected signal amplitude, such constraints will improve with time. Efforts are also underway to extend the frequency coverage to $\nu \approx 50$ MHz to access the transition from an absorption to emission signal. As Fig. 1 indicates, this transition is likely to be significantly larger in amplitude (~ 100 mK) than that at the end of reionization (\sim 20 mK).

B. Fluctuation history

The three-dimensional nature of the 21 cm signal makes it difficult to convey the evolution of the fluctuations with a single two-dimensional plot. We therefore plot the evolution of four individual comoving wavenumbers k = 0.01, 0.1, 1, and 10 Mpc^{-1} , spanning the range that might be observed. On large scales we expect contamination from foregrounds to limit the detection of the power spectrum. On small scales thermal broadening of the 21 cm line will smooth the signal. It is also to be expected that many of our approximations will break down as small-scale information about the sources becomes important (see for example [69] for the importance of higher order correlations on small scales during reionization). For the mean histories shown in Fig. 1, we calculate the evolution of the 21 cm angle-averaged power spectrum, which is plotted in Figs. 2–4, for models A, B, and C, respectively.

The evolution of Δ_{T_b} clearly shows three regimes: the post-reionization regime at low redshifts ($z < z_{reion}$) where the 21 cm fluctuations from residual hydrogen follow the matter power spectrum, an intermediate redshift regime ($x_{reion} < z < z_{trans}$), where Ly α coupling produces a large signal and complicated astrophysics leads to significant scale dependence, and a high redshift collisionally coupled regime where 21 cm fluctuations track the density field ($z > z_{trans} \approx 23$). For pedagogical purposes, let us describe the evolution on a single comoving scale (say, k =



FIG. 1. *Top panel:* Evolution of the CMB temperature T_{CMB} (dotted curve), the gas kinetic temperature T_K (dashed curve), and the spin temperature T_S (solid curve). *Middle panel:* Evolution of the gas fraction in ionized regions x_i (solid curve) and the ionized fraction outside these regions (due to diffuse x-rays) x_e (dotted curve). *Bottom panel:* Evolution of mean 21 cm brightness temperature T_b . In each panel we plot curves for Model A (thin curves), Model B (medium curves), and Model C (thick curves).



FIG. 2. Redshift evolution of the angle-averaged 21 cm power spectrum $\bar{\Delta}_{T_b}$ for Model A at k = 0.01 (solid curve), 0.1 (dotted curve), 1.0 (short-dashed curve), and 10.0 (long-dashed curve) Mpc⁻¹. Reionization at z = 6.5.



FIG. 3. Redshift evolution of the angle-averaged 21 cm power spectrum $\bar{\Delta}_{T_b}$ for Model B. Reionization at z = 9.8. Same line conventions as Fig. 2.

0.1 Mpc⁻¹) and draw attention to the main features. Thermal decoupling at $z \sim 200$ is a gradual process and, initially, $\bar{\Delta}_{T_b}$ grows due to a combination of the growth of



FIG. 4. Redshift evolution of the angle-averaged 21 cm power spectrum $\bar{\Delta}_{T_b}$ for Model C. Reionization at z = 11.8. Same line conventions as Fig. 2.

density fluctuations and the steady gas cooling below $T_{\rm CMB}$. As the gas rarifies and cools, collisional coupling becomes less effective and, at $z \sim 60$, $\bar{\Delta}_{T_b}$ begins to decrease in amplitude. Note that the continuing growth of structure offsets the turnover on Δ_{T_h} from the minimum of T_b , seen in Fig. 1 to occur at $z \approx 90$. As collisional coupling diminishes, the signal drops toward zero. This occurs while $T_K < 30$, a regime where $\kappa_{1-0}(T_K)$ drops exponentially with T_K [38] and results in a rapid drop of the signal at $z \leq 40$. Before the signal drops all the way to zero, significant star formation occurs, and the resultant Ly α production leads to the beginning of Ly α coupling by $z \approx 25$. The exponential increase in the global starformation rate at these redshifts is responsible for the rapid increase in T_b and $\bar{\Delta}_{T_b}$. With this rise in signal, we enter into a regime dominated by astrophysics and begin to see complicated scale dependence.

Initially, $Ly\alpha$ fluctuations boost the signal somewhat above the level of density fluctuations alone. However, xray heating follows not far behind and contributes to $\bar{\Delta}_{T_h}$ with the opposite sign (hotter regions produce a weaker absorption signal, see Pritchard and Furlanetto [35]). In this competition, x-ray driven temperature fluctuations dominate causing Δ_{T_h} to pass through a zero point (seen as a sharp dip at $z \sim 18$ in all three plots). Temperature fluctuations dominate as \bar{T}_K approaches $T_{\rm CMB}$ and \bar{T}_b vanishes. In proceeding to the emission regime, we note based on Fig. 1 that the brightness fluctuations emitted T_b are generically smaller than they were during the absorption regime, leading to a decreasing trend in Δ_{T_h} . As reionization gets underway, ionization initially causes $\Delta_{T_{k}}$ to drop leading to a pronounced dip in its evolution. This occurs as a result of the clustering of ionizing sources in over dense regions causing the ionized HII regions to "mask out" those dense regions that have been producing the strongest 21 cm signal. As reionization proceeds, the contrast between ionized and neutral regions comes to dominate and $\bar{\Delta}_{T_h}$, rises until $x_H \sim 0.5$ after which the contrast begins to drop.

Toward the end of reionization the signal drops sharply as very little gas is left neutral. The post-reionization signal grows slowly as the density field grows. Since by this time the gas is photo heated to $T_K \approx 30\,000$ K the thermal width of the 21 cm line is sufficient to smooth out the signal on wavenumbers $k \gtrsim 10$ Mpc⁻¹. This cutoff potentially acts as a thermometer of the gas after reionization giving information about the temperature of gas contained in dense clumps.

