ON POSSIBLE VARIATION IN THE COSMOLOGICAL BARYON FRACTION

GILBERT P. HOLDER\textsuperscript{1}, KENNETH M. NOLLETT\textsuperscript{2}, AND ALEXANDER VAN ENGELEN\textsuperscript{1}

\textsuperscript{1}Department of Physics, McGill University, Montreal, QC H3A 2T8, Canada
\textsuperscript{2}Physics Division, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

Received 2009 July 26; accepted 2010 April 9; published 2010 May 25

ABSTRACT

The fraction of matter that is in the form of baryons or dark matter could have spatial fluctuations in the form of baryon–dark matter isocurvature fluctuations. We use big bang nucleosynthesis calculations compared with observed light-element abundances as well as galaxy cluster gas fractions to constrain cosmological variations in the baryon fraction. Light-element abundances constrain spatial variations to be less than 26%–27%, while a sample of “relaxed” galaxy clusters shows spatial variations in gas fractions less than 8%. Larger spatial variations could cause differential screening of the primary cosmic microwave background (CMB) anisotropies, leading to asymmetries in the fluctuations, and ease some tension with the halo-star $^7$Li abundance. We also show that fluctuations within our allowed bounds can lead to “$B$-mode” CMB polarization anisotropies at a non-negligible level.

Key words: cosmological parameters – cosmology: miscellaneous – galaxies: abundances – galaxies: clusters: general – large-scale structure of universe

Online-only material: color figures

1. INTRODUCTION

Is the cosmological fraction of matter that is in the form of baryons a universal constant? There is no compelling theoretical model that would predict it to be observably non-constant, but there is little empirical evidence that it is indeed a constant to high precision.

The cosmological gravitational potential fluctuations are observed to be at the level of $10^{-5}$, but this measures fluctuations in the sum of baryonic and non-baryonic mass. The fraction of matter in the form of baryons could have much larger long wavelength fluctuations with little observable imprint in the cosmic microwave background (CMB). Such a modulation is a baryon–cold dark matter (CDM) isocurvature mode, and its imprint on the CMB is a second-order effect; for a detailed discussion of the physics of this isocurvature mode, see Gordon & Lewis (2003). On scales larger than the sound horizon at recombination there is almost no effect, while on scales that are sub-horizon at last scattering there would be an additional modulation of the acoustic peaks due to a varying sound speed. The extra pressure due to the baryons is negligible compared to the photon pressure.

In this paper, we investigate other effects of the variation of the baryon fraction. We find that large-scale modulation is not only allowed but would also actually alleviate some current mild tensions in the standard cosmology.

In particular, we below investigate constraints from big bang nucleosynthesis (BBN) in light of observed light-element abundances, observations of galaxy cluster gas fractions, and modulation of the optical depth to Thomson scattering of CMB photons. We also investigate the expected contribution to CMB polarization anisotropy and discuss other possible probes.

2. BIG BANG NUCLEOSYNTHESIS

The light-nuclide yields from BBN are functions only of the baryon density $\rho_B$ of the universe under the standard assumptions of uniform entropy per baryon, only standard-model particles, neutrino–antineutrino asymmetries not enormously larger than the corresponding baryon asymmetry, and no late additions of entropy (Schramm & Turner 1998). Observations of extragalactic deuterium provide a particularly tight constraint on $\rho_B$ based on the steep dependence of the deuterium yield on this parameter. Using $\Omega_B H_0^2 = 8\pi G \rho_B / 3$ to express the constraint in terms of the fraction $\Omega_B$ of the closure density provided by baryons, the average extragalactic D/H (≡ ratio by number of deuterium to hydrogen) gives $\Omega_B h^2 = 0.0213 \pm 0.0010$ (Pettini et al. 2008); $H_0 = h \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the Hubble constant. CMB anisotropy measurements with the Wilkinson Microwave Anisotropy Probe (WMAP) satellite now imply $\Omega_B h^2 = 0.02273 \pm 0.00062$ (Dunkley et al. 2009), in agreement with D/H. In these terms, the baryon fraction of the total mass density $\rho_M$ is $f_B = \rho_B / \rho_M$.

