Mon. Not. R. Astron. Soc. **388,** 1889–1898 (2008) doi:10.1111/j.1365-2966.2008.13546.x

A method to measure the mass of damped Lyα absorber host galaxies using fluctuations in 21-cm emission

J. Stuart B. Wyithe★
School of Physics, University of Melbourne, Parkville, Victoria 3010, Australia

Accepted 2008 June 6. Received 2008 June 6; in original form 2008 April 10

ABSTRACT

Observations of damped Lyα absorbers (DLAs) indicate that the fraction of hydrogen in its neutral form (H I) is significant by mass at all redshifts. This gas represents the reservoir of material that is available for star formation at late times. As a result, observational identification of the systems in which this neutral hydrogen resides is an important missing ingredient in models of galaxy formation. Precise identification of DLA host mass via traditional clustering studies is not practical owing to the small numbers of known systems being spread across sparsely distributed sightlines. However, following the completion of re-ionization, 21-cm surface brightness fluctuations will be dominated by neutral hydrogen in DLAs. No individual DLAs could be detected in 21-cm emission. Rather, observations of these fluctuations will measure the combined clustering signal from all DLAs within a large volume. We show that measurement of the spherically averaged power spectrum of 21-cm intensity fluctuations due to DLAs could be used to measure the galaxy bias for DLA host galaxies when combined with an independent measurement of the cosmological H I mass density from quasar absorption studies. Utilizing this technique, the low-frequency arrays now under construction could measure the characteristic DLA host mass with a statistical precision as low as 0.3 dex at $z \gtrsim 4$. In addition, high signal-to-noise ratio observations of the peculiar-motion-induced anisotropy of the power spectrum would facilitate measurement of both the DLA host mass and the cosmic H I density directly from 21-cm fluctuations. By exploiting this anisotropy, a second generation of low-frequency arrays with an order of magnitude increase in collecting area could measure the values of cosmic H I density and DLA host mass, with uncertainties of a few per cent and a few tens of per cent, respectively.

Key words: galaxies: high-redshift – intergalactic medium – cosmology: theory – diffuse radiation – large-scale structure of Universe.

1 INTRODUCTION

The primary avenue for study of the physical properties of the intergalactic medium (IGM) utilizes the Lyα absorption systems observed along the lines-of-sight towards high-redshift quasars. These absorbing systems include the Lyα forest (with column depth $n < 10^{17}\,\mathrm{cm}^{-2}$), Lyman limit systems ($10^{17} < n < 2 \times 10^{20}\,\mathrm{cm}^{-2}$) and damped Ly$\alpha$ absorbers (DLAs; $n < 2 \times 10^{20}\,\mathrm{cm}^{-2}$). The latter are self-shielded and have been shown to host >80 per cent of the neutral hydrogen (H I) gas during most of cosmic time $z \lesssim 5$ (Prochaska, Herbert-Fort & Wolfe 2005).

As the primary reservoirs of neutral gas, it is thought that DLAs provide the dominant sites for star formation, and represent the progenitors of modern galaxies (Wolfe, Gawiser & Prochaska 2005). Thus the identity of DLA hosts is crucial for our understanding of galaxy formation, and of the build-up of stellar mass in the Universe. DLA hosts have been identified in a few cases (e.g. Warren et al. 2001; Colbert & Malkan 2002), and observations of these (with some exceptions) are consistent with their being at the faint end of the Lyman break galaxy (LBG) population. On the other hand, these studies will more easily discover hosts in the most luminous systems, and could therefore be systematically biased in their conclusions. Within the cold dark matter model the masses of objects may be determined independently from their luminosity by studying their clustering properties and appealing to theoretical predictions of the bias relative to the underlying mass distribution. At high redshift $z < 3$, clustering analyses have been used in this way to study the host halo masses of quasars (Shen et al. 2007; White, Martini & Cohn 2007) and LBGs (Adelberger et al. 2005). However, the number of DLAs known is small, occurring only once in every few quasar spectra, and so no sample of sufficient size

