Lyman α forest–CMB cross-correlation and the search for the ionized baryons at high redshift

Rupert A. C. Croft,^{1,2*} A. J. Banday² and Lars Hernquist³

¹Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

²Max Planck Institute for Astrophysics, 85741 Garching, Germany

³Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Accepted 2006 March 9. Received 2006 March 8; in original form 2005 December 13

ABSTRACT

The intergalactic neutral hydrogen, which is responsible for the Ly α forest of quasar absorption lines, is a tracer of much larger amounts of ionized hydrogen. The ionized component has yet to be detected directly, but is expected to scatter cosmic microwave background (CMB) photons via the Sunyaev–Zel'dovich effect. We use hydrodynamic simulations of a Λ cold dark matter (CDM) universe to create mock quasar spectra and CMB sky maps from the same volume of space. We find that the high-redshift Ly α forest gas causes temperature fluctuations of the order of 1 µK rms in the CMB on arcmin scales. The kinetic and thermal Sunyaev-Zel'dovich effects have a similar magnitude at redshift 3, with the thermal effect becoming relatively weaker as expected at higher redshift. The CMB signal associated with lines of sight having H I column densities $> 10^{18}$ cm⁻² is only marginally stronger than that for lower column density sightlines. There is a much more significant dependence of rms temperature fluctuation on mean Ly α absorbed flux, however, suggesting that the CMB signal effectively arises in lower density material. We investigate the extent to which it is possible to cross-correlate information from the Ly α forest and the microwave background to detect the Sunyaev–Zel'dovich effect at redshifts 2-4. In so doing, we are able to set direct limits on the density of diffuse ionized intergalactic baryons. We carry out a preliminary comparison at a mean redshift z = 3 of 3488 quasar spectra from Sloan Digital Sky Survey (SDSS) Data Release 3 and the WMAP first year data. Assuming that the baryons are clustered as in a ACDM cosmology, and have the same mean temperature, the cross-correlation yields a weak limit on the cosmic density of ionized baryons $\Omega_{b,I}$. As a fraction of the critical density, we find $\Omega_{b,I} < 0.8$ at 95 per cent confidence. With data from upcoming CMB telescopes, we anticipate that a direct detection of the high-redshift ionized intergalactic medium will soon be possible, providing an important consistency check on cosmological models.

Key words: cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

The Ly α forest seen in quasar spectra is now thought to be produced by neutral hydrogen atoms in an intergalactic medium (IGM) which is in photoionization equilibrium [see e.g. Rauch (1998) for a review]. At low to moderate redshifts, the fraction of neutral atoms is small (1 in ~10⁶; e.g. Gunn & Peterson 1965), but the knowledge of the photoionization rate [from summing the contribution of known sources; e.g. Haardt & Madau 1996] makes it possible to extrapolate from this one part in a million measurement to an estimate of the total baryon density Ω_b (Rauch et al. 1997; Weinberg et al. 1997). Loeb (1996) has argued that the ionized component of the IGM should be directly observable by inverse Compton or Doppler scattering of cosmic microwave background (CMB) photons by free electrons [Sunyaev–Zel'dovich (1972, 1980), hereafter the SZ effect]. Such a detection would provide consistency checks on the currently favoured cosmological model for structure formation, and place constraints on the IGM density, temperature, ionization state and velocity field. In this paper, we use a hydrodynamic cosmological simulation to investigate how observations of the CMB temperature and Ly α spectra of quasars drawn from the same region of sky can be used in the search for the hitherto unseen 99.9999 per cent of the hydrogen associated with the Ly α forest gas at redshifts $z \approx 2-5$.

^{*}E-mail: rcroft@cmu.edu

In favoured models for structure formation, the optical depth for Ly α forest absorption arises in a continuously fluctuating medium (e.g. Cen et al. 1994; Zhang, Anninos & Norman 1995; Petitjean, Muecket & Kates 1995; Hernquist et al. 1996; Katz et al. 1996a; Wadsley & Bond 1997; Theuns et al. 1998; Davé et al. 1999). The density of neutral hydrogen is inversely proportional to the photoionization rate Γ :

$$\Gamma = \int_{\nu_{\rm HI}}^{\infty} d\nu \frac{4\pi J(\nu)}{h\nu} \sigma_{\rm HI}(\nu), \qquad (1)$$

and the optical depth to absorption is given by

$$\tau_{\rm Ly\alpha} = \frac{\pi e^2}{m_e c} f \lambda H^{-1}(z) n, \qquad (2)$$

where f = 0.416 is the Ly α oscillator strength, *n* is the neutral hydrogen number density and $\lambda = 1216$ Å (Gunn & Peterson 1965). In order for τ be of order unity (as seen, for example, at redshifts $z \sim 3$), *n* must be much smaller than the total density of hydrogen atoms, as noted earlier. The gas in photoionization equilibrium lies at densities $\sim 1-10$ times the mean, and obeys a power-law relationship between density and temperature (Gnedin & Hui 1998). At z = 3, according to hydrodynamic simulations (Davé et al. 2001), more than 90 per cent of the baryons are expected to be in this diffuse component. The tiny neutral fraction is readily detectable, but not the dominant ionized portion of the material. Owing to gravitational evolution, at z < 2 most of the baryons are expected to reside in a shock-heated IGM, with little neutral hydrogen (Cen & Ostriker 1999; Davé et al. 2001).

The interaction of CMB photons with free electrons in the intergalactic plasma was first considered by Sunyaev & Zel'dovich (1972). Inverse Compton scattering preferentially increases the energy of CMB photons, while conserving photon number, leading to a spectral distortion whose amplitude is proportional to the product of electron temperature and density (the thermal SZ effect, tSZ). Doppler scattering induces an intensity fluctuation with the same spectral shape as the CMB itself (the kinetic SZ effect, kSZ). The thermal effect is one of the main sources of secondary CMB anisotropies on small angular scales.

The perturbation in the CMB thermodynamic temperature resulting from scattering of non-relativistic electrons is

$$\frac{\Delta T}{T} = y \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) \tag{3}$$

$$\simeq -2y$$
 for $x \ll 1$, (4)

where $x = h\nu/kT_{\text{CMB}} \simeq \nu/56.85$ GHz is the dimensionless frequency, and the second expression is valid in the Rayleigh–Jeans limit, which we assume henceforth. The quantity *y* is known as the comptonization parameter and is given by

$$y \equiv \int dl \frac{n_e k (T_e - T_{\rm CMB})}{m_e c^2},$$
(5)

where the integral is performed along the photon path.

The kinetic SZ effect arises from the motion of ionized gas with respect to the rest frame of the CMB. The resulting temperature fluctuation is $\Delta T/T = -b$ where

$$b \equiv \sigma_T \int \mathrm{d}l n_e \frac{v_r}{c} \tag{6}$$

gives the magnitude of the effect and v_r is the component of the gas peculiar velocity along the line of sight (positive sign for receding gas and negative for approaching) to the observer.

The gas which causes $Ly\alpha$ forest absorption has a density near the cosmic mean, with a volume-weighted temperature around $2 \times$ 10^4 K at z = 3 (e.g. Schaye et al. 2000; McDonald et al. 2001). At this temperature, we expect the contribution from the kSZ to dominate the rms temperature fluctuations of the CMB from the Ly α forest. Using observed Ly α forest line counts, Loeb (1996) estimated that the rms temperature fluctuations from the kSZ would be $\Delta T/T \sim 10^{-6}$ for the Ly α forest lying between z = 2 and 5. This is for angular scales of the order of 1 arcmin, with a θ^{-1} decrease on larger scales. There is much competition from other signals on these small scales [including low-redshift thermal SZ from galaxy clusters, the Ostriker & Vishniac (1986) effect, patchy re-ionization (see, e.g. McQuinn et al. 2005) and foreground sources]. The suggestion to cross-correlate information from Sloan Digital Sky Survey (SDSS) Ly α forest spectra and the CMB was also made by Loeb (1996) in order to try to better extract the signal. In this paper, we measure the kSZ effect in simulations, as well as the thermal effect. The mass-weighted temperature of the IGM at $z \sim 3$ is expected to be closer to 10⁶ K (e.g. Springel, White & Hernquist 2001a), and so the thermal effect should be significant. How much of that hot gas is physically associated with $Ly\alpha$ absorption will affect how well the tSZ can be detected by cross-correlation.