There is an extensive literature on the effects of inhomogeneities in the baryon distribution during BBN, going back to at least Wagoner (1973) and Epstein & Petrosian (1975). For a review of baryon-inhomogeneous and other non-standard BBN models up to 1993, see Malaney & Mathews (1993).

Baryon-inhomogeneous scenarios of BBN may be divided into three cases, depending on the typical scale of the fluctuations: scales that are comparable to particle-diffusion horizons during BBN (such that the physics of BBN is affected), larger scales that are still small compared to the scale of our Galaxy (such that observations within our Galaxy are a blend of many different volumes that individually underwent nearly homogeneous BBN), and large scales today (such that cosmological scales sample distinct regions of uniform baryon density).

In the first case, which dominates the literature on inhomogeneous BBN, the inhomogeneities are comparable in length scale to the neutron diffusion length during BBN ($\sim 4 \times 10^{-5}$ comoving pc at the start or $\sim 0.08$ comoving pc at the end of BBN; Kurki-Suonio et al. 1997). Such short-length-scale inhomogeneities only have strong effects if the density contrasts are large (ratios of $\sim 10^6$ between high- and low-density domains), as was suggested in the 1980s on the basis of a strongly first-order quantum chromodynamics phase transition (Witten 1984).

A first-order phase transition of this kind has since fallen out of favor but is not completely ruled out (Boyanovsky et al. 2006).
The corresponding BBN models tend to overproduce lithium and underproduce deuterium relative to current measurements; for a recent assessment of constraints, see Lara et al. (2006). This is not the case that we consider in the present work.

Inhomogeneities longer than the diffusion length result in BBN at each location occurring independently at different $\rho_B$. These were investigated by Epstein & Petrosian (1975), and the most elaborate calculations were performed with a view toward constraining isocurvature baryon fluctuations as seeds of structure formation (Jedamzik & Fuller 1995; Copi et al. 1995; Kurki-Suonio et al. 1997). These authors assumed that any single observation is a blended sample containing matter drawn from the full distribution of $\rho_B$; the resulting effect on BBN is a “smearing out” of the standard BBN predictions of abundance versus mean baryon density. Calculations in the three works last cited mostly discarded perturbations exceeding the Jeans mass at recombination (which would not be observable sites of low metallicity today), with the effect of diminishing the contributions of high-$\rho_B$ domains. D/H then becomes much larger than observed, and the primordial $^4$He mass fraction $Y_P$ is reduced relative to the homogeneous case.

The main constraint on blended inhomogeneities arose from the flattening of the Li/H versus $\rho_B$ curve, which limits mixed inhomogeneities to $\delta \rho_B/\rho_B \lesssim 1$. If the Li/H observed in low-metallicity halo stars is undepleted, the primordial Li/H is constrained to lie near a minimum of the Li/H versus $\rho_B$ curve of standard BBN. Any mixing with regions of $\rho_B$ far from the minimum increases Li/H unacceptably. (See in particular Figures 1–3 of Kurki-Suonio et al. 1997 for these results.) The difference of the WMAP- and D/H-inferred $\rho_B$ from the Li/H minimum obviously complicates this inference.

The type of inhomogeneity to be addressed here is on length scales that did not mix during structure formation, so that abundance measurements are not blends. This possibility was noted in earlier work, particularly in the light of large differences between the early claims of extragalactic D/H measurements (Kurki-Suonio et al. 1997; Copi et al. 1998). It was not extensively pursued because most observational constraints (all Li/H and $^3$He/H, and all D/H before 1995) were within the Galaxy and presumed to constitute a single well-mixed sample, and because the higher claimed D/H values were eventually discounted.

We now examine the implications of the existing light-nuclide abundance data for unmixed inhomogeneities. The emphasis in the earlier literature was on how inhomogeneities affected the BBN upper bound on $\Omega_B$; we take $\Omega_B$ as a more or less settled quantity and ask how large fluctuations away from it can be.

2.1 $^2$H

The strongest BBN probe of $\rho_B$ is deuterium, with D/H $\propto \rho_B^{-1.6}$ (Burles et al. 2001). It also constrains site-to-site variations, since observations of deuterium believed to be primordial span a range of redshifts from about 2.5–3.6 at varying directions on the sky; the corresponding length scales are comparable to the present cosmological horizon.

