Measuring the Primordial Deuterium Abundance during the Cosmic Dark Ages

Kris Sigurdson* and Steven R. Furlanetto†
California Institute of Technology, Mail Code 130-33, Pasadena, California 91125, USA
(Received 10 May 2005; published 29 August 2006)

We discuss how measurements of fluctuations in the absorption of cosmic microwave background photons by neutral gas at redshifts $z \approx 7$–200 could reveal the primordial deuterium abundance of the Universe. The strength of the cross-correlation of brightness-temperature fluctuations in the redshifted 21-cm line of hydrogen with those in the redshifted 92-cm line of deuterium is proportional to the value of the deuterium-to-hydrogen ratio $[D/H]$ fixed during big bang nucleosynthesis. Although challenging, this measurement would provide the clearest possible determination of $[D/H]$, free from contamination by structure formation processes at lower redshifts. We additionally report our result for the thermal spin-structure cross section in deuterium-hydrogen scattering.

DOI: 10.1103/PhysRevLett.97.091301

PACS numbers: 98.80.Ft, 32.10.Fn, 34.50.–s, 95.30.Dr

Introduction.—After the cosmic microwave background (CMB) radiation decoupled from the baryons at a redshift $z \approx 1100$, most CMB photons propagated unfettered through the neutral primordial medium. This has allowed exquisite measurements of the temperature fluctuations in the primordial plasma at the surface of last scattering, and the statistical properties of these fluctuations have recently been used, in conjunction with other observations, to determine the cosmology of our Universe [1]. After the photons kinetically decoupled from the gas at $z \sim 200$, the latter cooled adiabatically with $T_g \propto (1 + z)^2$, faster than the $T_s \propto (1 + z)$ cooling of the CMB. This epoch, with most of the baryons in the form of relatively cold neutral atoms and before the first stars formed, is known as the cosmic dark ages.

The reason most CMB photons propagate unimpeded through the neutral primordial gas is elementary quantum mechanics—atoms absorb nonionizing radiation only at the discrete wavelengths determined by the differences of their atomic energy levels. One interesting example is the well-known 21-cm spin-flip transition [2], due to the hyperfine splitting of the ground state of the hydrogen (H) atom. At any given $z$, CMB photons with wavelength $\lambda_{21} = 21.1 \text{ cm}$ can resonantly excite this transition. By measuring brightness-temperature fluctuations due to density fluctuations in the neutral gas [3], radio telescopes observing at $\lambda = (1 + z)\lambda_{21}$ can probe the matter power spectrum at $z = 30$–200 [4].

In this Letter, we discuss another application of these measurements. Less well-known than the 21-cm transition of neutral H is the spin-flip transition of neutral deuterium (D) at $\lambda_{02} = 91.6 \text{ cm}$ [5]. We show below that cross-correlating brightness-temperature fluctuations at a wavelength $\lambda_{H} = (1 + z)\lambda_{21}$ with those at a wavelength $\lambda_{D} = (1 + z)\lambda_{02}$ allows a measurement of the primordial D abundance. In principle, this technique could constrain the primordial value of $[D/H] = n_D/n_H$ to better than 1%. While there is no physical obstacle to such a measurement, it would certainly be technically challenging and require a heroic experimental effort; simply confirming the presence of D via cross-correlation during the cosmic dark ages is a significantly easier goal.

Deuterium has long been recognized as our best “baryometer” because its primeval relic abundance is so sensitive to the baryon-to-photon ratio $\eta = n_b/s$. Moreover, big bang nucleosynthesis (BBN) [6] is the only known natural production mechanism, although mechanisms inside galaxies can destroy it [7]. The measurement we describe below could thus determine the true BBN abundance of D and, in principle, might improve BBN constraints to the baryon density of the Universe $\Omega_B h^2$.