We note that both being absorption phenomena, neither the SZ signal nor the Ly α forest opacity is directly affected by the inverse square law. The problem of finding bright quasi-stellar object (QSO) background sources aside, they therefore both have an advantage in the hunt for baryons at high redshifts. Indeed, the value of the SZ effect for finding galaxy clusters at the highest redshift has long been recognized [see e.g. Carlstrom, Holder & Reese (2002) for a review]. We aim to explore its potential for finding diffuse ionized material at z > 2.

The plan of this paper is as follows. In Section 2, we give details of the hydrodynamic cosmological simulation and how mock $Ly\alpha$ spectra and CMB temperature sky maps were constructed from the same volume of space. In Section 3, we compute statistical measures of the absorption and examine correlations between the two probes of the IGM. In Section 4, we directly compare observational data from the SDSS and *WMAP*, and in Section 5 we summarize and discuss our results.

2 SIMULATED SPECTRA AND SKYMAPS

2.1 Hydrodynamic simulation

We employ a hydrodynamical simulation of the popular A cold dark matter (CDM) cosmology to make our prediction of the Ly α and the SZ effects. In order to resolve both the Ly α features and the bulk velocities responsible for the kSZ effect, we require a large simulation volume with good mass resolution. The data snapshots we use are from the simulation 'G6', also used by Nagamine et al. (2005) to which we refer the reader for more details, which describes a representative cubic volume of space $100 h^{-1}$ Mpc on a side, in a cosmological constant-dominated CDM universe, consistent with measurements from the WMAP satellite (Bennett et al. 2003; Spergel et al. 2003). The relevant parameters are $\Omega_{\Lambda} = 0.7, \Omega_{m} = 0.3, \Omega_{b} =$ 0.04 and a Hubble constant $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The initial linear power spectrum is cluster-normalized with a linearly extrapolated amplitude of $\sigma_8 = 0.9$ at z = 0. The simulation was performed with the smoothed particle hydrodynamics (SPH) code GADGET-2 (Springel, Yoshida & White 2001b; Springel 2005), which is based on the entropy-conserving approach of Springel & Hernquist (2002), and is an extension of the 'G-series' runs of Springel & Hernquist (2003b).

The baryonic matter is initially represented by 486^3 SPH particles, and the collisionless dark matter by 486^3 *N*-body particles. The initial mass per gas particle is therefore 9.7×10^7 M_☉ and dark matter 6.3×10^8 M_☉. The gravitational softening length is $5 h^{-1}$ kpc. The calculation includes gas dynamics, cooling, a multiphase treatment for star formation (see Springel & Hernquist 2003a) and a uniform background of ionizing radiation in the manner of Katz, Weinberg & Hernquist (1996b), renormalizing the Haardt & Madau (1996) spectrum to reproduce the observed Ly α forest mean optical depth. A study of the SZ effect using this code was carried out by Springel, White & Hernquist (2002), who noted that the additional physics beyond adiabatic gas dynamics had a relatively small effect; changing, for example, the mean comptonization by 20 per cent.

2.2 Mock Ly α spectra

We use simulation outputs from redshifts z = 2, 3, 4, 5 and 6 to make mock Ly α spectra. For each output, this was done in the usual manner, by integrating through the SPH kernels of the particles to obtain the neutral hydrogen density field, and then convolving with the line of sight velocity field (see e.g. Hernquist et al. 1996). At each redshift, we compute 64² lines of sight, equally spaced on a square grid. In order to approximate longer spectra, we join together spectra after performing a random translation, reflection and choice of axis, as is done, for example, when making skymaps by ray-tracing through multiple copies of the simulation volume (e.g. da Silva et al. 2000). As our fiducial spectrum length, we take the distance between the Ly α and Ly β emission lines. Later in the paper, when we cross-correlate the Ly α forest and CMB at a particular redshift, our analysis will refer to this $Ly\alpha$ to $Ly\beta$ region, centred on the redshift in question. For reasons of computational simplicity, we choose to ignore evolution over the length of each spectrum, both in clustering and in the mean optical depth. The Ly α to $Ly\beta$ regions correspond to comoving spatial distances of 343, 343, 330 and 314 h^{-1} Mpc when centred on redshifts z = 2, 3, 4 and 5, respectively. We add partial spectra to the full spectra to ensure the correct overall length.

For each spectrum, we compute the mean effective optical depth over its entire length, i.e.

$$\tau_{\rm eff} = -\ln\langle F \rangle,\tag{7}$$

where $\langle F \rangle$ is the mean absorbed flux in the spectrum, $\langle e^{-\tau} \rangle$. Throughout most of this paper, we will use this quantity, τ_{eff} , to determine the relative amount of Ly α absorption in each sightline. Another possible measure is to compute the total column density of neutral hydrogen along each sightline from Ly α to Ly β . Being unaffected by saturation, this will tend to reflect the presence of individual high density clouds.

As noted earlier, the simulation was run using the ultraviolet (UV) ionizing background of Haardt & Madau (1996). In our subsequent analysis, we rescale the Ly α optical depths by a small amount after computing the spectra by adjusting them slightly so that $\langle F \rangle$ averaged over all spectra reproduce the observational results of Schaye et al. (2004), $\langle F \rangle = 0.878, 0.696, 0.447$ for z = 2, 3, 4.

2.3 CMB skymaps

At each redshift, we construct maps of the Compton y parameter and the Doppler b parameter by projecting the appropriate SPH kernel weighted quantities along one axis. Unlike Springel et al. (2001a),

for example, we do not project along lines which meet at a single z = 0 observer, as we will be interested only in the SZ signal generated by gas which falls within the same redshift interval as our Ly α forest spectra. In order to compare CMB maps and spectra for the same volume of space, we randomly translate and reflect our SZ maps in the same manner so that they match up with the spectra from the previous section. We generate 40 maps, which have an angular size which subtends the whole box at each particular redshift; i.e. 94.7 arcmin at redshift z = 2, and 77.2, 68.5, 63.1 and 59.5 arcmin at z = 3, 4, 5 and 6, respectively. We make maps at a resolution corresponding to 2048² pixels, but also average the pixels so that we have 64^2 , the same number as the Ly α spectra.

2.4 Examples

Looking at kSZ maps and Ly α spectra generated from the same region of space, we can see if there are any obvious visual relationships between the CMB temperature in pixels and the appearance of the absorption features. In Fig. 1, we show the Compton *y* parameter, from the thermal SZ effect at *z* = 2, and alongside it five spectra which are taken with the sightline along the axis perpendicular to the plane of the map. Here, in both maps and spectra, we do not plot the entire Ly α to Ly β region (we do this later, below), but just the material in one single cubic simulation volume. It is apparent from the map that much of the volume contributes to a very small *y* parameter, and that most of this redshift slice has *y* < 10⁻⁹ or below. There are also a number of obvious galaxy clusters, with *y* > 10⁻⁵.

Of course, a map of the y parameter from all material between z = 0 and the last scattering surface would have quite a different appearance, with the mean y parameter being higher, of the order of 10^{-6} , and with many of the obvious filaments seen in Fig. 1 washed out by projection effects. Fig. 2 of White, Springel & Hernquist (2002) demonstrates this, using a similar simulation.

Sightlines (a)–(c) in Fig. 1 were chosen from those among the 64^2 we made which have large values of y. Spectrum (a) is visibly taken from near a cluster, seen in the SZ map. There, however, does not seem to be much evidence of this in Ly α absorption, at least in terms of saturated absorption. The other two spectra, (b) and (c), are also fairly similar, even though they are both associated with large y values. Because there is not much saturated absorption, we can see that it is unlikely that column density will correlate well with SZ decrement.

Sightlines (d) and (e) in Fig. 1 have been chosen to have values of y near the bottom of the rank-ordered list of all pixels. From the SZ plot, it can be seen that the spectra lie on sightlines though obvious voids, and the y values are 10 000 times smaller than for Panels (a)–(c). On examination, these spectra appear to have very few wide absorption features compared to spectra (a)–(c), and the overall level of $Ly\alpha$ absorption is slightly lower [particularly (e)].

Turning now to the map of the kSZ effect (Fig. 2), we see that much of this redshift z = 2 slice has a Doppler |b| value of $\sim 10^{-8}$, although there are several clusters and filaments dense enough and moving fast enough to reach $|b| = 10^{-5}$. Clusters can be seen moving towards the observer (blue) and away (red). The large coherence length of the velocity field can also be seen, from the size of the red and blue structures. Projecting the kSZ effect for a larger volume, for example subtending z = 2-5, will lead to more structures but also more cancellation of positive and negative *b* values. The general level of fluctuations likely agrees at least roughly with the estimate of Loeb (1996) (rms $b \sim 10^{-6}$).