In the middle panel of Figure 1, we show likelihood as a function of $\Omega_B h^2$ derived from seven measurements of D/H in quasar absorption-line systems. These are the results tabulated in Kirkman et al. (2003; from Tytler et al. 1996; Burles & Tytler 1998a, 1998b; O’Meara et al. 2001; Pettini & Bowen 2001; Kirkman et al. 2000) plus those of O’Meara et al. (2006) and Pettini et al. (2008). Following the discussion in Kirkman et al. (2003), O’Meara et al. (2006), and Pettini et al. (2008), we have omitted from the figure and subsequent analysis the D/H toward Q0347-3819 (Levshakov et al. 2002) and the absorber at $z = 3.256$ toward PKS1937−1009 (Crighton et al. 2004; the error on the latter being claimed to be underestimated).

It is well established in the literature that there is dispersion within the D/H sample in excess of the reported uncertainties; for example, a jackknife estimate of the error on log D/H is about twice that derived by the inverse-variance estimate (Kirkman et al. 2003; O’Meara et al. 2006; Pettini et al. 2008). It may simply be that the errors have been underestimated. Alternatively, one can accept the published error estimates and use the data to limit variation in $\rho_B$ among the locations where deuterium is observed.

2.2 $^4$He

Because the primordial $^4$He mass fraction $Y_P$ depends very weakly on $\rho_B$, useful constraints require very difficult percent-level abundance determinations in sites of low metallicity. Operationally, this means observations of H I regions in blue compact dwarf galaxies (BCDs) or the Magellanic Clouds (most recently, Olive et al. 1997; Peimbert et al. 2000, 2007; Izotov et al. 2007). Since these observations are extragalactic, they might be sensitive to large-scale abundance variations. The sample of H I regions well measured for $Y_P$ extends only to $z \sim 0.05$, or $\sim 200 h^{-1}$ Mpc. Izotov et al. (2007) recently added to the available sample 271 BCDs identified from the Sloan Digital Sky Survey (SDSS); most of these are in the same redshift range as the previous sample, but there are tens of objects in the range 200–500 $h^{-1}$ Mpc (0.05 $\lesssim z \lesssim 0.13$).
The inferred $Y_p$ values are known to suffer from systematic difficulties. Disagreement among groups using different methods exceeds their claimed errors, and assessments of the situation are sometimes very pessimistic (Olive & Skillman 2004). After a recent revision of the atomic data on which $Y_p$ rests, some researchers find values in good agreement with WMAP (Peimbert et al. 2007), while others find $Y_p$ too high for a comfortable match (Izotov et al. 2007; Izotov & Thuan 2010).

The systematic problems with $Y_p$ should be relatively unimportant for present purposes, because galaxy-to-galaxy variation of $Y_p$ should show up as dispersion within any uniformly analyzed sample. Whatever the method of analysis, underlying variation among the data should introduce point-to-point scatter, and this is what we seek to constrain.

Dispersion within the data is not as simple as variation away from the mean $Y_p$, because truly primordial $Y$ is not measured anywhere. All H II regions observed for $Y_p$ have small but nonzero metal content, and the same stars that made the metals also made $^4\text{He}$. Even though systems observed for $Y_p$ are chosen for low metallicity, inferred values of $Y_p$ are almost always based on extrapolation to O/H = 0. One either applies a $Y$ versus O/H relation calibrated elsewhere or carries out a linear regression of $Y$ versus O/H among observed H II regions.

In our analysis, we place limits on dispersion within the “HeBCD” sample of 93 uniformly analyzed H II regions from Izotov et al. (2007). These data constitute the largest such sample and the only large sample analyzed with the new atomic data. We have not included the additional SDSS H II regions from the same publication. These were not subjected to the same data reduction as the HeBCD sample; they were selected differently; they are overall of higher metallicity; their regressed $Y_p$ are consistent with the HeBCD sample at the 2σ level, but regressions of the two sets have rather different slopes. The errors on the SDSS sample are also large enough such that a joint regression of both sets is still dominated by the smaller HeBCD sample (Izotov et al. 2007).