Hyperfine structure of H and D atoms.—The $\vec{p} \cdot \vec{B}$ interaction between the magnetic moments of the electron and the nucleus splits the ground state of single-electron atoms into eigenstates of the total spin operator $\vec{F} = \vec{S} + \vec{I}$ with eigenvalues $F_+ = I + 1/2$ and $F_- = I - 1/2$ and $\Delta E = (16/3)F_+ \mu_B(g_N \mu_N/a_0^2)$ (e.g., [8]). Here $S$ is electron spin, $I$ is nuclear spin, $a_0$ is the Bohr radius, $\mu_B$ is the Bohr magneton, $\mu_N$ is the nuclear magneton, and $g_N$ is the nuclear g factor ($g_p = 5.56$ for H; $g_D = 0.857$ for D). The proton, with $I = 1/2$, splits the H ground state into a triplet with $F_+ = 1$ and a singlet with $F_- = 0$. The deuteron, with $I = 1$, splits the D ground state into a quartet with $F_+ = 3/2$ and a doublet with $F_- = 1/2$.

The population of atoms in the excited spin state relative to the ground state $n_+ / n_- = (g_+/g_-) \exp(-T_+/T_\gamma)$ can be characterized by a spin temperature $T_\gamma$. Here $g_+ = 2F_+ + 1$ and $g_- = 2F_- + 1$ are spin degeneracy factors and $T_\gamma = \Delta E/k_B$. For H and D, we have $T_{\gamma H} = 0.0682 \text{ K}$, $T_{\gamma D} = 0.0157 \text{ K}$, $(g_d^2/g_d^H) = 3$, and $(g_d^D/g_d^H) = 2$.

Three factors determine $T_\gamma$: absorption of 21-cm CMB photons, absorption and reemission of Lyman-α photons [the Wouthuysen-Field (WF) effect [9,10]], and atomic spin-change collisions (free electrons are unimportant in these environments [10]). The first drives $T_\gamma$ toward $T_g$, while the latter two drive it toward the gas temperature $T_g$.

In equilibrium, the spin temperature of a species $X$ is $T_{\gamma X} = (1 + \chi^X)T_g/T_\gamma/(T_g + \chi^X T_\gamma)$, where $\chi^X = \chi^a + \chi^b$ is the sum of the equilibrium threshold parameters for spin-change collisions and for radiative coupling through the...
WF effect. Explicitly, \( \chi^{H} = (C_{\gamma}^{H} - T_{\gamma}^{H}) / (A_{\gamma}^{H} - T_{\gamma}) \) and \( \chi^{D} = (P_{\gamma}^{D} - T_{\gamma}^{D}) / (A_{\gamma}^{D} - T_{\gamma}) \), where \( C_{\gamma}^{H} \) is the collisional deexcitation rate, \( A_{\gamma}^{H} \) is an Einstein coefficient, and \( P_{\gamma}^{D} \propto P_{a} \), where \( P_{a} \) is the total Lyman-\( \alpha \) scattering rate. At \( z \approx 10 \), before the first galaxies formed, \( P_{a} \) is tiny and the WF effect can be neglected. However, it might have interesting consequences near \( z \approx 10 \).

**H-H and D-H collision rates.**—While the cross section for H-H spin-change collisions \( \Sigma_{\gamma}^{HH} \) is well-known [11–13], we could not locate the D-H spin-change cross section for the temperatures of interest [14] and computed \( \Sigma_{\gamma}^{DH} \) using standard partial-wave phase-shift methods. Our results agree with Refs. [11–13] for \( \Sigma_{\gamma}^{HH} \), with Ref. [13] for \( \Sigma_{\gamma}^{\mu H} \), and with Ref. [16] for \( \Sigma_{\gamma}^{DH} \) at 1 K (after accounting for our more recent molecular potentials).