As a result of the interplay between the radiation fields, as $\bar{\Delta}_{T_b}$ evolves it shows three peaks within the astrophysics-dominated regime. An important feature of this complicated evolution is that the maximum amplitude of $\bar{\Delta}_{T_b}$ occurs at different *k* values for different redshifts. Accurate observation and modeling of this complicated evolution may provide important information about the early radiation fields.

It is helpful to get a sense of how the amplitude of the signal compares with galactic foregrounds. We take the sky noise to be $T_{\rm sky} \approx 180 \text{ K}(\nu/180 \text{ MHz})^{-2.6}$ (appropriate for galactic synchrotron emission [10]), noting that the normalization depends upon the region of sky being surveyed. In Figs. 2–4, we plot $rT_{\rm sky}(\nu)$ where *r* ranges from $10^{-4}-10^{-9}$. We see that reducing foregrounds by a factor of $\sim 10^{-5}$ is required to observe fluctuations during reionization and cosmic twilight. The difficulty increases if reionization occurs early, which has the effect of compressing the signal at high redshifts (Model C). The signal from astrophysics in these three models begins at $\nu \approx 60 \text{ MHz}$ and continues to $\nu \approx 150 \text{ MHz}$, although this upper limit is very sensitive to the details of reionization.

Removing foregrounds to the rather low level of $\sim 10^{-7}$ is required to have a hope of observing the Dark Ages in the range z = 30-50 (25 MHz $\leq \nu \leq 60$ MHz). Some hope is restored by noticing that the rapid evolution of $\bar{\Delta}_{T_h}$ between z = 30-50 mimics the strong frequency dependence of the foregrounds. If one can remove foregrounds to the level needed to observe z = 30, then the same level of foreground subtraction would allow z = 50 to be observed. Beyond z = 50, the amplitude of $\overline{\Delta}_{T_h}$ decreases rapidly and becomes progressively more difficult to observe. Modifying the thermal history between z = 30 and recombination by introducing exotic physics can greatly change this high redshift behavior. Energy injection by decaying dark matter [19] can increase the residual electron fraction in the IGM, while maintaining collisional coupling to lower redshifts boosting the high redshift 21 cm signal and improving the chance of observing the Dark Ages.

We end this section by pausing to discuss the robustness of our modeling and highlight some of the important assumptions. Fluctuations well before star formation fall well within the linear regime and may be calculated with the same level of reliability as calculations of the CMB anisotropy (although see [23] for a more detailed calculation). At redshifts after reionization, we have assumed a simplistic prescription by using a constant bias factor. Although generally correct on large scales, the details of UV ionization of neutral hydrogen in dense clumps will cause corrections on small scales.

More subject to uncertainty is our modeling of fluctuations arising from the effects of luminous sources. One major assumption in our modeling has been that the starformation rate and hence the emissivity of the sources tracks the rate at which matter forms collapsed objects. This would be modified if any of f_* , ζ , f_X , or f_α evolves in time. However, it is likely that the evolution of the starformation rate is dominated by the near exponential growth of the collapse fraction and, as we discuss further in Sec. III D, the ordering of events should be highly robust. Our approach, assumes that the large-scale radiation field is largely unaffected by the small-scale properties of the sources. In reality, the field is likely to be highly non-Gaussian, so that the power spectrum contains only a fraction of the total information contained in the brightness temperature fluctuations.

A weakness of our current approach is in handling the cross-correlation between different fluctuations. The same collapsed objects are likely to host sources of x-rays and Ly α photons, so that there will be a strong correlation between the x-ray and Ly α radiation fields. We have assumed that fluctuations in each correlates perfectly to the density field leading to a perfect correlation between the two, so that $P_{\alpha\delta}/\sqrt{P_{\alpha\alpha}P_{\delta\delta}} = 1$ and similarly for $P_{T\delta}$. Significant stocasticity in the sources on small scales would reduce this correlation and so decrease the depth of the trough produced by temperature fluctuations in the absorption regime. Better understanding of this point will require simulations (see [27]), but it could decrease the distinctness of the peaks arising from Ly α and temperature fluctuations.

Our treatment of the propagation of radiation ignores the scattering of photons off resonance and assumes that radiation propagates through an IGM at mean density. Including these effects will reduce the mean free path of photons leading to greater inhomogeneity in the radiation field. This would allow regions far from sources to remain in absorption for longer than our model would predict, since x-ray heating will be more concentrated, and so increase the amplitude of the temperature fluctuation contribution to the power spectrum. We leave this and a detailed study of the interaction between UV and x-ray ionization for future work.

C. Decomposition of fluctuations

The total 21 cm signal is composed of fluctuations in five quantities. Thus, when we form the power spectrum from Eq. (7), we must calculate 15 distinct correlations and cross correlations. We plot the different angular power spectra for Model A in Fig. 5. The μ^4 term is the simplest, arising solely from the VV term, so that it tracks $P_{\delta\delta}$ modulated only by the mean T_b , making it ideal for trying to measure cosmology.

The μ^{0} contribution contains contributions from 10 different terms, and so we have not attempted to show its full decomposition. Instead, we plot the contribution only from the autocorrelations, which gives some idea of where different contributions are important. Note that the contribution of cross terms is significant and that many cross terms have negative signs over some range of redshifts. Hence, $\Delta_{\mu^{0}}$ can be smaller than the sum of the autocorrelations would indicate. Since the μ^{0} term dominates $\bar{\Delta}_{T_{b}}$, its behavior is very similar in form to that plotted in Fig. 2. Temperature Ly α and ionization show clear overlapping regions of contribution.