★E-mail: swyithe@unimelb.edu.au

to allow for an autocorrelation clustering analysis exists (Cooke et al. 2006). As a result a quantitative measurement of DLA host mass has remained illusive. Some progress has been made through comparison of clustering with the LBG population through cross-correlation. Cooke et al. (2006) have performed a survey for LBGs in regions surrounding known DLA systems. This sample allows for measurement of the cross-correlation between DLAs and LBGs. Cooke et al. (2006) find the cross-correlation length to be comparable to the correlation length (Adelberger et al. 2005) among the large LBG sample of Steidel et al. (2003), and so conclude that DLA systems have similar masses to LBGs. They argue for a host halo mass of $10^9 \lesssim M \lesssim 10^{12}\,\mathrm{M}_\odot$. This interpretation is consistent with modelling studies carried out in a cosmological framework (Nagamine et al. 2007).

Alternatively, the presence of cold gas is also traced by Mg II absorption. Indeed, Mg II absorbers have been shown to be associated with neutral hydrogen absorbers over a range of column densities, including DLAs (e.g. Rao, Turnshek & Nestor 2006), and so Mg II absorption is used as a proxy to study the H I galaxy population at $z \lesssim 1.5$. The host halo masses of Mg II absorbers have been estimated at $z \sim 0.5$ via cross-correlation with luminous red galaxies in the Sloan Digital Sky Survey Data Release 3 (SDSS DR3; Bouché et al. 2006), yielding a host mass of $M \sim 10^{12}\,\mathrm{M}_\odot$. The large number of objects available (both Mg II absorbers and galaxies) yields a statistical accuracy of a factor of 2 in host halo mass, much better than is available using comparable techniques for DLAs at higher redshift (Cooke et al. 2006). In addition, host masses have been estimated via direct kinematic measurement in spatially resolved follow-up spectroscopy (Bouché et al. 2007), yielding estimates that are consistent with clustering results.

The small number of known DLAs is due to the low density of bright background sources. An alternative approach is therefore to observe the DLA population directly through 21-cm emission. There is an extensive literature describing the use of redshifted 21-cm observations as a probe of the process of re-ionization in the high-redshift IGM (see Furlanetto, Oh & Briggs 2006, for an extensive discussion). Re-ionization starts with ionized (H II) regions around galaxies, which later grow to surround groups of galaxies. The process of re-ionization is completed when these H II regions overlap (defining the so-called *overlap* epoch) and fill-up most of the volume between galaxies. The conventional wisdom has been that the 21-cm signal disappears after the *overlap* epoch, because there is little neutral hydrogen left through most of intergalactic space. However, the following simple estimate can be used to demonstrate that 21-cm emission should be significant even after re-ionization is completed (Wyithe & Loeb 2008). As mentioned above, DLA systems are believed to contain the majority of H I at high redshifts. Indeed, observations of DLAs out to a redshift of $z \sim 4$ show the cosmological density parameter of H I to be $\Omega_{\rm H\,I} \sim 10^{-3}$ (Prochaska et al. 2005). In the standard cosmological model the density parameter of baryons is $\Omega_{\rm b} \sim 0.04$, so that the mass-averaged neutral hydrogen fraction at $z \sim 4$ (long after the end of the H II overlap epoch) is $x_{\rm H\,I} \equiv \Omega_{\rm H\,I}/\Omega_{\rm b} \sim 0.02$.