The collisional deexcitation rate is \( C_{\gamma}^{H} = \tilde{v}_{XH} \Sigma_{\gamma}^{XH} n_{H} \), where \( \tilde{v}_{XH} = \sqrt{8k_{B}T_{g}/\mu_{XH}} \) is the thermal velocity, \( \Sigma_{\gamma}^{XH} \) is the thermally averaged spin-change cross section, and \( n_{H} \) is the number density of H atoms. In Fig. 1, we plot \( \Sigma_{\gamma}^{HH} \), \( \Sigma_{\gamma}^{DH} \), and \( \Sigma_{\gamma}^{\mu H} \). While \( \Sigma_{\gamma}^{DH} \) falls off for \( T_{g} \approx 100 \) K, \( \Sigma_{\gamma}^{DH} \) continues to rise to a peak near \( T_{g} \approx 1 \) K. This occurs because of low-energy s-wave and p-wave contributions to D-H scattering. These do not appear for the H-H system due to its different reduced mass \( (\mu_{HH} = m_{H}/2) \) while \( \mu_{DH} = 2m_{H}/3 \). The discussion of D-H spin change in Ref. [10] did not account for this and incorrectly concluded that \( \Sigma_{\gamma}^{DH} \approx \Sigma_{\gamma}^{HH} \).

**Spin-temperature evolution.**—In Fig. 2, we plot \( T_{\gamma}^{H} \), \( T_{\gamma}^{D} \), and \( T_{\gamma}^{P} \) as a function of \( z \). After the gas cools below \( T_{\gamma} \), collisions keep \( T_{\gamma}^{H} \) and \( T_{\gamma}^{D} \) coupled to \( T_{\gamma} \). Near \( z \approx 30 \), collisions become inefficient for H and \( T_{\gamma}^{H} \) returns to \( T_{\gamma} \). \( T_{\gamma}^{D} \) remains coupled to \( T_{\gamma} \) to significantly lower redshift both because the lifetime of the excited state of D is relatively long \( (A_{\gamma}^{H} / A_{\gamma}^{D} = 61.35) \) and because \( \Sigma_{\gamma}^{DH} \gg \Sigma_{\gamma}^{HH} \) at low \( T_{\gamma} \).

**Brightness-temperature fluctuations.**—When the spin temperature of a given species is less than \( T_{\gamma} \), it will absorb CMB photons. The brightness temperature is \( T_{B}^{X} = aT_{g}^{X} - T_{\gamma}^{X} \), where \( T_{g}^{X} = g_{X}^{e}c^{2}A_{\gamma}^{X}n_{X} / [8(g_{X}^{e} + g_{X}^{s})\pi^{2}k_{B}T_{g}^{X}H(z)] \) is the optical depth of the spin-flip transition in question, \( a = 1/(1 + z) \), and \( H(z) \) is the Hubble parameter. We are interested in correlations between brightness-temperature fluctuations \( \delta T_{B}^{X}(\mathbf{n}, a) = \beta^{X}(\mathbf{n})T_{B}^{X}(\mathbf{n})\delta(\mathbf{n}, a) \) observed in a direction \( \mathbf{n} \) at wavelengths differing by a factor \( \lambda_{92}/\lambda_{21} \). Here \( \beta^{X} = 1 + \lambda_{X}^{H}/\lambda_{X}^{H} + \Gamma(T_{g}^{H} / T_{g}^{X} - 1) + (\lambda_{X}^{H} / \lambda_{X}^{H})d\ln(C_{\gamma}^{X}) / d\ln(T_{\gamma}) \). \( \gamma(\mathbf{n}, a) \) is the fractional density contrast, and \( \lambda^{X} = \lambda^{H}(1 + \lambda^{X}) \). At high \( z \), when \( T_{g} = T_{\gamma} \), \( \Gamma \to 0 \) due to residual Thomson scattering with free electrons [19] (fluctuations are isothermal), but, as the gas begins to cool adiabatically, \( \Gamma \to 0 \) [20]. We have neglected the contributions to \( \delta T_{g}^{X} \) from fluctuations in the neutral fraction (likely to be small at high \( z \)) and, for simplicity, fluctuations in the gradient of the radial velocity \( \delta_{v, i}^{X} \) [20,21]. The latter will enhance our signal by a factor of \( \sim 1 - 2 \).