Again, the kSZ plot can be compared to full lightcone maps (e.g. fig. 2 of Springel et al. 2001a), which show a similar pattern



Figure 1. The tSZ CMB *y* parameter (left-hand panel) for a projection of one simulation volume (side length 100 h^{-1} Mpc). We show the positions of five pixels for which we have drawn Ly α sightlines (see the text). The five spectra are shown in panels (a–e). Black curves in the right-hand panels are the mock transmission spectra, while the dashed red curves are the peculiar velocity. Sightlines (a), (b) and (c) have been chosen so that *y* is large [for (a), $y = 5 \times 10^{-6}$, for (b), $y = 1.4 \times 10^{-5}$ and for (c), $y = 3.3 \times 10^{-6}$], all among the top 1 per cent of *y*-values. Sightlines (d) and (e) have small values of *y*, $y = 5.13 \times 10^{-10}$ and 3.8×10^{-10} , respectively. Sightline (b) is the same as sightline (b) in Fig. 2. The Ly α spectra are of length 9900 km s⁻¹ (one simulation box length).



Figure 2. Projection through the simulation box showing the kSZ CMB Doppler parameter, *b* (left-hand panel). Pixels with a positive values of *b* contain material moving on average away from the observer, and negative *b* values towards the observer. We show the positions of five pixels for which we have drawn Ly α sightlines (see the text). The five spectra are shown in panels (a)–(e). Black curves in the right-hand panels are the mock transmission spectra, while the dashed red curves are the peculiar velocity. Sightlines (a) and (b) have been chosen so that *b* is large and positive [for (a), $b = 7 \times 10^{-6}$ and for (b), $b = 4 \times 10^{-6}$], both among the top 1 per cent of *b* values. Sightline (c) is the pixel with the smallest value of |b|, -8×10^{-12} . Sightlines (d) and (e) have large negative values of b, -3×10^{-6} and -6×10^{-6} , respectively. The Ly α spectra are of length 9900 km s⁻¹ (one simulation box length).

of approaching and receding material, also with less obvious filamentary structure. Sightlines (a) and (b) in Fig. 2 were chosen to have large positive Doppler *b* values, and (d) and (e) negative. It can be seen that their spectra contain fairly substantial amounts of absorption, and one can understand how the mass-weighted velocity will give rise to the *b* values seen. Panel (c) has a velocity field which cancels in the left and right halves of the spectrum, yielding a *b* parameter much closer to zero ($b \sim -10^{-11}$).

The number of sightlines passing close to galaxy groups and clusters in Figs 1 and 2 is relatively small, at least in this redshift interval. This means that any clustering statistic which weights spectra equally will tend to be sensitive to physical conditions in the regions with relatively small overdensity.

Looking at a few individual sightlines, it seems likely that there is a connection between $Ly\alpha$ forest absorption and SZ-induced temperature fluctuations. One way to examine this further would be to look at skymaps of both the $Ly\alpha$ forest and CMB. Of course, with real observations it is unlikely that we would have a dense enough grid of background quasars to make a true Ly α map, but with the simulations, we can use such a technique to examine the spatial pattern of absorption. In Fig. 3, we show examples of such maps, plotting the thermal and kinetic SZ effects alongside angular maps of the Ly α optical depth and column density. It should be noted that we plot the absolute value of the temperature change produced by the kinetic effect. Unlike the previous two figures, here we show a projection through the volume of space corresponding to an entire Ly α to Ly β region (not just one simulation box). The results are centred on redshift z = 2, and the angular size of the maps is 95 arcmin. Because we have a grid of 64^2 Ly α spectra, we have rebinned the CMB maps so that they have the same resolution (1.5 arcmin pixels).

The grey-scales of the SZ maps are the same, and so we can see from the left two panels of Fig. 3 that the thermal and kinetic effects trace similar features, with the thermal effect being more prominent



Figure 3. Sky maps of CMB temperature (left-hand panel) owing to the kSZ effect at z = 2 (for a volume equivalent to the entire Ly α to Ly β spectral width). The width of the panel is 95 arcmin. We also show |T|, mean flux $\langle F \rangle$ in Ly α spectra and total column density in Ly α spectra. In each map, there are 64×64 pixels.

at this redshift. Because of the differing frequency dependence of the two SZ effects, cross-correlation of the two maps may eventually be possible using observational data, as suggested, for example, by da Silva et al. (2000). In the present context, we are interested in the potential cross-correlation of the Ly α forest spectra and the SZ effects. In the third panel of Fig. 3, we show the mean effective optical depth, given by equation (7) for spectra through the centre of each pixel. It can clearly be seen that the patches of absorption largely follow the SZ temperature well, although there are several regions with large differences. The Ly $\alpha \tau_{eff}$ fluctuations have a somewhat lower contrast, varying by about a factor of 2 across the map. The obvious groups and clusters in the SZ maps do not show up as prominently in the Ly α map. This is likely to at least partly owe to the fact that regions with large amounts of Ly α absorption are saturated.

This is not the whole story, however, as we can see by looking at the rightmost panel of Fig. 3. Here, we show the H_I column density in each spectrum integrated from Ly α to Ly β . There is much smaller scale structure present, as the dense knots of neutral hydrogen are picked out. The range of column densities varies from $N_{\rm H_I} = 10^{14}$ to 10^{19} cm⁻². The structures in $N_{\rm H_I}$ do not correlate well with the CMB maps. The smooth, large-scale structures in the SZ effect maps (and Ly $\alpha \tau_{\rm eff}$) are not readily seen. This observation will guide our use of $\tau_{\rm eff}$ as the main Ly α statistic to be used.

In Fig. 4, we show a scatter plot of the Ly $\alpha \tau_{eff}$ against the thermal SZ temperature decrement, at redshifts 2, 3 and 4, again from maps with 64² pixels (1.5, 1.2 and 1.1 arcmin size for z = 2, 3, 4). The correlation between the two quantities is apparent, and in particular at each redshift there are few lines of sight with very low values of τ_{eff} but substantial SZ decrements. There is a large amount of scatter, however, indicating that even if foreground/background noise is not an issue, it will be necessary to average over many sightlines to obtain statistically significant results.

3 STATISTICAL MEASURES

In this section, we use the set of skymaps and spectra generated from our simulation to statistically examine the CMB fluctuations caused by $Ly\alpha$ forest gas.



Figure 4. Thermal SZ temperature decrement T_{tSZ} versus Ly α forest τ_{eff} for spectra at three different redshifts.

3.1 Rms temperature fluctuations

In Fig. 5, we show the rms temperature fluctuations averaged in top-hat circles on the plane of the sky, calculated from maps such as those shown in the left two panels of Fig. 3. We show results centred on integer redshifts between z = 2 and 6, in each case, calculating only the CMB fluctuations caused by gas within the Ly α to Ly β region of the spectrum. The thermal and kinetic SZ effects are shown separately, and it can be seen that at z = 2, $T_{\rm rms}$ is approximately three times larger for the tSZ effect, as we expect from Fig. 3. At $\theta = 1$ arcmin, the rms *T* fluctuations are $\sim 2 \,\mu$ K for the tSZ at this redshift.

The curves for all redshifts have an angular dependence given by $\sim \theta^{0.6}$. At z = 3, the two SZ effects are comparable, and beyond that the kSZ effect stays approximately constant out to z = 5. The



Figure 5. The rms CMB temperature fluctuations in top-hat circles as a function of radius (*x*-axis) for three different redshifts. In each case, we plot the rms fluctuations associated with a volume which would contain material with H I absorption wavelengths from Ly α to Ly β centred on the indicated redshift. Thin lines are kinetic SZ and thick lines thermal SZ.

 Table 1. The mass density and electron number weighted gas temperature and velocity dispersion in the simulation at different redshifts.