FIG. 5. Angular decomposition of 21 cm power spectrum for model A at wave number $k = 0.1 \text{ Mpc}^{-1}$. *Top panel:* $\Delta_{\mu^0}(k)$ (solid curve). We also show the separate contribution from the *bb* (dotted curve), $\alpha \alpha$ (long dashed curve), TT (short-dashed curve), and XX (dotted-dashed curve) terms. Note that there are additional cross terms, many of which contribute to the power spectrum with negative sign, not shown here. *Middle panel:* $\Delta_{\mu^2}(k)$ (solid curve). We also plot the contribution from the bV (dotted curve), αV (long-dashed curve), TV (short-dashed curve), and XV (dotted-dashed curve) terms, indicating the sign of the contribution to the power spectrum as positive (thick curves) or negative (thin curves). *Bottom panel:* $\Delta_{\mu^4}(k)$ (solid curve).

Finally, the μ^2 contains contributions from the bV, XV, α V, and TV cross terms. We have plotted these in the middle panel of Fig. 5 to illustrate where different effects become important. Since these terms differ only in one element of the cross correlation, Δ_{μ^2} is more sensitive to the form of the different fluctuations than Δ_{μ^0} . Around $z \approx 15$ all components drop toward zero briefly as $\overline{T}_b = 0$. This imprints an extra dip in the Δ_{μ^2} power spectrum that is not seen in Δ_{μ^0} , where the TT term dominates at this point. Utilizing the different redshift evolution of these power spectra is likely to greatly increase the chance of separating different astrophysical sources of fluctuations.

D. Separation of fundamental physics from astrophysics at high redshifts

We next turn our attention to characterizing the redshift at which the 21 cm signal becomes affected by the astrophysics of star formation. Although, in principle, this could occur when any of $\beta_T \delta_T$, $\beta_x \delta_x$, or $\beta_\alpha \delta_\alpha$ become large in comparison with $\beta_b \delta_b$, as we have seen in the preceding section, the onset of Ly α coupling is likely to provide the first astrophysical modification of the signal. In order to quantify this transition, we will consider the redshift $z_{\text{trans},\alpha}$ at which $\beta_{\alpha}\delta_{\alpha} = \beta_b\delta_b$ for the first time. Although cosmological parameters may be extracted from components other than the *bb* term, this is the dominant component at high redshift.

Figure 6 shows the dependence of $z_{\text{trans},\alpha}$ on the parameters f_X and f_{α} . In order to separate out the different effects, we consider two artificial situations. First, we vary f_{α} using the fiducial value $f_X = 1$, but neglecting Ly α photons produced by x-rays. Next, we vary f_X and neglect the stellar Ly α component, i.e. $f_{\alpha} = 0$. Thus, the plotted values of $z_{\text{trans},\alpha}$ should be regarded as lower limits, since including the Ly α flux of the missing component will cause the transition to occur slightly earlier. All other parameters are set to those of Model A. We plot $z_{\text{trans},\alpha}$ for three different wavenumbers, k = 0.01, 0.1, and 1 Mpc⁻¹, noting that there is little difference between the three curves. For comparison, we also plot the redshift at which Ly α and collisional coupling are comparable, defined by $x_{\alpha} = x_c$.

The basic conclusion is that for a Gaussian field of initial density perturbations, $z_{\text{trans},\alpha}$ depends only logarithmically on f_x and f_{α} . This dependence arises because the fraction of mass in galaxies grows exponentially with cosmic time on the high-density tail of the Gaussian. Exploring 4 orders of magnitude in f_{α} causes $z_{\text{trans},\alpha}$ to vary by a redshift interval of $\Delta z = 8$ ($\Delta \nu \approx 20$ MHz). The same variation in



FIG. 6. Evolution of $z_{\text{trans},\alpha}$ with f_{α} (thin curves) and f_X (thick curves) for k = 0.01 (dotted curves), 0.1 (short dashed curve), and 1 Mpc⁻¹ (long-dashed curve). Also plotted is the redshift at which $x_{\alpha} = x_c$ (solid curves).

JONATHAN R. PRITCHARD AND ABRAHAM LOEB

 f_X causes $z_{\text{trans},\alpha}$ to vary by $\Delta z = 14$ ($\Delta \nu \approx 40$ MHz). This extra dependence arises because modifying f_X also changes T_K . Note that we can relate the uncertainty in the star-formation efficiency f_* to these results, provided we remember that changing f_* will modify the production of both UV and x-ray photons.

We can understand this parameter dependence, by considering the redshift at which $Ly\alpha$ coupling first becomes important. Obtaining $x_{\alpha} = 1$ requires a flux $J_{\alpha} =$ $0.0767(1 + z/20)^2 Ly\alpha$ photons per baryon. We can convert this into the fraction of baryons in stars needed for $Ly\alpha$ coupling

$$\eta_{\alpha} \equiv \frac{J_{\alpha}}{N_{\alpha}} = 8.2 \times 10^{-6} \left(\frac{1+z}{20}\right)^{-2} \left(\frac{6950}{N_{\alpha}}\right).$$
(10)

To a good approximation, we can find the redshift of transition by setting $\eta_{\alpha} = f_{coll}f_*$ to find the point when enough baryons are in stars for the transition to occur. This will slightly overestimate the redshift of transition, as we are not accounting for photons that redshift out of the Ly α resonance. Since f_{coll} initially grows exponentially with redshift we obtain a logarithmic dependence on f_*N_{α} . A similar result will apply for Ly α photons from x-ray sources.

In more extreme astrophysical models, it is conceivable that intense x-ray heating might affect the 21 cm signal before Ly α coupling became important. We therefore also consider the redshift $z_{\text{trans},T}$ at which $\beta_T \delta_T = \beta_b \delta_b$ for the first time. This is plotted in Fig. 7 for the same models used above. For comparison, we also plot the redshift at which the gas is heated to $T_K = T_{\text{CMB}}$, which occurs significantly later and marks the transition from absorption to emission. Unsurprisingly, $z_{\text{trans},T}$ shows very little dependence on f_{α} . The dependence on f_X is more pronounced. In no part of the parameter space explored does $z_{\text{trans},T}$ exceed $z_{\text{trans},\alpha}$, suggesting that $Ly\alpha$ coupling will always be the first source of astrophysical fluctuations. This need not be the case if gas heating does not also produce $Ly\alpha$ photons, for example, if shock heating dominates in the early Universe [70]. We also note that values of $f_* f_X \ge 0.01$ are required to ensure that $T_K > T_{\text{CMB}}$ in the redshift range z < 12 that will be probed by the first low-frequency observatories. These results suggest that the ordering and distinctness of the peaks seen in the power spectrum as a function of redshift is fairly robust with respect to our model parameters.