In Fig. 3, we plot \( T_{B}^{H}, T_{B}^{D}, \sigma_{T_{B}^{H}}, \) and \( \sigma_{T_{B}^{D}} \) (where \( T_{B}^{X} = \epsilon T_{B}^{Y}, \sigma_{T_{B}^{X}} = \epsilon \sigma_{T_{B}^{Y}}, \) and \( \epsilon = [D/H] \)). We see that \( T_{B}^{D} \) and \( \sigma_{T_{B}^{D}} \) peak at much lower \( z \) than their H counterparts because, as discussed above, \( T_{B}^{D} \) is coupled to \( T_{B}^{H} \) to lower \( z \). The lower set of curves in Fig. 3 have \( P_{a} = 0 \) throughout (appropriate for rapid late reionization), while the upper assume strong WF coupling and simultaneous heating at \( z = 14 \) (appropriate if gradual reionization begins at about that time).

**D-H cross-correlations.**—We now calculate the cross-correlation of brightness-temperature fluctuations across frequencies in the ratio \( \lambda_{92}/\lambda_{21} \). We write the brightness-temperature fluctuation due to H or D as \( \mathbb{H}(\mathbf{n}, a) = \beta^{H}(\mathbf{n})T_{B}^{H}(\mathbf{n})\delta(\mathbf{n}, a) \) and \( \epsilon D(\mathbf{n}, a) = \epsilon \beta^{D}(\mathbf{n})T_{B}^{D}(\mathbf{n})\delta(\mathbf{n}, a) \),
Product of observables at frequencies \( \nu = 0.0015 \) signal contributed from a pixel in a frequency band \( \nu = 0.0001 \) all the pixels observed on the sky) of most terms vanish as \( \nu = 0.0014 \). Assuming that \( \nu = 0.0027Tb \) and \( \nu = 0.0023h \) respectively. A radio telescope observing at a frequency \( \nu = 0.0023 \) and \( \nu = 0.0023 \) to-noise contributed from the bands \( \nu_h \) and \( \nu_d \) in the D band, the total signal-to-noise ratio \( S/N_{\nu} = 16f_{sky}/(\theta_d^2) \) pixels for an instrument with maximum baseline \( L \) and angular resolution \( \theta_d = \lambda_d/L = (1+z)\lambda_d/L \) in the D band that observes a fraction \( f_{sky} \) of the sky. The signal-to-noise ratio contributed from the bands \( \nu_h \) and \( \nu_d \) at redshift \( z = 14 \) is thus

\[
S
\left(\frac{\nu}{N_{\nu}}\right) = \frac{\epsilon(HD_{\nu})}{\sqrt{\sigma_{H}^2 + \sigma_{D}^2 + \sigma_{H}^2 + \sigma_{D}^2}}.
\]

If \( S/N(z) \) did not depend on \( z \) and measurements were made over a total bandwidth \( B_h \) in the H band \([B_h = (\nu_{92}/\nu_{21})B_h \) in the D band], the total signal-to-noise ratio would just be Eq. (2) multiplied by a factor \( \sqrt{B_h}/\Delta\nu_{\nu} \). In practice, as \( S/N(z) \) varies with \( z \), we calculate \( S/N(z) \) total by summing \( S/N(z) \) over all redshift bins in quadrature.