Ζ	<i>Τ</i> _ρ (K)	Т _е (К)	$v_{\mathrm{rms},\rho}$ (km s ⁻¹)	$v_{\mathrm{rms},e}$ (km s ⁻¹)
6	7.1×10^{3}	1.9×10^{5}	136	166
5	2.57×10^{4}	2.60×10^{4}	149	148
4	6.86×10^{4}	6.89×10^{4}	164	164
3	1.77×10^{5}	1.76×10^{5}	185	185
2	5.04×10^{5}	5.04×10^{5}	214	214

rms peculiar velocity of the gas increases slightly with decreasing redshift, as detailed in Table 1. This is balanced by the decrease in angular size for a given comoving length which shifts the curves to the left. The main change in the kSZ at z = 6 is caused by the switching on of the UV ionizing background in our model, and so the ionized fraction is small then. The temperature of the gas influences the fluctuations owing to the thermal effect directly, and this has a much more rapid dependence on redshift. In Table 1, we can see that at z = 6, the electron-weighted mean temperature is $> 10^5$ K, as the free electrons are concentrated in the few collisionally ionized

structures present at that redshift.¹ By z = 5, when most of the IGM is photoionized, the electron-weighted temperature is 2.6×10^4 K, and as gravitational evolution continues, shock heating increases the temperature by another factor of 20 by z = 2. The rise in $T_{\rm rms}$ owes mostly to this, with a small part coming from the factor of 2.3 increase in linear density fluctuations themselves between the two redshifts.

We note that although the thermal SZ effect is dominant at redshifts z < 3, the signal that is actually generated in gas that directly produces the Ly α forest is a small fraction of the total thermal SZ signal at that redshift. This is because the temperature of Ly α forestproducing gas at the mean density at z = 3 is $\sim 10^4$ K, low enough that the thermal SZ effect should be extremely small (e.g. Loeb 1996). The thermal SZ signal we are instead seeing is largely generated in the Warm–Hot Intergalactic Medium (WHIM; e.g. Davé et al. 2001) at these redshifts which is collisionally ionized, and is responsible for the higher mass-weighted temperature seen in Table 1. As we discuss below, the thermal SZ signal at this redshift is strongly correlated with the Ly α absorption signal, indicating that the WHIM signatures are physically associated with the same large-scale structures as the Ly α forest, tracing their higher density regions.

3.2 Cross-correlation function

We have seen that IGM gas at redshifts z = 2-5 yields rms CMB temperature fluctuations on the order of a few μ K. One could calculate the angular power spectrum, the C_1 s, and then see whether it could be separated from other small angle fluctuations. Owing to the multitude of other possible signals, however, it is likely to be more useful to cross-correlate with material at a known redshift, and use this to extract information. As the Ly α forest contains most of the baryons at these redshifts, and the mean effective Ly α optical depth appears to correlate visually with CMB temperature in Fig. 3, we use this quantity in our cross-correlation.

We compute the cross-correlation separately for the tSZ and kSZ effects. For the kSZ, because the magnitude of the effect can be positive or negative, we work with $|T_{tSZ}|$. Although this is not easily done with observational data, in the current section we are interested in making a relative comparison between the magnitude of the tSZ and kSZ effects, where this is a useful approach. In order to do the cross-correlation for each Ly α spectrum, we first calculate $\Delta \tau_{eff} = -(\tau_{eff} - \langle \tau_{eff} \rangle)/\langle \tau_{eff} \rangle$, where $\langle \tau_{eff} \rangle = -\ln \langle F \rangle$ and we use the average mean transmitted flux for all spectra. The cross-correlation we compute is $\langle \Delta \tau_{eff} \Delta T \rangle$, which has units of μ K. For the kSZ, we use $\Delta T = |T| - \langle |T| \rangle$ and for the tSZ, $\Delta T = T - \langle T \rangle = T$. The results for redshifts z = 2 - 4 are shown in Fig. 6. We have not rebinned the CMB temperatures to the angular resolution of the spectrum grid, so that we are able to show results down to angular scales below 1 arcmin.

Unlike the case for the rms *T* fluctuations, one can see that the cross-correlation statistic for the kSZ effect depends on redshift, driven by evolution in the Ly α forest. On first inspection, this is somewhat puzzling, as Ly $\alpha \tau_{eff}$ fluctuations should be increasing towards high redshift, as the mean optical depth increases, but instead they are smaller. This is because we have divided out $\langle \tau_{eff} \rangle$, and so the fractional fluctuations are indeed smaller. Measurements

¹ We ignore any difference between the electron and ion temperatures (see Yoshida, Furlanetto & Hernquist 2005).



Figure 6. Cross-correlation of |T| (for kSZ) or T (for tSZ) and $F/\langle F \rangle$ as a function of pixel–pixel distance, for three different redshifts. Results for tSZ are shown in panel (a) and kSZ in panel (b).

based on the flux F, instead divide out $\langle F \rangle$ which is smaller at high redshift (see e.g. Croft et al. 1998).

This statistic and its spatial dependence could in principle be measured observationally for the thermal SZ effect. We will return to this point in Section 5.

3.3 One-dimensional distributions

We have seen from Fig. 4 that there is a dependence of SZ decrement and the Ly $\alpha \tau_{eff}$ for spectra in the same pixel. We will average this relationship over many pixels and bin as a function of τ_{eff} , in order to visualize the mean trend of T versus τ_{eff} . This simple statistic is one on which we will concentrate on in this section, and use in Section 4 below to compare to observational data. The mean trend of T versus τ_{eff} does depend on the resolution of the CMB map, and we present results for the same resolution as the spectrum grid (64² pixels, each of width 1.2 arcmin at z = 3) as well as smoothing with a Gaussian filter with full width at half-maximum (FWHM) 13 arcmin (the *WMAP W*-band resolution; see Bennett et al. 2003). We will also plot the rms temperature fluctuations in pixels caused by the kinetic SZ effect.

Fig. 7 shows our results for three different redshifts, with the top panels being the tSZ decrement. We can see that the overall level of the temperature decrement is strongest for z = 2 and weakest for z = 4, as expected. At z = 2, pixels with Ly $\alpha \tau_{\text{eff}} = 0.07$ have

tSZ temperature decrements of ~0 μ K, and the pixels with $\tau_{eff} = 0.2$ have decrements ~ -1.5μ K. Between these two extremes, the trend of T with τ_{eff} is close to linear. The Ly α spectra with the very highest levels of absorption, $\tau_{eff} \gtrsim 0.2$, exhibit a weakening of the decrement, likely owing to radiative cooling of the gas in these high-density regions producing relatively more neutral hydrogen. The signal smoothed at the *WMAP* resolution is somewhat weaker, but still easily measurable from the simulations.

In the middle panels of Fig. 7, we show the kSZ-induced rms temperature fluctuations. They exhibit a rise with increasing τ_{eff} , in accordance with the rise in density probed by the sightlines, again with a turndown for the spectra with the very highest absorption levels. As with the rms in circles plotted in Fig. 5, the level of fluctuations changes little with redshift. Smoothing the map to the *WMAP* resolution does affect the level strongly, however, reducing the magnitude of the signal by roughly a factor of 10. As rms fluctuations are already more difficult to measure in the presence of noise than the mean level, this means that it is likely that the thermal rather than the kinetic SZ effect will be easiest to target when considering the analysis of observational data.

One can ask whether the signal from the tSZ and kSZ effects arises in low- or high-density gas, and whether we are really sampling the low-density IGM probed by the Ly α forest. Because our measure of the τ_{eff} -T correlation is angular pixel weighted (we are using a histogram of pixel values), the fact that most of the sightlines we use will contain relatively little gas comes into play. For example, only a few of the pixels in Fig. 3 pass close to clusters and groups, and most of the spectra with the lowest 90 per cent of τ_{eff} values are likely to pass through only void-like regions in the Ly α to Ly β interval. This means that while the thermal SZ does directly probe the massweighted (or rather electron number-weighted) IGM temperature, and so is dominated by virialized objects, by breaking up the signal as a function of τ_{eff} , we introduce sensitivity to what is occurring in low-density regions.

In Fig. 3, we saw that there appeared to be a much better correspondence between CMB temperature and τ_{eff} than H I column density. Nevertheless, from looking at the maps it was not immediately obvious that the very highest density objects are in fact associated with large amounts of HI and have a significant SZ signal. In Fig. 8, we plot CMB T versus $N_{\rm H_{I}}$, and see that there is a slight increase in the SZ signal for sightlines with $N_{\rm H_{I}} = 10^{14} - 10^{16} \text{ cm}^{-2}$, for all redshifts. The curves, however, are relatively flat at higher column densities. It appears that information from saturated regions will not be helpful in carrying out Lya forest CMB correlations, and so there is little need for high-resolution spectra. In order to shed further light on this, in Fig. 9 we have plotted values of $N_{\rm HI}$ versus τ_{eff} for individual sightlines, as a scatter plot (as always, we average over the Ly α to Ly β region in each spectrum). We expect there to be an envelope, a minimum $N_{\rm H1}$ for sightlines with a given $\tau_{\rm eff}$. This minimum level would correspond to a uniform distribution of absorbing material (see e.g. Weinberg et al. 1997). We can see such an envelope, but also a large scatter towards higher values of $N_{\rm H_{I}}$. The fractional scatter in $N_{\rm H_{I}}$ increases for larger $\tau_{\rm eff}$, so it is not surprising that for high τ_{eff} there is little evidence of a trend in Fig. 8.