In the same fashion as the onset of Ly α coupling, we can seek to understand the requirement for heating the gas above the CMB. Assuming that $T_K \ll T_{\text{CMB}}$ at the onset of heating, an energy of $\Delta E \approx 3k_B T_{\text{CMB}}/2$ per baryon is required. Using parameters appropriate for x-ray emission for starburst galaxies and assuming that x_e is approximately the primordial value, yields the fraction of baryons in stars required to heat the gas



FIG. 7. Evolution of $z_{\text{trans},T}$ with the same line conventions as Fig. 6. Also plotted is the redshift at which $T_K = T_{\text{CMB}}$ (solid curves).

$$\eta_T \equiv \frac{\Delta E}{\epsilon_X} = 1.15 \times 10^{-4} \left(\frac{1+z}{20}\right) \left(\frac{560 \text{ eV}}{\epsilon_X}\right) \left(\frac{0.073}{f_{\text{heat}}}\right).$$
(11)

Comparison with Eq. (10) shows that heating the IGM to $T_{\rm CMB}$ with x-rays requires considerably more star formation than producing Ly α coupling. While generalizing this argument rigorously from the mean history to the case of fluctuations is difficult, the basic point remains. Finally, in the absence of significant clumping reionization requires approximately one UV photon per baryon in stars so that we may write

$$\eta_X \equiv \frac{1}{N_{\text{ion,IGM}}} = 2.5 \times 10^{-3} \left(\frac{4000}{N_{\text{ion}}}\right) \left(\frac{0.1}{f_{\text{esc}}}\right).$$
 (12)

Thus, reionization should occur after $Ly\alpha$ coupling is significant and the gas has been significantly heated.

We conclude that for astrophysical parameters not too far from our fiducial ones $f_X \approx f_\alpha \approx 1$ and $f_* \approx 0.1$, Ly α fluctuations will first become important at $z \approx 25$ ($\nu \approx$ 55 MHz) and temperature fluctuations at $z \approx 19.5$, although this latter number is considerably more uncertain. This defines the minimum requirement for a low-frequency observatory hoping to probe the cosmology dominated regime before the first sources affect the 21 cm signal. For the mean history, Ly α coupling becomes relevant at $z \approx 23$ and $T_K > T_{\text{CMB}}$ at $z \approx 14$. These results are somewhat sensitive to the cosmology used, especially the value of σ_8 , which affects the collapse fraction. The location of these landmarks will be important for attempts to probe the mean 21 cm signal.

IV. OBSERVATIONAL PROSPECTS

In Sec. III B, we showed that detecting the 21 cm signal would require foreground removal to the level of 10^{-5} for reionization and cosmic twilight and 10^{-7} to reach the Dark Ages. Next we explore the experimental capabilities necessary to achieve the required sensitivity for an accurate measurement of the 21 cm power spectrum. Before presenting our calculations, let us digress and consider the angular scales probed by these experiments. Until this point our calculations have been presented in terms of comoving wave number k, appropriate for a 3D power spectrum. For observers, the more natural quantity is the angular scale on the sky Θ . Figure 8 shows a conversion between the two using $\Theta = (2\pi/k)/d_A(z)$, with $d_A(z)$ the angular diameter distance to redshift z. We also relate angles to multipole index ℓ in the spherical harmonics decomposition of the sky, based on the approximate relation $l \approx \pi/\Theta$. Although this conversion is somewhat crude, it allows a quick comparison with quantities that might be more intuitive to the CMB researchers. Astrophysics primarily affects the 21 cm power spectrum on comoving scales k = 0.1-1 Mpc⁻¹ corresponding to angular scales of several arc minutes.

We now calculate the signal-to-noise ratio $(S/N = \sqrt{P_{21}/\delta P_{21}})$ for three fiducial experiments: (i) a pathfinder class experiment like the *Murchison Widefield Array*



FIG. 8. Redshift evolution of angular scales. $\Theta = (2\pi/k)/d_A(z)$, which we map to angular moments using $l \approx \pi/\Theta$. This is meant to be an approximate guide, rather than an exact conversion.

(MWA); (ii) a fully fledged EoR instrument like the Square Kilometer Array (SKA); and (iii) a futuristic lunar array (LA) [71]. The parameters used to describe these instruments are summarized in Table I. In each case, we specify the number of antennae N_a , and the total collecting area $A_{tot} = N_a A_e$ for an experiment optimized to observe z = 8. We also specify the bandwidth B and total observing time t_{int} . Note that while we have chosen parameters that correspond roughly to proposed experiments, we take liberties with the antennae configurations to maximize the sensitivity of each experiment at each redshift. We are interested in the most favorable capabilities of each experiment given the parameters in Table I, and so we optimize the associated array configuration for each redshift window. For ease of reference, we will label these experiments MWA, SKA, and LA. Current designs for the actual MWA and SKA limit their frequency range (≥ 80 MHz) and antennae distribution, and so the above labels should not be associated with the actual instruments being designed or constructed. The labels are simply meant to denote different scales of experimental effort.