We note that our choice for \( T_{sys} \) is only an estimate, and the noise (ultimately due to Galactic synchrotron radiation) varies strongly across the sky. Potential systematic effects including instrument calibration and foreground removal will undoubtedly make a detection more challenging and will necessitate the use of more sophisticated statistical estimators of the cross-correlation signal than the simple estimator we describe here. For instance, as we propose to cross-correlate brightness-temperature fluctuations that vary rapidly as a function of frequency, most foreground signals (which are expected to vary smoothly in frequency space) can be removed by applying a high-pass filter in cofrequency space to multiband observations (which would reject any signal smooth in frequency space) \([22,23]\). In fact, as we are looking for a cross-correlation signal, this type of measurement should be robust to even more pathological foreground and systematic effects. For example, a foreground or instrumental effect which leads to correlations between the \( \nu_h \) band and the \( \nu_d \) band may also lead to related correlations between the \( \nu_h \) band and brightness-temperature fluctuations:

\[
\epsilon(HD_{\nu}) = \sigma\frac{c}{B_h} \beta_{\nu}^h \beta_{\nu}^d. \]

Here \( \sigma^2 \) is the variance of density fluctuations averaged over the flat cylindrical volume of each pixel. Explicitly,

\[
\sigma^2 = \frac{2}{\pi^2} \int_0^\infty dkkP(k) \int_0^\infty dkj\beta^2(k,\rho) \frac{J_0^2(\sqrt{k^2-k^2\rho})}{(k^2-k^2\rho)^2},
\]

for a pixel at redshift \( z \) with comoving radius \( \rho \) and thickness \( 2\xi \). The noise in each pixel is a combination in quadrature of the random noise \( \langle N^2 \rangle = \sigma^2_N \) and the confusion arising from the autocorrelation of the H and D fluctuations. In practice, since \( \epsilon^2(D^2) \ll (H^2) \), we can always neglect the confusion due to D fluctuations in the noise. The average noise per pixel in the cross-correlation measurement is thus

\[
\sigma^2 = \frac{2}{\pi^2} \int_0^\infty dkkP(k) \int_0^\infty dkj\beta^2(k,\rho) \frac{J_0^2(\sqrt{k^2-k^2\rho})}{(k^2-k^2\rho)^2},
\]

for a pixel at redshift \( z \) with comoving radius \( \rho \) and thickness \( 2\xi \). The noise in each pixel is a combination in quadrature of the random noise \( \langle N^2 \rangle = \sigma^2_N \) and the confusion arising from the autocorrelation of the H and D fluctuations. In practice, since \( \epsilon^2(D^2) \ll (H^2) \), we can always neglect the confusion due to D fluctuations in the noise. The average noise per pixel in the cross-correlation measurement is thus

\[
\sigma^2 = \frac{2}{\pi^2} \int_0^\infty dkkP(k) \int_0^\infty dkj\beta^2(k,\rho) \frac{J_0^2(\sqrt{k^2-k^2\rho})}{(k^2-k^2\rho)^2},
\]

for a pixel at redshift \( z \) with comoving radius \( \rho \) and thickness \( 2\xi \). The noise in each pixel is a combination in quadrature of the random noise \( \langle N^2 \rangle = \sigma^2_N \) and the confusion arising from the autocorrelation of the H and D fluctuations. In practice, since \( \epsilon^2(D^2) \ll (H^2) \), we can always neglect the confusion due to D fluctuations in the noise. The average noise per pixel in the cross-correlation measurement is thus

\[
\sigma^2 = \frac{2}{\pi^2} \int_0^\infty dkkP(k) \int_0^\infty dkj\beta^2(k,\rho) \frac{J_0^2(\sqrt{k^2-k^2\rho})}{(k^2-k^2\rho)^2},
\]

for a pixel at redshift \( z \) with comoving radius \( \rho \) and thickness \( 2\xi \). The noise in each pixel is a combination in quadrature of the random noise \( \langle N^2 \rangle = \sigma^2_N \) and the confusion arising from the autocorrelation of the H and D fluctuations. In practice, since \( \epsilon^2(D^2) \ll (H^2) \), we can always neglect the confusion due to D fluctuations in the noise. The average noise per pixel in the cross-correlation measurement is thus

\[
\sigma^2 = \frac{2}{\pi^2} \int_0^\infty dkkP(k) \int_0^\infty dkj\beta^2(k,\rho) \frac{J_0^2(\sqrt{k^2-k^2\rho})}{(k^2-k^2\rho)^2},
\]
another frequency band $\nu_a$. However, a statistical estimator for $\epsilon = [D/H]$ can be devised that rejects correlations between the $\nu_a$ band and $\nu_d$ band if the fluctuations in these bands are also significantly correlated with the fluctuations in the $\nu_a$ band or any other band. The extent to which these or other techniques will be needed will become clearer in the years to come as 21-cm experiments gain real-world experience with these issues.