4 LIMITS FROM SDSS AND WMAP DATA

4.1 Method

The micro-Kelvin fluctuations produced by the $Ly\alpha$ forest gas are likely too small to be detected directly in current data (e.g.



Figure 7. Top panels: dependence of mean temperature from the thermal SZ effect on $\tau_{eff} = -\log \langle F \rangle$ for spectra running through the pixels. Middle panels: dependence of rms temperature fluctuation in pixels (1.5, 1.2 and 1.1 arcmin size for z = 2, 3, 4) on $\tau_{eff} = -\log \langle F \rangle$ for spectra running through the pixels, for three different redshifts (left- to right-hand side). The round symbols, which make up the upper curve in each case, show results for pixels, unsmoothed and the triangles for maps smoothed with a Gaussian filter of FWHM 13 arcmin, similar to the *WMAP W*-band resolution. Error bars are from the scatter between 10 fields in each case (size 95, 77, 69 arcmin for z = 2, 3, 4). Bottom panels: the pdf of $\tau_{eff} = -\log \langle F \rangle$ for spectra from Ly α to Ly β .



Figure 8. Top panels: Dependence of rms temperature fluctuation in pixels (same size as Fig. 7) on the neutral hydrogen density, $N_{\rm H_{I}}$, for spectra running through the pixels, for three different redshifts (left- to right-hand side). The round symbols which make up the upper curve in each case show results for pixels, unsmoothed, and the triangles for maps smoothed with a Gaussian filter of FWHM 13 arcmin, similar to the resolution of the *WMAP* satellite. Error bars are from the scatter between 10 fields in each case (size 95, 77, 69 arcmin for z = 2, 3, 4). Bottom panels: the pdf of $N_{\rm H_{I}}$ for spectra from Ly α to Ly β .

Bennett et al. 2003; Kuo et al. 2004). However, it is possible that by comparing observational statistics with simulation results we can come up with interesting limits on the population of ionized baryons, as well as estimating the properties of the observations necessary to make a direct detection. Based on our investigations in the previous sections, a good strategy will be to measure τ_{eff} for the Ly α forest as well as the CMB temperature in pixels where there are Ly α spectra. For the present limited observational study, we choose to investigate the one-dimensional distributions of these statistics, as in Section 3.3. We do this for reasons of simplicity, and because breaking up the signal into sightlines with high- and low-absorption levels will enable us to look for an obvious distinct signature (unlike a full spatial cross-correlation analysis which would have more discriminating power, but be more difficult to interpret).

In our analysis, we calculate the mean and rms CMB temperature for pixels binned using the τ_{eff} of the spectra that pass through



Figure 9. $N_{\rm HI}$ versus $\tau_{\rm eff}$ for spectra at three different redshifts.

them. We use a jackknife estimator (e.g. Bradley 1982) to compute the error bars, breaking up the data into 50 subsamples. We have also computed error bars by randomly rotating the CMB data with respect to the $Ly\alpha$ forest data and find them to be broadly consistent.

4.2 Ly α forest data

We use the Ly α forest of quasar spectra taken from Data Release 3 of the SDSS (Abazajian et al. 2005). Our data sample consists of 3488 quasar spectra with quasar redshifts greater than 2.4, pruned to remove Broad Absorption Line quasars. The particular data set was compiled from the publically available data by Scott Burles and kindly made available to us (Burles, private communication). Owing to the relatively high low-redshift limit for the quasars, the spectra consist of the full Ly α to Ly β region of the forest, lying above the atmospheric cut-off at ~3200 Å. This will be useful in order to be consistent with what was done in the simulations.

The one-dimensional distribution of τ_{eff} requires that the data be divided into broad bins (we will use five bins). In order to calculate $\tau_{\rm eff}$ for each spectrum, we need to fit the continuum over the full Ly α to Ly β region. Historically, the averaged value of the mean transmitted flux from many spectra has been used to calculate a global value of τ_{eff} (e.g. Kirkman et al. 2005). Fluctuations in τ_{eff} for entire spectra have not been estimated, however, presumably owing to problems with calculating a mean value on large scales where continuum fluctuations are seen to be comparable to fluctuations in the absorption [see e.g. the power spectrum of the raw spectrum in Hui et al. (2001) or McDonald et al. (2004)]. The largest scale on which fluctuations intrinsic to the Ly α forest have been measured is 153 Mpc at redshift z = 1.9 by Tytler et al. (2004), who found that Lyman limit systems and metal lines make large contributions to fluctuations in the mean flux at this relatively low redshift. In our case, we will be working at z = 3, where these effects should be less important. Nevertheless, we expect that there will be a large contribution to the variation in τ_{eff} from sightline to sightline owing to continuum estimate errors as well as noise in the spectra and non-Ly α absorption. We leave the question of how best to fit the

continuum on the largest scales to other work (e.g. Suzuki et al. 2005), and in the meantime, we will add a Gaussian distribution of errors to our simulation measurement to account for these effects when comparing our simulations with observations. The relatively large bin size should serve to mitigate effects of τ_{eff} errors in this first illustrative comparison.

We fit the continuum using the often-employed polynomial fitting method (e.g. Croft et al. 1998). A low-order polynomial (third order in our case) is fit to the Ly α forest region. We avoid the Ly α and Ly β emission lines by restricting our fit to the rest wavelength range 1030–1190 Å. We then remove all points more than 2σ below the continuum and refit, iterating this procedure until convergence is achieved. Averaging over all Ly α pixels in our data set, we find a mean redshift z = 2.91. At this stage, we divide out the continuum and compute a value of F. For all spectra, we find F = 0.8. Measurements of the mean flux and global value of τ_{eff} from higher resolution data sets better suited to this measurement find a lower value of F (e.g. Schaye et al. 2004; Kirkman et al. 2005). Our measurement of F is on average biased high because we miss absorption in flatter low-density regions around the continuum owing to our low resolution. We correct for this bias in a crude way by multiplying the value of the fit continuum in each case by 1/0.85. This simple correction is adequate for our current purposes, having been set so that we reproduce the global τ_{eff} value of Schaye et al. (2004) at z =3. We make the assumption that any random variations in this bias from spectrum to spectrum will simply act as an additional source of noise which we can model in our simulation data.

In the bottom panel of Fig. 10, we have plotted a histogram of the $\tau_{\rm eff}$ values derived from the SDSS spectra. The mode of the distribution is $\tau_{\rm eff} = 0.35$, the same as the simulation results at z = 3 plotted in Fig. 7. In the observational data, there is a tail extending down to $\tau_{\rm eff} \sim 0.1$ and up to $\tau_{\rm eff} \sim 1$ which owes to noise and the continuum fitting uncertainties described above.

4.3 CMB data

We make use of the *WMAP* first year data (Bennett et al. 2003) in our analysis. For the illustrative application of Ly α -CMB crosscorrelation in this paper, we employ the foreground cleaned version of the data presented by Tegmark et al. (2004). The CMB temperature is given in HEALPIX format (12 × 512²) pixels, at the resolution of the *WMAP W*-band data (13-arcsec FWHM).