The variance of a 21 cm power-spectrum estimate for a single **k** mode with line-of-sight component $k_{\parallel} = \mu k$ is given by [72]

$$\sigma_P^2(k,\mu) = \frac{1}{N_{\text{field}}} \left[\bar{T}_b^2 P_{21}(k,\mu) + T_{\text{sys}}^2 \frac{1}{Bt_{\text{int}}} \frac{D^2 \Delta D}{n(k_\perp)} \times \left(\frac{\lambda^2}{A_e} \right)^2 \right]^2.$$
(13)

We restrict our attention to modes in the upper-half plane of the wave vector **k**, and include both sample variance and thermal detector noise assuming Gaussian statistics. The first term on the right-hand side of the above expression provides the contribution from sample variance, while the second describes the thermal noise of the radio telescope. The thermal noise depends upon the system temperature T_{sys} , the survey bandwidth *B*, the total observing time t_{int} , the conformal distance D(z) to the center of the survey at redshift *z*, the depth of the survey ΔD , the observed wavelength λ , and the effective collecting area of each antennae tile A_e . The effect of the configuration of the antennae is encoded in the number density of baselines $n_{\perp}(k)$ that observe a mode with transverse wave number

TABLE I. Low-frequency radio telescopes and their parameters. We specify the number of antennae N_a , total collecting area A_{tot} , bandwidth *B*, and total integration time t_{int} for each instrument.

Array	N_a	$A_{\rm tot}(10^3 {\rm m}^2)$	B (MHz)	$t_{\rm int}$ (hr)
MWA	500	7.0	6	1000
SKA	5000	600	6	1000
LA	7800	3600	8	12000

 k_{\perp} [26]. Observing a number of fields N_{field} further reduces the variance.

Estimates of the error on a power-spectrum measurement are calculated using the Fisher matrix formalism, so that the $1 - \sigma$ errors on the model parameter λ_i are $(\mathbf{F}_{ii}^{-1})^{1/2}$, where

$$F_{ij} = \sum_{\mu} \frac{\epsilon k^3 V_{\text{survey}}}{4\pi^2} \frac{1}{\sigma_P^2(k,\mu)} \frac{\partial P_{T_b}}{\partial \lambda_i} \frac{\partial P_{T_b}}{\partial \lambda_j}.$$
 (14)

In this equation, $V_{\text{survey}} = D^2 \Delta D(\lambda^2/A_e)$ denotes the effective survey volume of our radio telescopes, and we assume wave number bins of width $\Delta k = \epsilon k$. We will be interested in the cases where $\lambda_i = \{\bar{P}_{T_b}\}$ and $\lambda_i = \{P_{\mu^0}, P_{\mu^2}, P_{\mu^4}\}$.

Instrumental sensitivities have been shown to depend upon the distribution of antennae [26,73]. One typical model is to assume a close packed core with filling fraction close to unity surrounded by antennae distributed with a r^{-2} dependence out to a sharp cutoff at the edge of the array. Lidz et al. [72] showed that significant gains in sensitivity can be achieved by using a "super core" configuration, condensing the array so that the filling fraction is unity throughout the array. Since most of the longest baselines in the r^{-2} configuration only poorly sample the largest k modes anyway, this enhances the signal to noise, while losing measurements on only a few large k modes. Since we wish to maximize the sensitivity of our instrument, we will use a similar "super core" configuration. The effective collecting area of the antennae is a strong function of observed frequency (scaling as ν^{-2} or λ^2), so long as antennae are more widely spaced than $\lambda/2$. We optimize our arrays at each redshift by setting the minimum baseline to $\lambda/2$ and then close packing the antennae. Our calculations should then give a reasonable idea of the best performance that can be hoped for given our chosen specifications. In practice, an array will be optimized for a single redshift and will suffer performance degradation when probing higher redshifts, where geometrical shadowing becomes important. Additionally, some fraction of antennae will be reserved to provide the long baselines needed for foreground removal.

Figure 9 shows three redshift slices of the 21 cm power spectrum for Model A. These illustrate the shape of the sensitivity of the various instruments. The sensitivity is degraded by foregrounds on large scales and thermal noise on small scales. In each case, those scales closest to the foreground cutoff (which is determined by the line-of-sight resolution set by the experimental bandwidth) on large scales are the most likely to retain some foreground contamination. Unfortunately, these tend to be the modes that are measured with the highest sensitivity. The shape of the power spectrum on intermediate scales shows significant evolution as different sources of fluctuations become im-



FIG. 9. Errors on the angle-averaged power spectrum $\overline{\Delta}_{T_b}$ (solid curves) for our optimized experiments MWA (dotted curve), SKA (short dashed curve), and LA (long-dashed curve), as functions of wavenumber *k* at redshifts z = 7.9 (top panel), 15.7 (middle panel), and 30.2 (bottom panel).

portant. At z = 30.2, $\overline{\Delta}_{T_b}$ is dominated by density fluctuations. At z = 15.7, temperature fluctuations lead to a trough on a scale $k \approx 1 \text{ Mpc}^{-1}$ and a peak at $k \approx$ 0.1 Mpc^{-1} . The location of this trough evolves at higher redshift to lower k, where it might be more easily detectable. By z = 8, ionization fluctuations flatten the power spectrum significantly. The evolution of the shape of the power spectrum during reionization has been examined in more detail by Lidz *et al.* [72], who explored the ability of the MWA to measure the slope and amplitude of the power spectrum. While the SKA is a significant improvement over the MWA, it takes the futuristic LA to directly observe the trough in $\overline{\Delta}_{T_b}$ at z = 15.7 or to detect fluctuations at z = 30.2.