If collisions dominate the coupling between $T_3^H$ and $T_g$ down to $z \approx 7$, then we estimate that a value $[D/H] \approx 3 \times 10^{-5}$ could be detected at $1-2\sigma$ by a benchmark experiment with $L \sim 7.5$ km and $\Delta \nu_d \sim 100$ kHz in $\sim 6$ years. If, however, the first generation of stars created a flux of Lyman-$\alpha$ photons which coupled $T_3^H$ to $T_g$ until $z \sim 7$ through the WF effect (without heating the gas), a similar detection might be made by a smaller experiment with $L \sim 2.5$ km and $\Delta \nu_d \sim 1$ kHz. Although it strongly enhances the 21-cm signal, the WF effect does not improve the [$D/H$] measurement by the same margin once $\sigma_{\nu_d} \gg \sigma_{\nu_a}$. In that regime, it only makes the H fluctuations a better template. Finally, we note that an experiment capable of mapping 21-cm brightness-temperature fluctuations out to $l_{\text{max}} \approx 10^5$ (where it may be a powerful probe of the small-scale matter power spectrum [4]) could measure [$D/H$] to a precision as good as $\sim 1\%$—or even $\sim 0.1\%$ if the WF effect coupling is efficient.

For these estimates, we have assumed a $\Lambda$CDM cosmology with $n_s = 1$ and that a significant fraction of the Universe remains neutral until $z \sim 7$. The largest contribution to the signal originates from $z \sim 10$, where the D signal peaks and the variance in the density fluctuations is largest. Complete and sudden reionization at $z_r = 10$ (14) would reduce the observable signal by a factor of 5 (20). However, a long phase of partial ionization would have much less severe effects (only reducing the number of available pixels by the filling factor of partially neutral gas). Moreover, any heating from these ionizing sources would boost the signal by strengthening the H template (upper curves in Fig. 3): If heating begins at $z = 14$, $(S/N)_{\text{tot}}$ decreases by only a factor of 2 relative to our fiducial case if $z_r = 10$—and it even increases the signal by a factor of 2 if $z_r = 7$.

Discussion.—Despite the obvious technical challenges in observing this signal, it has the virtue of providing the cleanest possible measurement of the primordial [D/H], free from contamination by structure formation processes at lower $z$. Via the window of BBN, this would allow radio telescopes to peer into the first few minutes of the Universe. We believe future searches for cosmic 21-cm fluctuations should bear this possibility in mind.

We also note that $^3\text{He}^+$ has a hyperfine transition (with $\lambda = 3.46$ cm) that can be used in a similar fashion; it has the advantage of much lower foreground contamination at higher frequencies. This line will show a strong anticorrelation with the 21-cm signal during reionization. If the details of reionization can be understood well enough, the cross-correlation of this line with the 21-cm line could supplement the D-H experiment, especially if the Universe has a long period of partial ionization.

We thank A. Dalgarno, J. Gair, M. Hayden, M. Kamionkowski, T. Pearson, M. Reynolds, D. Scott, and B. Zygelman for discussions. K. S. acknowledges the support of the Canadian NSERC. This work was supported in part by NASA Grant No. NAG5-9821 and DOE Grant No. DE-FG03-92-ER40701.

*Electronic address: ksigurds@tapir.caltech.edu
†Electronic address: sfurlane@tapir.caltech.edu

[14] It has been computed at higher temperatures in Ref. [15], at 1 K in Ref. [16], and measured at 1 K in Ref. [17].