The Tegmark et al. map is chosen by us because it combines the five WMAP channels into a map with the highest angular resolution possible. However, because it is derived from channels with different frequencies and the thermal SZ effect has a weak frequency dependence, the map will have a scaled value of the SZ effect present, dependent on the weights used to combine the channels. In order to understand what impact this will have, we can use equation (3) to examine the thermal SZ temperature fluctuations in the different bands. Equation (3) can be written in the form $\Delta T/T = f(v) \times$ y, and at the lowest frequencies, $f(v) \rightarrow -2.0$. The five WMAP channels are centred at frequencies of 23, 33, 41, 61 and 94 GHz. The values of f(v) for the channels are then -1.97, -1.94, -1.91, -1.81 and -1.56, respectively. The weights for the Tegmark map vary spatially and are not publically available, but one can see the maximum difference of the effect compared to our Rayleigh-Jeans approximation [f(v) = -2.0] by considering the 94 GHz W-band map on its own, which gives approximately a 22 per cent reduction in the predicted tSZ signal. If we instead assume that the weights will have a similar effect to those used by the WMAP team itself for their linearly combined map, we find by summing over the five



Figure 10. SDSS τ_{eff} versus *WMAP T* for z = 3. We show jackknife errors (10 subsamples). We have subtracted off a ring for background from 15 to 25 arcmin around the pixel that the spectrum is in. The lines show predictions from Λ CDM after broadening the τ_{eff} values with observational errors. The solid line is for the Λ CDM baryon density, dotted for Λ CDM \times 5 (i.e. $\Omega_{b,I} = 0.4$) and dot–dashed Λ CDM \times 20 (i.e. $\Omega_{b,I} = 0.8$).

bands that $f_{\text{effective}} = \sum_{i}^{5} W_i f_i(\nu) = -1.79$, or a ~10 per cent reduction. The effect of frequency dependence is therefore small enough and uncertain enough that we neglect it in our illustrative application to observational data. In the future, however, as we have mentioned before, the frequency dependence can be used to better extract the signal. We will return to this point.

For each pixel centred on one of the SDSS quasar spectra, we extract the corresponding CMB temperature of that pixel. In order to crudely mitigate the effects of primary CMB anisotropies, which act as noise for our signal, we subtract a local background consisting of the CMB temperature averaged over all pixels in an annulus of inner radius 15 arcmin and outer radius 25 arcmin, centred on the pixel in question. Our motivation for this background subtraction is the fact that the SZ signals fall off rapidly on large scales (e.g. see Fig. 6), unlike the CMB primary signal. Because we will only be computing an upper limit to the SZ signal, this subtraction is a conservative assumption.

The presence of CMB primordial anisotropies will produce fluctuations on large scales. For example, there will be coherent fluctuations on the 100-arcmin scales which our simulated sky maps cover. The annulus subtraction technique represents a crude way to deal with this. As we will see below, the noise level on the pixel scale is sufficiently large with the *WMAP* data that we are not limited by the primary fluctuations in our analysis. However, in the future, with better data it will become imperative to model and remove the effect of the primordial anisotropies. This could be done with matched filtering (e.g. Schaefer et al. 2004), and the technique tested with simulations of the CMB sky which include large-scale anisotropy, foregrounds, noise and an instrumental beam response. We leave this for future work, and in Section 5.2 we will discuss other more sophisticated ways to deal with the CMB data.

4.4 Results

We find the rms temperature fluctuations in all 3488 pixels containing spectra to be 60 μ K. In the middle panel of Fig. 10, we show the rms temperature in each of 5 τ_{eff} bins. The error bars range from 1.3 to 6.1 μ K. This is close to the Poisson error estimate in each case, which is not surprising because the pixels are sampled quite sparsely from the data. Given that we expect ~1 μ K rms fluctuations at most from the kinetic SZ effect, it is clear that this particular data set will fall short of giving us any meaningful constraint from the kSZ. Because the kSZ signal will add in quadrature to the rms from the noise and CMB primary signal, the highest τ_{eff} bin might have a total rms of $\sqrt{60^2 + 1^2} \mu$ K and the lowest $\sqrt{60^2 - 1^2} \mu$ K, a difference of 0.02 μ K. In practice, because the CMB data are smoothed, we expect at least a factor of 10 less signal (see the middle panel of Fig. 7). The kSZ signal is therefore at least 500 times too small to be detectable in this way with these data.

If instead we turn to the thermal SZ signal, the situation is much more promising. As a mean temperature rather than a variance, it is easier to constrain this signal from noisy data. In the top panel of Fig. 10, we show the mean temperature in pixels for the five bins of Ly $\alpha \tau_{eff}$. Owing to the background subtraction, the mean value of the pixels is zero. When comparing to the simulation data, we will renormalize so that the same is true. The observational error bars vary from 1.3 to 8.0 μ K per bin.

From Fig. 7, we expect the mean temperature decrement from the thermal SZ effect at redshift z = 3 to exhibit a total variation of $\sim 0.6 \,\mu\text{K}$ amongst the pixels with highest and lowest Ly α forest τ_{eff} . This is for CMB maps at the WMAP resolution, but unfortunately not including noise in the Ly α τ_{eff} estimates. As detailed in Section 4.2 above, we model the effects of noise in the spectra and continuum fitting uncertainties by adding a Gaussian-distributed error to the $\tau_{\rm eff}$ from simulations. We add a $\tau_{\rm eff}$ value randomly drawn from a normal distribution with standard deviation $\sigma = 0.1$. Doing this to the simulated values of τ_{eff} enables us to reproduce the width of the histogram of observed τ_{eff} values. In the top panel of Fig. 10, we plot the curve of simulation CMB temperature $\langle T \rangle$ versus simulated $au_{
m eff}$ after this noise has been added. The line which lies closest to the $\langle T \rangle = 0$ axis is this prediction. As we can see, the predicted variation of $\langle T \rangle$ with τ_{eff} is barely noticeable on this scale, indicating that tight constraints lie only in the future with larger observational data sets.

Nevertheless, the comparison is interesting because it constitutes an attempt to detect the ionized baryons at z = 3, and the sensitivity increase required for success, if our current cosmological models are correct, is only about one order of magnitude. In order to place an upper limit on the temperature and density of ionized baryons using existing data, one would ideally use a grid of simulations with different values of the baryon density, with the amplitude of matter fluctuations normalized so that their clustering and other properties are consistent with observations (e.g. Spergel et al. 2003; Tegmark et al. 2004). In this manner, both the clustering of baryons and the variation of mean IGM temperature with Ω_b could be marginalized over to yield a constraint on Ω_b from ionized baryons, $\Omega_{b,I}$. The effects of varying Ω_b on structure at high redshift were explored with a suite of simulations by Gardner et al. (2003). These authors found, for example, that Ly α forest absorption is virtually unaffected by changing Ω_b (by a factor of ~6), as long as the ionizing background intensity is adjusted to reproduce the same value of τ_{eff} .

In the present case, because of the large error bars, the value of $\Omega_{b,I}$ consistent with observations is likely to be so high that a simpler and more crude approach than a full simulation grid is all that is warranted. In order to model the effect of increasing $\Omega_{\rm h}$, we simply multiply the simulation value of SZ $\langle T \rangle$ by a constant factor, assumed to be equivalent to multiplying $\Omega_{\rm b}$ by the same factor. We do not adjust the Ly α observables, imagining that any increase in baryon density can be cancelled out by the presence of a UV background radiation field higher in intensity by the same factor. Moreover, we assume that increasing $\Omega_{\rm b}$ does not affect the amplitude of matter fluctuations and that the temperature of the IGM stays the same as in the $\Omega_{\rm b} = 0.04$ simulation. A model with a very large value of $\Omega_{\rm b}$ is likely to have many other differences related to gas physics and cooling, such as the number of collapsed baryonic objects. Additionally, of course, it will be incompatible with big bang nucleosynthesis (BBN) constraints (e.g. Burles & Tytler 1998), and measurements of $\Omega_{\rm b}$ from the CMB acoustic peaks (Spergel et al. 2003). None of these problems concerns us here, as our comparison to simulations is essentially illustrative.

Under these assumptions, a χ^2 fit of the simulation data to the *WMAP*/SDSS data, both binned in the same manner, yields $\Omega_{b,I} < 0.55$ at 90 per cent confidence and $\Omega_{b,I} < 0.80$ at 95 per cent confidence. The limits on the fraction of the critical density in ionized baryons at z = 3 are therefore very weak. This is as one would expect from looking at Fig. 10 where we show the simulation curves for various values of $\Omega_{b,I}$. It is possible to conclude that the Universe is not closed by ionized baryons, but not much else. This situation is likely to be much improved with future observational data, as we will discuss below.

Much of the statistical power in the measurement comes from the bin with the lowest value of τ_{eff} . Most of the pixels have moderate values of τ_{eff} , but these pixels in the lowest bin also have low SZ decrement amplitudes. In this manner, the CMB–Ly α forest cross-correlation can be said to be probing the low-density IGM.

5 SUMMARY AND DISCUSSION

5.1 Summary

Using a cosmological hydrodynamic simulation, we have investigated the CMB temperature fluctuations caused by the diffuse IGM at redshifts z = 2 - 6, also responsible for the Ly α forest. Our main findings can be summarized as follows.