We assume that the $Y_p$ for each object in the HeBCD sample may be found by subtracting from its $Y$ the product of O/H with the regression-line slope found by Izotov et al. (2007) for that sample. In the analysis of Section 4, this is equivalent to assuming that the difference between the local $Y_p$ and the mean $Y_p$ is equal to the difference between the measured $Y$ and the regression line.

This procedure has obvious limitations, since the occurrence of stellar nucleosynthesis in discrete stars can introduce scatter in both $Y$ and O/H away from the mean trend. It is even conceivable that galaxy-to-galaxy differences of $Y_p$ could affect the slope of $Y$ versus O/H through effects on stellar nucleosynthesis. Our simple procedure is consistent with the existing literature (cf. Peimbert et al. 2007, wherein an assumed $Y$–O/H relation is used to find $Y_p$ for each H II region individually). We believe that this level of sophistication is reasonable in the light of the known systematic problems and allows a rough, if optimistic, estimate of the degree to which $^4\text{He}$ constrains variations in $\rho_B$.

Constraints on $\Omega_b h^2$ from the resulting $Y_p$ for each of the 93 H II regions are shown in the lower panel of Figure 1. It has been argued in the literature (Olive & Skillman 2004; Aver et al. 2010) that the error bars within our $^4\text{He}$ sample have been underestimated. The effect of larger errors on individual $Y$ inferences would be to broaden the corresponding likelihood curves and to loosen the constraint on variation of $\rho_B$ from $^4\text{He}$. Although we do not examine the other large collection of more or less uniformly analyzed BCD $^4\text{He}$ abundances (Olive et al. 1997), we note that signs of excess scatter about the best-fit $Y$ versus O/H relation for those data have been sought and were not found (Pagel et al. 1992; Olive et al. 1997). We reiterate that systematic differences between data sets greatly exceed both scatter and error estimates within data sets.

2.3. $^7\text{Li}$

Measurements of primordial lithium are confined to the atmospheres of metal-poor stars in the Galactic halo. Since our present interest is in variations of $\rho_B$ on much larger length scales, all of these stars together constitute a single sample.

Dispersion of Li/H among the stars of the “Spite plateau”—the flat region in the graphs of Li/H versus both effective temperature and metallicity—has been actively sought ever since the discovery of the plateau and its identification with the primordial Li/H (Spite & Spite 1982). This is because star-to-star variation along the plateau would be a strong indication that the observed lithium has been depleted (Deliyannis et al. 1993), the dispersion having arisen from varying amounts of depletion. Some authors claimed dispersion of 0.04–0.1 dex (Thorburn 1994; Deliyannis et al. 1993), but most have stressed its absence (Ryan et al. 1999; Charbonnel & Primas 2005; Asplund et al. 2006; Bonifacio et al. 2007). Very recent efforts indicate that stars of the lowest metallicities show increased scatter of Li/H below the plateau inferred at higher metallicities (Frebel et al. 2008; Aoki et al. 2009; Sbordone et al. 2010). This effect seems most naturally explained by stellar depletion; it cannot be explained by variations of $\rho_B$ because the scatter extends below the minimum producible at any density. In the upper panel of Figure 1, we show the galactic halo as a single sample of Li/H, using the values from Asplund et al. (2006) and Bonifacio et al. (2007). Primordial Li/H values published over a 30 year span all agree within the quoted errors.

These observations at face value indicate that our lone sample of primordial Li/H is a factor of 3 below the Li/H predicted by the $\Omega_b h^2$ from WMAP. (See Cyburt et al. 2008 for an overview of the current situation.) Spatial variations about the WMAP mean would help to explain this discrepancy if we live in a region of low $\rho_B$. However, inspection of Figure 1 shows that variations in excess of 50% would be needed for a complete explanation. This seems implausible in the light of constraints derived below.

2.4. $^3\text{He}$

There is no convincingly primordial observed $^3\text{He}$/H: all observations of $^3\text{He}$/H are near solar metallicity, and a strong increase of $^3\text{He}$/H with time is predicted from models of stellar nucleosynthesis. Measurements of $^3\text{He}$/H in various locations within the Galaxy (Geiss & Gloeckler 1998; Gloeckler & Geiss 1996; Bania et al. 2002, 2007) fail to see any evolution of $^3\text{He}$/H and are all near the WMAP-inferred primordial value. Presumably some mechanism of $^3\text{He}$ destruction is active, and its outlines seem clear (Hogan 1995; Wasserburg et al. 1995; Charbonnel 1995). Among other things, it causes difficulties for the traditional (D + $^4\text{He}$)/H constraint applied in the earlier literature on inhomogeneous BBN.