Figure 10 shows achievable S/N ratio for measurements of the spherically averaged signal $\bar{\Delta}_{T_b}$ for the three instruments as a function of redshift for three different wave numbers k = 0.1, 1, and 10 Mpc⁻¹. These wave numbers span the scales to which these instruments are sensitive. We see that a well-optimized MWA achieves high signal-tonoise ratio S/N for only the k = 0.1 Mpc⁻¹ mode and moderate S/N for k = 1.0 Mpc⁻¹. SKA achieves the sample-variance limit at k = 0.1 Mpc⁻¹ over a wide range of scales and shows sensitivity to k = 1.0 Mpc⁻¹ through most of the EoR. Only the LA is capable of detecting the signal from the *Dark Ages* at z > 30. These curves are optimistic, since the instruments are assumed to be optimized to the specific frequency of observation. At very low



FIG. 10. Signal-to-noise ratio as a function of redshift for our optimized experiments. Curves indicate k = 0.1 (dotted curve), 1.0 (short-dashed curve), and 10 Mpc⁻¹ (long-dashed curve). *Top panel:* MWA. *Middle panel:* SKA. Bottom panel: LA.

redshifts ($z \leq 3$), the bandwidth that we have chosen is inadequate to the task of removing foregrounds on the scale $k = 0.1 \text{ Mpc}^{-1}$, hence the sharp cutoff in sensitivity seen in Fig. 10.

We now turn to the possibility of performing the angular separation by powers of μ discussed earlier. There are two main reasons for performing this separation. First, as discussed in Sec. III, the 21 cm fluctuations arising from different astrophysical models tend to overlap, giving $\bar{\Delta}_{T_b}$ a complicated shape. In order to constrain astrophysics, it may be advantageous to turn to the Δ_{μ^2} term. This term shows more structure than the spherically averaged signal (see Fig. 11), while being simpler to analyze as it is composed of fewer terms. Second, the complexity of the astrophysical fluctuations obscures the underlying cosmological information contained in the density contribution to 21 cm fluctuations. This problem can be circumvented by measuring Δ_{μ^4} , which is determined by the density field alone [29]. Although the use of this term is most important during the astrophysics-dominated regime, it might also be important during the Dark Ages. If exotic energy injection mechanisms, such as decaying dark matter [19], occur during the Dark Ages they would produce fluctuations in temperature and ionization fraction that would interfere with the extraction of cosmology from the density fluctuations. The angular separation would provide extra information about these processes and still allow cosmological constraints to be extracted.



FIG. 11. Experimental sensitivity to the separation of powers. Plotted is the redshift evolution of the power spectrum (thick solid curve) at $k = 0.1 \text{ Mpc}^{-1}$ and the corresponding sensitivity for MWA (thin-solid curve), SKA (thin-dotted curve), and LA (thin-dashed curve). We also plot sensitivity for LA with its observing time split between 16 separate fields (thin dotteddashed curve). Top panel: Δ_{μ^0} . Middle panel: Δ_{μ^2} . Bottom panel: Δ_{μ^4} . Calculations are for Model A.

Figure 11 shows the sensitivity of our fiducial experiments to the angular components as a function of redshift for wave number $k = 0.1 \text{ Mpc}^{-1}$. All three experiments show some sensitivity to Δ_{μ^0} , which is to be expected from our analysis of the spherically averaged signal. Only SKA and LA show any sensitivity to Δ_{μ^2} and Δ_{μ^4} .

Since Δ_{μ^0} dominates the signal, the errors in Δ_{μ^2} and Δ_{μ^4} track its shape. This results in only a few windows where observation of these components is possible. During the end of reionization the largely isotropic ionization signal dominates [26] making observations of the μ^2 and μ^4 components difficult. This is also true earlier when temperature and Ly α fluctuations dominate the signal. Observational prospects for the angular separation are best when these fluctuations are small, at the beginning of the emission period but before reionization is well advanced ($8 \le z \le 12$ for Model A). There is also a possibility for measuring Δ_{μ^2} during the absorption period $(15 \le z \le 21 \text{ for Model A})$ although this relies somewhat upon the scale-dependent cancellation between temperature and Ly α fluctuations. We note that a high S/N detection is required for the angular separation to be feasible. For this reason, none of the instruments are able to perform the angular separation during the Dark Ages. The angular decomposition also becomes more difficult on small scales.

JONATHAN R. PRITCHARD AND ABRAHAM LOEB

Since the SKA and LA experiments are essentially sample-variance limited at $k = 0.1 \text{ Mpc}^{-1}$, it would seem that performing the full angular separation will be exceedingly difficult except in a few redshift windows. This is a direct consequence of the isotropic part of the power spectrum dominating the uncertainty in measurement of the μ^2 and μ^4 terms. To improve on the results illustrated here it would be necessary to beat down cosmic variance by observing multiple fields on the sky. Since the variance scales as $N_{\text{field}}^{-1/2}$, while the contribution from thermal noise scales as t_{int}^{-1} , there is a trade-off involved in splitting observing time into different fields. The experiment will gain sensitivity on scales to which it is samplevariance limited, while losing sensitivity where thermal noise is important. Only if the experiment has sufficient instantaneous sensitivity, will it be worth observing multiple fields. Note though that for some interferometer designs, e.g. LOFAR, multiple fields can be observed simultaneously, improving the S/N, while adding only to the computational cost of the required correlations. For an experiment like LA, which has the sensitivity to reach sample-variance limited measurements, experimental design will be key in distributing antennae and observing time to minimize errors.

Since performing the angular separation may be vital to obtaining both astrophysical and cosmological information, we briefly consider a modified version of the LA in which we redistribute the same fixed total observing time from deep integration of a single field, necessary to probe the Dark Ages, into shallower integrations of 16 separate fields. This is plotted in Fig. 11 (dotted-dashed curve). We see that considerable gains in sensitivity are achievable, dramatically increasing the range of redshifts over which the angular separation is possible at high S/N on this scale. However, this version of the LA is unable to probe the Dark Ages and is sample-variance limited over a smaller range of wave numbers. Since observation strategies for direct observation of the *Dark Ages* versus separating Δ_{μ^4} are orthogonal to one another, detailed consideration of which approach is better for cosmology will be important for the design of future instruments.