(i) At redshift z = 2, the kinetic Sunyaev–Zel'dovich effect from gas within the Ly α to Ly β region of the forest is predicted to cause CMB fluctuations of amplitude 0.8- μ K rms averaged in a top-hat filter of angular radius 1 arcmin. This remains roughly constant for higher redshifts.

(ii) The thermal SZ at redshift z = 2 yields a rms T of 2.4 μ K. At higher redshift, the relative fluctuation owing to the thermal SZ compared to kinetic SZ declines as expected because of the decreasing mean temperature. However, there is still a substantial $T_{\rm rms}$ at higher redshift, 0.4 μ K at z = 4.

(iii) There is a correlation between Ly α forest properties of quasar spectra with CMB temperature fluctuations caused by both the kinetic and thermal SZ effects. In particular, there is a strong relation between the mean effective optical depth τ_{eff} averaged over the entire Ly α forest region and the CMB temperature. Pixels (of size \sim 1 arcmin) in the top 10 percentile of τ_{eff} are predicted to have

a mean CMB temperature of 1.5 μ K lower than the bottom 10 percentile at z = 2, owing to the thermal SZ effect. The quantity τ_{eff} is relatively sensitive to density fluctuations in the diffuse IGM.

(iv) There is, however, little correlation between the total H_I column density in spectra and the CMB temperature, with no relation at all seen for column densities $N_{\rm HI} > 10^{16} \rm \, cm^{-2}$. This statistic is most sensitive to the neutral hydrogen density in saturated regions, indicating that SZ effects correlate best with relatively diffuse unsaturated gas.

(v) We have compared the correlation of Ly $\alpha \tau_{eff}$ with CMB T seen in simulations of the Λ CDM cosmology with observational data from the SDSS and *WMAP*. Because the kinetic SZ signal and pixel noise add in quadrature, no constraints are possible on the IGM velocity field with the present data. However, the sensitivity is within an order of magnitude of that necessary to detect the predicted thermal SZ effect from the Ly α forest. We are able to place a weak limit on the density in ionized baryons at z = 3 of $\Omega_{B,i} < 0.80$ at 95 per cent confidence. Below, we outline future developments which would make a direct detection possible.

5.2 Discussion

In the present paper, we have examined the relationship between $Ly\alpha$ forest absorption and the SZ effect in general, without the expectation that a detection of a correlation can be made from current data. In the future, the quality and size of data sets will increase, and it will become worthwhile to plan a strategy for the cross-correlation that makes the best use of the data. Two obvious improvements could be made in the future. In our illustrative comparison of *WMAP* and SDSS data, we did not use the full angular or frequency information, both of which would help with the sensitivity of the test.

As we have seen that the thermal SZ effect offers the best hope for the detection of the ionized baryons, making use of the spectral difference between the SZ decrement and the primary CMB signal will enable easier detection of the effect. One could imagine fitting the τ_{eff} -T correlation in different CMB frequency bands to do this. For example, in the work of Huffenberger, Seljak & Makarov (2004), a method was derived for combining the WMAP channels in an optimal manner to remove the primary anisotropy, effects of noise and contaminating point sources, enabling the thermal SZ signature to be constrained directly. In this paper, we have concerned ourselves with predictions for the Rayleigh-Jeans part of the spectrum, which is reasonable given the large observational uncertainties in the present data. We note that at higher frequencies the thermal SZ signal becomes positive (e.g. at 330 GHz the magnitude is the same as the Rayleigh-Jeans signal but with the opposite sign). Bolometers tend to be more sensitive than radiometers at present, and they operate at frequencies >100 GHz. These differences could be used profitably to increase the signal-to-noise ratio of any detection. Also, with the WMAP data, we have limited our analysis to single pixels. Making use of the full spatial cross-correlation in our detection of the signal will be useful when CMB data sets with higher angular resolution become available. Another possibility could be to use wavelets to isolate the relevant physical scales in a future analysis (see e.g. Vielva, Martínez-González & Tucci 2006).

One can ask whether emission from the QSOs themselves might create a signal, as the SED of QSOs in the microwave band is not well constrained (Perna & Di Matteo 2000; White & Majumdar 2004). Apart from the possible intrinsic microwave emission of QSOs, there will also be an intrinsic thermal SZ effect from galaxy clusters which host QSOs. Neither of these will necessarily bias any measurement of the $Ly\alpha$ -CMB correlation unless QSOs with more Ly α absorption in their spectra have more intrinsic microwave emission. This seems unlikely in the former case, although in the case of a SZ signal associated with clusters in which the QSOs lie one could argue that this is part of the signal that we are trying to measure. Otherwise, it could be removed if we so desire by not using the part of the spectrum closest to the QSO.

The method we have outlined in this paper is one way to potentially detect ionized baryons in the IGM at high redshift. In general, one would like to know if there are any other methods that can be used. At lower redshifts, for example (z < 1), the ionized baryons are mainly collisionally ionized, with such a low neutral fraction that they cannot be seen in the Ly α forest. Their higher temperature, however, means that this WHIM (see e.g. Cen & Ostriker 1999; Davé et al. 2001) can potentially be detected by looking for X-ray absorption (e.g. Fang et al. 2002) or emission (Yoshikawa et al. 2004; Fang et al. 2005) or perhaps in Ly α emission (e.g. Furlanetto et al. 2003, 2005). The thermal SZ signal at low redshifts is also one way to detect the WHIM. At higher redshifts and lower IGM temperatures considered in this paper, however, the X-ray signal (see e.g. Croft et al. 2001) is not likely to be observable.

The overall *y*-distortion in the CMB spectrum from the hot IGM should eventually be detectable. Limits from the *COBE* Far Infrared Absolute Spectrometer (FIRAS) instrument (e.g. Mather et al. 1990) are $y < 2.5 \times 10^{-5}$ at 95 per cent confidence, more than an order of magnitude above the mean *y* caused by the high-*z* IGM. However, separating out the contribution from the high-*z* IGM will be difficult without a spatial correlation of the type we have proposed here. One could also use other tracers of the high-*z* density field for a cross-correlation, such as Lyman-break galaxies, or QSOs themselves, although their connection to the IGM would be less easy to interpret.

Another interesting question to ask is how likely it is that we will be able to constrain the high-*z* IGM velocity field, given that the kinetic SZ signal from the photoionized IGM is expected to be relatively important (Loeb 1996). Because of the possible positive or negative nature of the effect, depending on the velocity, we cannot use a mean signal, unlike the thermal SZ. By using a rms fluctuation as our statistic, we therefore need a noise level which is extremely low. Since the spectrum of the kSZ is the same as the CMB, the primary fluctuations become even more important sources of noise than with the thermal effect. Given that the rms fluctuations on scales of 1 arcmin from the kSZ are predicted to be $\sim 5 \,\mu\text{K}$ at z = 2, one would need the 'noise' from primary anisotropies and other sources to be less than this in order to detect a signal.

The thermal SZ is more promising. Future experiments, which will be useful for improving constraints on $\Omega_{b,1}$, include the Planck satellite mission,² the Sunyaev–Zel'dovich Array (SZA³), the Atacama Cosmology Telescope (ACT, see Kosowsky 2003) and the South Pole Telescope (SPT⁴). Given the expected noise characteristics and angular resolution/sky coverage of these instruments, we expect substantial improvements over the constraints derived using *WMAP* in this paper. We have seen in the *WMAP* case (Section 4) that we are limited both by the noise level (~60 µK rms) and by the resolution of the data. As an example, the ACT angular resolution is projected to be ~1 arcmin with a rms detector noise level of ~2 µK per pixel in the 145-GHz band. Even without allowing for the higher resolution or using the full angular cross-correlation information, one would therefore expect to extract the same level

Ly α forest thermal SZ signal using (60/2)² fewer spectra (assuming the primary anisotropy can be filtered out in frequency space). The ACT signal will be lower by about 50 per cent at 145 GHz than the Rayleigh–Jeans prediction, but the other frequency bands can be used to increase the significance of the detection. In order to yield a 2σ detection of Ω_b , using pixel–pdf statistics, as in this paper, we estimate that ~300 quasar Ly α forest spectra would be needed in the ACT survey area, assuming the BBN value of $\Omega_b h^2$ (0.019; Burles & Tytler 1998) Using angular cross-correlation information would increase the significance of the detection. The lack of overlap between the ACT fields and the SDSS would necessitate a southern quasar survey, although it should be borne in mind that only extremely low resolution spectra are needed.