Since all existing $^3\text{He}$/H measurements are within the Galaxy, we are near the WMAP-inferred prediction, and suffer from large uncertainties in the post-BBN evolution, they do not provide constraints on variation of $\rho_B$. 

No. 2, 2010 ON POSSIBLE VARIATION IN THE COSMOLOGICAL BARYON FRACTION 909
3. GALAXY CLUSTER GAS FRACTIONS

Measurements of the galaxy cluster gas fraction can provide strong constraints on the cosmic baryon fraction, provided the galaxy clusters adequately sample the cosmic baryon fraction distribution (White et al. 1993). Galaxy clusters emit X-rays through a combination of thermal bremsstrahlung and line emission, allowing detailed reconstructions of both density and temperature profiles. The electron density can be directly integrated to obtain the total gas mass, while the temperature information can be added to obtain an estimate of the total mass (assuming hydrostatic equilibrium).

The sample of Allen et al. (2008) has gas fraction determinations of 42 galaxy clusters covering the redshift range $z \sim 0.06–1$. This is a large homogeneously analyzed sample of galaxy clusters that was primarily selected to include objects that appear to be relatively undisturbed (relaxed).

In practice, there are real concerns about the importance of the assumption of spherical or biaxial symmetry (even triaxial ellipsoidal symmetry is a strong assumption), deviations from equilibrium, non-thermal pressure support, non-equilibration of electron and proton temperatures, and other issues. However, empirically, this method has been found to be relatively robust when applied to simulated clusters with many real-world complications (Nagai et al. 2007), and the restriction to only include relatively undisturbed clusters should significantly reduce many of these complications.

It is found that galaxy cluster gas fractions are below the cosmic mean when measured well within the virial radius (Vikhlinin et al. 2006; Allen et al. 2008). For example, Allen et al. (2008) find $f_{\text{gas}} = 0.1105 \pm 0.0016$, while the baryon fraction reported in the WMAP five-year release (WMAP5) is $0.17 \pm 0.01$. Star formation and feedback from accretion onto black holes both act to reduce the gas fraction in galaxy clusters well below the cosmic mean (Puchwein et al. 2008; Nagai et al. 2007), by both locking baryons into stars and pushing gas outside the virialized region. In spite of these processes that can strongly affect the mean gas fraction, there is little evidence for cluster-to-cluster scatter; this can be seen by eye in Figure 2, where most of the individual cluster probabilities are consistent with the mean gas fraction. This will be quantitatively investigated below.

4. CONSTRAINTS ON BARYON FRACTION VARIATIONS

A maximum likelihood fit was done to each data set, varying the mean, $m$, and intrinsic scatter, $s$. The fits were done to the gas fractions and abundance determinations. The figure of merit used was

$$-2 \ln L = \sum_i \left[ \frac{(d_i - m)^2}{s^2 + \sigma_i^2} + \ln \left( s^2 + \sigma_i^2 \right) \right],$$

with $d_i$ the data and $\sigma_i$ their errors.

Marginalizing over the mean, an estimate of the likelihood of the intrinsic scatter was obtained for each data set. For the galaxy cluster gas fractions, this was translated directly into scatter in the fractional baryon fraction. For each BBN abundance, a value of $s$ was computed and then translated to scatter in baryon density by the application of the fits of Burles et al. (2001), expanded around the best-fit WMAP5 value for $\Omega_B h^2$ (Dunkley et al. 2009). These provided

$$\frac{\Delta \Omega_B h^2}{\Omega_B h^2} \sim \frac{\Delta Y_P}{0.0087},$$

for the helium abundance and

$$\frac{\Delta \Omega_B h^2}{\Omega_B h^2} \sim \frac{\Delta [\log_{10}(D/H)]}{0.69},$$

for the deuterium abundance.