V. DISCUSSION AND CONCLUSIONS

Future low-frequency radio observations of the redshifted 21 cm line hold the potential to greatly increase our knowledge of the high redshift Universe. Theoretical understanding of fluctuations in the 21 cm brightness temperature has reached the point where it is possible to begin making predictions for the evolution of the signal from recombination to the present day. In this paper, we have combined the present understanding of fluctuations in density, neutral fraction, $Ly\alpha$ flux, and x-ray heating, to calculate the three-dimensional 21 cm power spectrum as a function of redshift. This draws attention to the three different epochs mentioned in the introduction. Early reionization compresses the reionization and twilight signal into a narrow plateau, likely complicating attempts to separate different fluctuations and learn about astrophysics.

Our comprehensive analysis provides a useful foundation for optimizing the design of future arrays whose goal is to separate the particle physics from the astrophysics, either by probing the peculiar velocity distortion of the 21 cm power spectrum during the epoch of reionization, or by extending the 21 cm horizon to $z \ge 20$ or $z \le 6$.

We have shown that the signal above z = 25 is likely to be relatively uncontaminated by astrophysics. This sets the minimum redshift that a low-frequency radio observatory must probe in order to seek cosmology most simply. This same calculation indicates that experiments with sensitivity down to 60 MHz should be able to observe the full span of the EoR signal.

Observations should be sensitive to both the emission and absorption regimes of the 21 cm signal during the EoR. This is in part due to the greater amplitude of T_b during the absorption regime and in part to the enhancement of the signal from temperature and Ly α fluctuations. Additionally, experiments sensitive to z < 5 should have a chance of detecting the post-reionization signal.

In order for experiments to detect 21 cm fluctuations from the EoR they will need to first remove foregrounds to the level of 10^{-5} . To probe the *Dark Ages*, heroic efforts to obtain the level of 10^{-7} will be required. If this can be achieved then the SKA should be able to measure the EoR signal over several decades of wave number. Significant improvement in either collecting area or, for example, going to the moon for longer integration times will be necessary to access the *Dark Ages*.

Interestingly, the 21 cm emission from residual hydrogen after reionization ($z \leq 6$) offers excellent prospects for probing fundamental physics [24], because of the steep decline in the galactic foreground brightness with decreasing wavelength ($\propto \lambda^{2.6}$). The pockets of self-shielded hydrogen are expected to trace the distribution of matter, and allow a precise determination of the matter power spectrum. This in turn will provide an unprecedented probe of non-Gaussianity and running of the spectral index of the power spectrum of primordial density fluctuations from inflation. Detection of small-scale fluctuations can also be used to infer the existence of massive neutrinos [13] and other subdominant components in addition to the commonly inferred cold dark-matter particles. The scale of the baryonic acoustic oscillations in the power spectrum can be used as a standard ruler to constrain the equation of state of the dark energy through 99% of the cosmic history [22,25,60].

The use of 21 cm cosmology to study fundamental physics is not restricted to either high redshifts ($z \ge 25$; $\nu \le 50$ MHz) or low redshifts ($z \le 6$; $\nu \ge 200$ MHz). The implications for inflation and dark matter can be

EVOLUTION OF THE 21 CM SIGNAL THROUGHOUT ...

separated from the astrophysical effects of star formation through an angular decomposition of the 21 cm power spectrum, even during reionization ($6 \le z \le 25$). The angular term Δ_{μ^4} isolates the contribution of gravitationally induced peculiar velocities. We have found that with an appropriately designed instrument, separating the Δ_{μ^4} term should be possible over a wide range of redshifts.

PHYSICAL REVIEW D 78, 103511 (2008)

ACKNOWLEDGMENTS

J. R. P. is supported by NASA through Grant No. HST-HF-01211.01-A from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under Contract No. NAS 5-26555. We thank the anonymous referees for their comments and suggestions.

- U. Seljak, A. Slosar, and P. McDonald, J. Cosmol. Astropart. Phys. 10 (2006) 14.
- [2] A. Loeb and S. Wyithe, Phys. Rev. Lett. 100, 161301 (2008).
- [3] R.S. Ellis, arXiv:astro-ph/0701024 [Astrophysics (Engl. Transl.) (to be published)].
- [4] D.N. Spergel *et al.*, Astrophys. J. Suppl. Ser. **170**, 377 (2007).
- [5] J.E. Gunn and B.A. Peterson, Astrophys. J. 142, 1633 (1965).
- [6] X. Fan, C. L. Carilli, and B. Keating, Annu. Rev. Astron. Astrophys. 44, 415 (2006).
- [7] Z. Haiman and L. Hui, Astrophys. J. 547, 27 (2001).
- [8] A. Loeb, arXiv:0711.3463.
- [9] A. Loeb, arXiv:astro-ph/0603360 [Astrophysics (Engl. Transl.) (to be published)].
- [10] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Phys. Rep. 433, 181 (2006).
- [11] R. Barkana and A. Loeb, Rep. Prog. Phys. 70, 627 (2007).
- [12] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. 92, 211301 (2004).
- [13] A. Loeb and S. Wyithe, Phys. Rev. Lett. 100, 161301 (2008).
- [14] http://www.haystack.mit.edu/ast/arrays/mwa/.
- [15] http://www.lofar.org/.
- [16] http://astro.berkeley.edu/~dbacker/EoR/.
- [17] http://web.phys.cmu.edu/~past/.
- [18] http://www.skatelescope.org/.
- [19] S. R. Furlanetto, S. P. Oh, and E. Pierpaoli, Phys. Rev. D 74, 103502 (2006).
- [20] K. J. Mack and D. H. Wesley, arXiv:0805.1531.
- [21] A. M. Wolfe, E. Gawiser, and J. X. Prochaska, Annu. Rev. Astron. Astrophys. 43, 861 (2005).
- [22] J. S. B. Wyithe, A. Loeb, and P.M. Geil, Mon. Not. R. Astron. Soc. 383, 1195 (2008).
- [23] A. Lewis and A. Challinor, Phys. Rev. D 76, 083005 (2007).
- [24] S. Wyithe and A. Loeb, arXiv:0708.3392.
- [25] T.-C. Chang, U.-L. Pen, J. B. Peterson, and P. McDonald, Phys. Rev. Lett. **100**, 091303 (2008).
- [26] M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto, Astrophys. J. 653, 815 (2006).
- [27] M. G. Santos, A. Amblard, J. Pritchard, H. Trac, R. Cen, and A. Cooray, arXiv:0708.2424.
- [28] I. T. Iliev, P. R. Shapiro, G. Mellema, U.-L. Pen, P. McDonald, and M. A. Alvarez, arXiv:0710.2451.
- [29] R. Barkana and A. Loeb, Astrophys. J. Lett. 624, L65

(2005).