We note that in our analysis we have assumed that the ratio of neutral to ionized hydrogen is that expected if the ionizing background radiation field does not vary spatially in intensity. If there are coherent spatial fluctuations in the photoionization rate (Γ in equation 1), then the correlation between Ly α optical depth and CMB temperature decrement will be affected by an additional source of scatter. This is likely to be small, however with the spatial cross-correlation reduced by 5 per cent or less (see e.g. Croft 2004) and could in principle be calibrated out using theoretical modelling in the manner of McDonald et al. (2005).

Finally, one can ask why it is important to try to detect the ionized component of the baryons at high redshift when the neutral hydrogen tracer has already been well measured (e.g. Rauch et al. 1997). One reason is that a detection would act as a consistency check and enable us to directly verify the baryon inventory at these redshifts. Also, at the moment, the best constraints on the baryon density in observed structures (if we set aside for the moment relatively indirect CMB and BBN measurements, e.g. Spergel et al. 2003; Walker et al. 1991) comes from looking at the mean absorption level in the Ly α forest (e.g. Rauch et al. 1997; Weinberg et al. 1997; Hui et al. 2002; Tytler et al. 2004). However, these measurements are degenerate with constraints on the ionizing background radiation intensity J. Often, $\Omega_{\rm b}$ from BBN or CMB is assumed, and then the Ly α forest measurement is used to constrain the ionizing radiation intensity (e.g. Kirkman et al. 2005). If we use actual measurements of J (from e.g. the proximity effect; e.g. Scott et al. 2000) then constraints on Ω_b directly seen in the Ly α forest become weaker (J is only known to within roughly a factor of 2). Measurements of the ionized baryon fraction such as those we propose may therefore become competitive.

ACKNOWLEDGMENTS

We thank Scott Burles for providing the modified SDSS Data Release 3 data set used in this paper and also Volker Springel for allowing us to use the cosmological simulation data. RACC thanks Tiziana Di Matteo, Volker Springel, Jeff Peterson and Xuelei Chen for useful discussions and the hospitality of Simon White and the Max Planck Institute for Astrophysics, where this work was largely carried out. RACC acknowledges partial support from NASA ATP contract ATP05-39.

REFERENCES

Abazajian K. et al., 2005, Astron J., 129, 1755

- Bennett C. L. et al., 2003, ApJS, 148, 1
- Bradley E., 1982, The Jackknife, the Bootstrap and other Resampling Plans. SIAM, Philadelphia

² http://planck.esa.int

³ http://astro.uchicago.edu/sza

⁴ http://astro.uchicago.edu/spt

Burles S., Tytler D., 1998, ApJ, 499, 699

- Carlstrom J. E., Holder G., Reese E. D., 2002, ARA&A, 40, 643
- Cen R., Ostriker J. P., 1999, ApJ, 514, 1
- Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, ApJ, 437, L9
- Croft R. A. C., 2004, ApJ, 610, 642
- Croft R. A. C., Weinberg D. H., Katz N., Hernquist L. 1998, ApJ, 495, 44
- Croft R. A. C., Di Matteo T., Davé R., Hernquist L., Katz N., Fardal M. A., Weinberg D. H., 2001, ApJ, 557, 67
- da Silva A. C., Barbosa D., Liddle A. R., Thomas P. A., 2000, MNRAS, 317, 37
- Davé R., Hernquist L., Katz N., Weinberg D. H., 1999, ApJ, 511, 521
- Davé R. et al., 2001, ApJ, 552, 473
- Fang T., Marshall H. L., Lee J. C., Davis D. S., Canizares C. R., 2002, ApJ, 572, 127
- Fang T., Croft R. A. C., Sanders W. T., Houck J., Davé R., Katz N., Weinberg D. H., Hernquist L., 2005, ApJ, 623, 612
- Furlanetto S., Schaye J., Springel V., Hernquist L., 2003, ApJ, 599, L1
- Furlanetto S., Schaye J., Springel V., Hernquist L., 2005, ApJ, 622, 7
- Gardner J. P., Katz N., Hernquist L., Weinberg D., 2003, ApJ, 587, 1
- Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44
- Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
- Haardt F., Madau P., 1996, ApJ, 461, 20
- Hernquist L., Katz N., Weinberg D., Miralda-Escudé J., 1996, ApJ, 457, L51
- Huffenberger K. M., Seljak U., Makarov A., 2004, Phys. Rev. D, 70, 063002 Hui L., Burles S., Seljak U., Rutledge R. E., Magnier E., Tytler D. 2001,
- ApJ, 552, 15
- Hui L., Haiman Z., Zaldarriaga M., Alexander T., 2002, ApJ, 564, 525
- Katz N., Weinberg D. H., Hernquist L., Miralda-Escudé J., 1996a, ApJ, 457, L57
- Katz N., Weinberg D. H., Hernquist L., 1996b, ApJS, 105, 19
- Kirkman D. et al., 2005, MNRAS, 360, 1373
- Kosowsky A., 2003, New Astron. Rev., 47, 939
- Kuo C. L. et al., 2004, ApJ, 600, 32
- Loeb A., 1996, ApJ, 471, L1
- Mather J. C. et al., 1990, ApJ, 354, L37
- McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., 2001, ApJ, 562, 52
- McDonald P. et al., 2004, ApJS, 163, 80
- McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471
- McQuinn M., Furlanetto S. R., Hernquist L., Zahn O., Zaldarriaga M., 2005, ApJ, 630, 643
- Nagamine K., Cen R., Hernquist L., Ostriker, J. P., Springel V., 2005, ApJ, 627, 608
- Ostriker J. P., Vishniac E. T., 1986, ApJ, L51

- Perna R., Di Matteo T., 2000, ApJ, 542, 68
 - Petitjean P., Muecket J. P., Kates R., 1995, A&A, 295, 9
 - Rauch M., 1998, ARA&A, 36, 267
- Rauch M. et al., 1997, ApJ, 489, 7
- Schaefer B. M., Pfrommer C., Hell R., Bartelmann M., 2004, MNRAS, submitted (astro-ph/0407090)
- Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, MNRAS, 318, 817
- Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2004, ApJ, 596, 768
- Scott J., Bechtold J., Dobrzycki A., Kulkarni V., 2000, ApJS, 130, 67
- Spergel D. N. et al., 2003, ApJS, 148, 175
- Springel V., 2005, MNRAS, 364, 1105
- Springel V., Hernquist L., 2002, MNRAS, 333, 649
- Springel V., Hernquist L., 2003a, MNRAS, 339, 289
- Springel V., Hernquist L., 2003b, MNRAS, 339, 312
- Springel V., White M., Hernquist L., 2001a, ApJ, 549, 681
- Springel V., Yoshida N., White S. D. M., 2001b, New Astron., 6, 79
- Springel V., White M., Hernquist L., 2002, ApJ, 579, 16
- Sunyaev R. A., Zel'dovich Ya. B., 1972, ARA&A, 4, 173
- Sunyaev R. A., Zel'dovich Ya. B., 1980, ARA&A, 18, 537
- Suzuki N., Tytler D., Kirkman D., O'Meara J. M., Lubin D., 2005, ApJ, 618, 592
- Tegmark M. et al., 2004, Phys. Rev. D, 69, 3501
- Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
- Tytler D. et al., 2004, ApJ, 617, 1
- Vielva P., Martínez-González E., Tucci M., 2006, MNRAS, 365, 891
- Wadsley J. W., Bond J. R., 1997, in Clarke D., West M., eds, ASP Conf. Ser., Vol. 123, The 12th Kingston Meeting: Computational Astrophysics. Astron. Soc. Pac., San Francisco p. 332
- Walker T. P., Steigman G., Kang H., Schramm D. M., Olive K. A., 1991, ApJ, 376, 51
- Weinberg D. H., Miralda-Escudé J., Hernquist L., Katz N., 1997, ApJ, 490, 564
- White M., Majumdar S., 2004, ApJ, 602, 565
- White M., Springel V., Hernquist L., 2002, ApJ, 579, 16
- Yoshida N., Furlanetto S., Hernquist L., 2005, ApJ, 618, L91
- Yoshikawa K. et al., 2004, PASJ, 56, 939
- Zhang Y., Anninos P., Norman M. L., 1995, ApJ, 453, L57

This paper has been typeset from a TEX/LATEX file prepared by the author.