Results of the likelihood calculations are shown in Figure 3. Upper limits on intrinsic scatter in the baryon fraction were obtained by integrating the likelihoods starting at $s = 0$ until 95% of the probability was enclosed. Constraints from $^4\text{He}$ and D/H allow 95% confidence upper limits on intrinsic scatter of $\delta \rho_B/\rho_B < 0.27$ and $\delta \rho_B/\rho_B < 0.26$, respectively. Because of the systematic problems noted in Section 2.2, we consider the D/H constraint to be more robust than the $^4\text{He}$ constraint. From
The galaxy cluster gas fractions, this limit is only 0.08. The best-fit $s$ for $\log_{10} D/H$ at $s \sim 0.07 (\delta \rho_B/\rho_B \sim 0.10)$ in Figure 3 is in agreement with a similar analysis by Pettini et al. (2008).

The scales probed by these measurements vary greatly: the $^4\text{He}$ measurements lie at low redshift (less than 0.1), the galaxy clusters are at intermediate redshift (less than 1), and the D/H constraints are at $z \sim 2.5–3.5$.

The possible selection effects are also very different. The $^4\text{He}$ constraints come from a large sample of nearby low-metallicity galaxies experiencing star formation today. The D/H measurements require a strong signature of D in a quasi-stellar object (QSO) absorption spectrum (hence a large hydrogen column density) and favorable conditions for fitting D and H column densities (hence a simple velocity structure with minimal nearby absorption and a bright QSO).

The galaxy clusters are selected for being preferentially “relaxed” in their appearance. This means appearing undisturbed and relatively circular. In a universe with varying baryon fractions, such selection could be problematic: low-baryon regions are likely to have larger density fluctuations on small scales (Eisenstein & Hu 1998), and clusters are likely to form earlier there. If a sample is selected for clusters that are apparently dynamically more mature (relaxed), this sample could be skewed toward regions of lower baryon fraction, where clusters formed earlier. A sample of galaxy clusters selected in this way would thus show less scatter than that of the universe as a whole, and would be systematically biased toward regions of lower baryon fraction. Detailed numerical simulations of galaxy cluster formation will be required to better understand this important selection effect.

In summary, there is no evidence for cosmological variation in the baryon fraction; the amplitude is constrained to be no larger than roughly 10%–30%. For comparison, baryon–photon or CDM–photon isocurvature modes have been constrained to have amplitudes on horizon scales at the level of $10^{-5}$ or smaller (e.g., Sollom et al. 2009 and references therein).

5. DIFFERENTIAL THOMSON SCATTERING

In recent years, there have been indications of a possible large-scale asymmetry in CMB fluctuations (Eriksen et al. 2004, 2007; Hansen et al. 2004, 2009; Bernui et al. 2006; Hoftuft et al. 2009).

Large-scale baryon fraction fluctuations would cause such an asymmetry through differential Thomson scattering. If there are significantly more electrons along some lines of sight, this leads to a differential damping of primary CMB fluctuations. The observed dipole asymmetry in the CMB fluctuations is found to be $0.072 \pm 0.022$ (68% confidence; Hoftuft et al. 2009) in the temperature fluctuation amplitude, and the current best fit for the total Thomson optical depth is $0.087 \pm 0.017$ (Dunkley et al. 2009). Thus, to explain the observed anomaly would require variations of $0.8 \pm 0.3$ (at 68% confidence) in the baryon fraction on cosmological scales. Such a large variation (but only 2.7$\sigma$ from zero) is observed in neither D/H abundances nor galaxy cluster gas fractions, but smaller variations could be consistent with both the observed asymmetry and the light-element constraints.

Thomson scattering by a region with a fluctuating baryon fraction would also lead to polarization anisotropies in the CMB. The polarization anisotropy can be expressed as a sum of “E modes” and “B modes” (Zaldarriaga & Seljak 1997), where $E/B$ refer to curl-free and curl components of the polarization field. The primary CMB components generated by scalar perturbations have zero B modes, with the dominant primary contributions to B modes sourced by gravitational radiation.