- [30] M. Zaldarriaga, S.R. Furlanetto, and L. Hernquist, Astrophys. J. **608**, 622 (2004).
- [31] S.R. Furlanetto, M. Zaldarriaga, and L. Hernquist, Astrophys. J. 613, 1 (2004).
- [32] R. Barkana and A. Loeb, Astrophys. J. 626, 1 (2005).
- [33] L. Chuzhoy, M. A. Alvarez, and P. R. Shapiro, Astrophys. J. Lett. 648, L1 (2006).
- [34] J. R. Pritchard and S. R. Furlanetto, Mon. Not. R. Astron. Soc. 367, 1057 (2006).
- [35] J. R. Pritchard and S. R. Furlanetto, Mon. Not. R. Astron. Soc. 376, 1680 (2007).
- [36] S. Bharadwaj and S. S. Ali, Mon. Not. R. Astron. Soc. 352, 142 (2004).
- [37] A.C. Allison and A. Dalgarno, Astrophys. J. 158, 423 (1969).
- [38] B. Zygelman, Astrophys. J. 622, 1356 (2005).
- [39] S.R. Furlanetto and M.R. Furlanetto, Mon. Not. R. Astron. Soc. 374, 547 (2007).
- [40] C. M. Hirata and K. Sigurdson, Mon. Not. R. Astron. Soc. 375, 1241 (2007).
- [41] S. A. Wouthuysen, Astron. J. 57, 31 (1952).
- [42] G.B. Field, Proc. IRE 46, 240 (1958).
- [43] X. Chen and J. Miralda-Escudé, Astrophys. J. 602, 1 (2004).
- [44] C. M. Hirata, Mon. Not. R. Astron. Soc. 367, 259 (2006).
- [45] L. Chuzhoy and P.R. Shapiro, Astrophys. J. 655, 843 (2007).
- [46] S. R. Furlanetto and J. R. Pritchard, Mon. Not. R. Astron. Soc. 372, 1093 (2006).
- [47] S.R. Furlanetto, Mon. Not. R. Astron. Soc. 371, 867 (2006).
- [48] S. C. O. Glover and P. W. J. L. Brand, Mon. Not. R. Astron. Soc. 340, 210 (2003).
- [49] V. Bromm and R.B. Larson, Annu. Rev. Astron. Astrophys. 42, 79 (2004).
- [50] R. Barkana and A. Loeb, Phys. Rep. 349, 125 (2001).
- [51] O. Zahn, A. Lidz, M. McQuinn, S. Dutta, L. Hernquist, M. Zaldarriaga, and S. R. Furlanetto, Astrophys. J. 654, 12 (2007).
- [52] H. Trac and R. Cen, Astrophys. J. 671, 1 (2007).
- [53] M. McQuinn, S. R. Furlanetto, L. Hernquist, O. Zahn, and M. Zaldarriaga, Astrophys. J. 630, 643 (2005).
- [54] P. Madau, A. Meiksin, and M. J. Rees, Astrophys. J. 475, 429 (1997).
- [55] X. Chen and J. Miralda-Escude, Astrophys. J. 684, 18 (2008).

JONATHAN R. PRITCHARD AND ABRAHAM LOEB

- [56] B. Semelin, F. Combes, and S. Baek, Astron. Astrophys. 474, 365 (2007).
- [57] L. Chuzhoy and Z. Zheng, Astrophys. J. 670, 912 (2007).
- [58] S. Naoz and R. Barkana, arXiv:0707.3146.
- [59] S. Naoz and R. Barkana, Mon. Not. R. Astron. Soc. 362, 1047 (2005).
- [60] R. Barkana and A. Loeb, Mon. Not. R. Astron. Soc. 363, L36 (2005).
- [61] H.-J. Seo and D. J. Eisenstein, Astrophys. J. 598, 720 (2003).
- [62] J. R. Pritchard, S. R. Furlanetto, and M. Kamionkowski, Mon. Not. R. Astron. Soc. 374, 159 (2007).
- [63] P.R. Shapiro, M.L. Giroux, and A. Babul, Astrophys. J. 427, 25 (1994).
- [64] N. Y. Gnedin, Astrophys. J. 542, 535 (2000).
- [65] D.J. Eisenstein and W. Hu, Astrophys. J. 496, 605 (1998).

- [66] A. Cooray and R. Sheth, Phys. Rep. 372, 1 (2002).
- [67] A. P. Chippendale, E. Subrahmanyan, and R. Ekers, *New Techniques and Results in Low Frequency Radio Astronomy*, Hobart, Australia, 2005 (unpublished).
- [68] J.D. Bowman, A.E.E. Rogers, and J.N. Hewitt, arXiv:0710.2541.
- [69] A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, S. Dutta, and L. Hernquist, Astrophys. J. 659, 865 (2007).
- [70] S. R. Furlanetto and A. Loeb, Astrophys. J. 611, 642 (2004).
- [71] C.L. Carilli, J.N. Hewitt, and A. Loeb, arXiv:astro-ph/ 0702070 [Astrophysics (Engl. Transl.) (to be published)].
- [72] A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, and L. Hernquist, Astrophys. J. 680, 962 (2008).
- [73] J. D. Bowman, M. F. Morales, and J. N. Hewitt, Astrophys. J. 638, 20 (2006).