At $z \sim 4$, the brightness temperature contrast of redshifted 21-cm emission will be $\Delta T \sim 0.5\,\mathrm{mK}$. Moreover, on the $R \sim 10$ comoving Mpc scales relevant for upcoming 21-cm experiments, the *root-mean-square* amplitude of density fluctuations at $z \sim 4$ is $\sigma \sim 0.2$. Hence, we expect fluctuations in the 21-cm intensity field due to DLAs to be at least $\sim$0.1 mK on 10 comoving Mpc scales, with a boost in this signal if the DLAs are hosted by biased (i.e. massive) dark matter haloes (Wyithe & Loeb 2008). The fluctuations in the 21-cm emission signal after re-ionization are therefore expected to be only an order of magnitude or so smaller than the largest fluctuations predicted at any time during the entire re-ionization era (e.g. Wyithe & Morales 2007). Moreover, the sky temperature, which provides the limiting factor in the system noise at the low frequencies relevant for 21-cm studies, is proportional to $(1+z)^{2.6}$, and so is a factor of $\sim 3.4\,[(1+z)/5]^{2.6}$ smaller at low redshifts than for observations at $z \sim 7$.

Several experiments are currently under development which aim to detect the 21-cm signal during re-ionization [including Murchison Widefield Array (MWA),[1] Low Frequency Array (LOFAR),[2] PAPER,[3] 21-cm Array (21CMA)[4]] and more ambitious designs are being planned [Square Kilometer Array (SKA)[5]]. In addition, the first statistical detection of 21-cm fluctuations due to discrete, unresolved clumps of neutral gas was recently made (Pen et al. 2008) through cross-correlation of the H I Parkes All Sky Survey (HIPASS; Barnes et al. 2001) 21-cm observations of the local universe with galaxies in the 6 degree field galaxy redshift survey (Jones et al. 2004, 2005). This detection represents an important step towards using 21-cm surface brightness fluctuations to probe the neutral gas distribution in the IGM (both during and after re-ionization), as well as the mass power spectrum (PS) and cosmology (McQuinn et al. 2006; Chang et al. 2008; Loeb & Wyithe 2008; Mao et al. 2008; Pritchard & Loeb 2008; Wyithe, Loeb & Geil 2008).

The first generation of low-frequency radio arrays will have low spatial resolution, and insufficient sensitivity to detect individual DLAs in emission. On the other hand, the redshifted 21-cm emission is sensitive to the total (mass-weighted) optical depth of this neutral gas. Observations of the redshifted 21-cm signal would therefore detect the total neutral hydrogen content in a volume of IGM dictated by the observatory beam and frequency bandpass, and as a result each beam will contain the combined emission from a large number of DLAs. It is important to note that observation of DLAs via 21-cm surface brightness fluctuations would be fundamentally different from quasar absorption line studies. The surface brightness fluctuations would be sensitive to the statistical properties of the population (particularly the clustering), but could not be used to study individual DLAs.

Although the 21-cm emission after H II overlap is dominated by dense clumps of gas rather than by diffuse gas in the IGM as is the case before re-ionization is complete, we do not expect 21-cm self-absorption to impact the level of 21-cm emission. This conclusion is based on 21-cm absorption studies towards DLAs at a range of redshifts between $z \sim 0$ and $\sim$3.4, which show optical depths to absorption of the background quasar flux with values less than a few per cent (Kanekar & Chengalur 2003; Curran et al. 2007). The small optical depth for self-absorption is also supported by theoretical calculations of the 21-cm optical depth of neutral gas in high-redshift minihaloes (Furlanetto & Loeb 2002). Moreover, DLAs have a spin temperature that is large relative to the temperature of the cosmic microwave background (CMB) radiation, and will therefore have a level of emission that is independent of the kinetic gas temperature (e.g. Kanekar & Chengalur 2003). These factors combine to make prediction of the 21-cm signal from DLAs robust against poorly understood astrophysical details of galaxy formation.