The CMB polarization signature from an inhomogeneous distribution of free electrons during and after reionization has two components: that due to differential screening and that due to rescattering (Dvorkin et al. 2009). The differential screening term arises because modulation of the CMB polarization at recombination by a direction-dependent $e^{-\tau(\hat{n})}$ generates additional $B$ and $E$ modes. To first order in the anisotropic part of the depth, the induced $B$ modes are given by Dvorkin et al. (2009) as

$$C^{BB}_{\ell} = e^{-2\tau} C^{EE}_{\ell} \sum_{\ell'\ell''} C^{\tau\tau}_{\ell} C^{E\ell'}_{\ell},$$

where $C^{\tau\tau}_{\ell}$ is the power spectrum of the optical depth fluctuations on the sky, $C^{EE}_{\ell}$ describes the $E$ modes at recombination, $\tau$ is the isotropic part of the optical depth, and $o_{\ell\ell'\ell''} = 1$ for $(\ell + \ell' + \ell'')$ odd, and zero otherwise.

The other contribution is due to the scattering of incident quadrupole radiation on free electrons, which generates both $E$- and $B$-mode polarization if the scatterers are inhomogeneous. On small scales (where the spherical curvature of the sky can be ignored), the $B$ modes from rescattering trace the angular fluctuations in the optical depth on the sky $C^{\tau\tau}_{\ell}$ as (Hu 2000; Dvorkin & Smith 2009)

$$C^{BB}_{\ell}(\text{scatt}) = C^{EE}_{\ell}(\text{scatt}) \simeq \frac{3}{100} Q_{\text{rms}}^2 e^{-2\tau_{\text{eff}}} C^{\tau\tau}_{\ell},$$

where $Q_{\text{rms}}$ is the rms quadrupole, which we take to be 22 $\mu$K, and $\tau_{\text{eff}}$ is a characteristic optical depth accounting for the low redshift screening of the generated anisotropies.

Although a primordial baryon–CDM isocurvature mode could in principle take any form, with the details set by the physics of how it is generated, a natural choice is the scale-invariant $k^3 P_{\delta}(k) \propto k$ constant, where $P_{\delta}(k) = (\delta_{\text{void}}(k) / \delta_{\text{map}}(k))$. The integral along the line of sight selects a plane of modes in the three-dimensional (3D) Fourier space. This is simply a slice through the original 3D Fourier space, so the statistics of the selected modes are unchanged (the “Limber approximation”; Kaiser 1992), and $P_{\delta}(k)$ maintains the same scaling with wavenumber after projection. Thus, we assume $C^{\tau\tau}_{\ell} \propto \ell^{-3}$.

Figure 4 shows the scattered and the screening $B$-mode polarizations for a large total fractional optical depth fluctuation of 20%. For the rescattering, we show only the result for $\ell > 50$; lower multipoles require a more sophisticated treatment to include the effects of the spherical geometry.

Both the rescattered and the screened $B$ modes in this scenario are well below the $B$ modes induced from E modes through gravitational lensing by foreground structure. However, lensing estimation techniques could reduce the lensing $B$-mode noise by factors of at least 40 (Seljak & Hirata 2004), making the rescattered $B$ modes potentially observable on intermediate angular scales for the $\tau$ fluctuations we have assumed. Rescattered $B$ modes could consequently be a contaminant in the search for primordial gravitational waves using CMB polarization, if the tensor-to-scalar ratio is sufficiently low.
Figure 4. B-mode CMB polarization anisotropy induced by a reionized region containing a power-law baryon inhomogeneity with optical depth fluctuations of 20%. The B modes induced due to the rescattering effect (dot-dashed blue curve) trace the optical depth fluctuations with a constant prefactor in the flat-sky regime, and can become significant at intermediate multipoles. The B modes due to the patchy screening effect (green solid curve) remain well below the lens-induced B modes (red dotted curve), even when the lensed modes are reduced by a factor of 40, as motivated by delensing studies (red dashed curve). The magenta curves show inflation-induced B modes with tensor-to-scalar ratios of 0.1 and 0.001 (dash-triple-dotted and long dashed, respectively).

(A color version of this figure is available in the online journal.)

6. FUTURE TESTS

With current CMB data, it should be possible to look for variations in the baryon fraction at recombination. One could either divide the sky into many segments and analyze each section independently (Hansen et al. 2009) or use a quadratic estimator (as was done for lensing detection in Smith et al. 2007) to map out baryon fractions on the sphere.

Large-scale structure tests are a potentially powerful probe of variations in the baryon fraction. It has been shown that fluctuations in the high-redshift hydrogen 21 cm intensity could constrain variations at the $10^{-3}$ level (Gordon & Pritchard 2009). Large-scale structure studies can also constrain large-scale variations in the baryon fraction through baryonic effects on the matter transfer function. One could use galaxy surveys or QSO counts (Hirata 2009) to put limits on large-scale variations in the matter transfer function. This will be complicated by the possibility of the baryon fraction fluctuation being correlated (or anti-correlated) with the potential fluctuations, but it provides strong constraints. There is an additional complication in that one needs to understand how the astrophysics depends on the baryon fraction. For example, the number density of high-redshift quasars (Hirata 2009) is very sensitive to the matter power spectrum, but it could also be extremely sensitive to the baryon fraction. Cooling processes often scale as density squared, so formation times could be strongly affected by a varying baryon density.

Assuming that the varying baryon fraction does not affect the astrophysics, we can translate the results of Hirata (2009) into limits on baryon fraction fluctuations. Baryons affect the matter power spectrum (Eisenstein & Hu 1998) such that a 1% variation in $\sigma_8$ would be caused by a variation in the baryon fraction on the order of 5%. The constraints of Hirata (2009) on large-scale gradients in the high-redshift QSO number density (horizon scale gradients in $\ln \sigma_8$ less than 0.03 at 99% confidence) are thus comparable to the galaxy cluster gas fraction constraint.

Significant improvement on the intrinsic scatter D/H is possible with a larger sample of QSO absorption systems. Deuterium is an excellent measure of baryon density; with a modest improvement on current samples, it is likely that D/H will be the most robust constraint on large-scale variations in the baryon fraction.

With future radio interferometers such as the Square Kilometer Array, direct measurements of D/H and $^3$He/H will be possible over a cosmological volume through radio lines (Sigurdson & Furlanetto 2006; McQuinn & Switzer 2009; Bagla & Loeb 2009). The differential Thomson optical depth can, in principle, be determined from maps of the neutral hydrogen density based on redshifted 21 cm emission or absorption using the same interferometers; very large-scale variation will be extremely difficult to separate from foreground emission.

7. DISCUSSION AND CONCLUSION

We have shown that large-scale variation in the cosmological baryon fraction is constrained, although not strongly, by non-CMB observations. Scatter in observed light-element abundances constrains cosmological variations in the baryon density to be less than 26% (95% confidence). Variations are constrained to be less than 8% (95% confidence) by galaxy cluster measurements, provided there are no selection effects at work that lead to an artificially low-observed scatter in the gas fraction compared to the cosmic distribution.

Within these bounds of variation, there is an opportunity to ease the tension between the observed Li and the standard BBN calculation. In addition, there is evidence of an asymmetry in the large-scale CMB fluctuations which could be caused by differential screening effects. However, in both cases the observed effects appear to require fluctuations that are much larger than allowed by galaxy cluster measurements and well above the upper limits from BBN. It is possible that our Galaxy is in an especially baryon-deficient part of the universe, and that there is an anomalously large baryon asymmetry on scales of the comoving distance to $z \sim 7$, but this would require fine tuning.

Variations in the baryon fraction constitute a mechanism for generating B modes in CMB polarization; a detection of B modes on large scales would therefore not be an unambiguous detection of primordial gravitational waves. The existence of such large-scale baryon fraction modulation would be most simply explained through an inflation-like mechanism, as there would be super-horizon correlations in the large-scale baryon fraction.

In general, cosmological variations in the baryon fraction (baryon–CDM isocurvature modes) are allowed at a level that could be important for galaxy and active galactic nucleus formation and evolution, and could also be a foreground for future CMB experiments searching for the signature of primordial gravitational radiation. Detection of such a variation would be extremely interesting for fundamental physics, as well.

We thank the Kavli Institute for Cosmological Physics in Chicago, where a significant fraction of this work was done, and Lloyd Knox, Kendrick Smith, Wayne Hu, and Olivier Dore for useful discussions. We acknowledge support from the NSERC Discovery Grant program, the Canadian Institute for Advanced Research Cosmology & Gravity program, the Canada Research Foundation Cosmology & Gravity program, the Canada Research Foundation Cosmology & Gravity program.