[1] http://www.haystack.mit.edu/ast/arrays/mwa/
[2] http://www.lofar.org/
[3] http://astro.berkeley.edu/ dbacker/EoR/
[4] http://web.phys.cmu.edu/ past/
[5] http://www.skatelescope.org/

In this paper we consider the PS of 21-cm intensity fluctuations due to DLA systems at $2.5 \lesssim z \lesssim 5.5$ where there are complementary quasar absorption line studies (Section 2). We show that the clustering could be accurately determined using low-frequency radio telescopes currently under construction (Section 3), and that this clustering will allow for an accurate measurement of the DLA mass (Sections 4–5). We summarize our conclusions in Section 6. In our numerical examples, we adopt the standard set of cosmological parameters (Spergel et al. 2007), with values of $\Omega_{\rm m} = 0.24$, $\Omega_{\rm b} = 0.04$ and $\Omega_\lambda = 0.76$ for the matter, baryon and dark energy fractional density, respectively, and $h = 0.73$ for the dimensionless Hubbles constant. We compute the PS of matter fluctuations using CMBFAST (Seljak & Zaldarriaga 1996) with a variance linearly extrapolated to the present day of $\sigma_8 = 0.76$ inside spheres of radius $8\,h^{-1}$ Mpc.

2 THE 21-cm PS OF DLAS

We begin by describing the 21-cm PS following the completion of re-ionization. The amplitude of fluctuations in the 21-cm signal depends on the mass-averaged neutral hydrogen fraction in the IGM, and on the clustering bias of DLA systems. The latter is related to the galaxy bias of the DLA host dark matter haloes. If more massive haloes tend to house more neutral hydrogen, we would expect there to be a mean relation between DLA column density and the host halo mass. However, DLA systems have a range of column densities, and so we would also expect significant scatter around this mean relation, both because of variation in the H I content from galaxy to galaxy, and because the column density for a particular galaxy will depend on the line-of-sight probed by a particular observation. In addition, there is a small contribution to the neutral hydrogen content of the IGM from Lyman limit systems and the Lyα forest.

For a DLA host with halo mass M, the galaxy bias b may be approximated using the Press–Schechter (Press & Schechter 1974) formalism (Mo & White 1996), modified to include non-spherical collapse (Sheth, Mo & Tormen 2001):

$$b(M,z) = 1 + \frac{1}{\delta_{\rm c}}\left[\nu'^2 + b\nu'^{2(1-c)} - \frac{\nu'^{2c}/\sqrt{a}}{\nu'^{2c} + b(1-c)(1-c/2)}\right], \quad (1)$$

where $\nu \equiv \delta_{\rm c}/\sigma(M)$, $\nu' \equiv \sqrt{a}\nu$, $a = 0.707$, $b = 0.5$ and $c = 0.6$. Here $\sigma(M)$ is the variance of the density field smoothed on a mass-scale M at redshift z. This expression yields an accurate approximation to the halo bias determined from N-body simulations (Sheth et al. 2001).

Our approach to computing the DLA PS is to assume that there is a probability distribution $p(b)\,{\rm d}b$ for the neutral hydrogen mass-weighted clustering bias (b) of DLAs relative to the underlying dark matter. This distribution has a mean $\langle b\rangle$, and a characteristic mass $M_{\langle b\rangle}$, where $\langle b\rangle = b(M_{\langle b\rangle})$. The quantity $\langle b\rangle$ is the typical galaxy bias of haloes hosting neutral hydrogen. Since DLAs dominate the neutral hydrogen mass, and are thought to reside in massive, biased haloes, we refer to this typical bias in the 21-cm PS as the bias of DLA hosts, and $M_{\langle b\rangle}$ as the DLA host mass throughout this paper.

Our goal is to discuss the statistical precision with which $M_{\langle b\rangle}$ can be determined via the 21-cm PS. As noted above, while most neutral hydrogen is located in the DLAs, a fraction of H I in the IGM is observed in Lyman limit systems with column densities below $2 \times 10^{20}\,{\rm cm}^{-2}$. If these systems are found in the same host halo population as the DLAs, then $M_{\langle b\rangle}$ will faithfully represents the DLA host mass. On the other hand, if Lyman limit systems are located in smaller less biased haloes, then equation (2) will lead to an underestimate of the DLA host mass. We do not expect this underestimate to be significant because the contribution to the observed PS is weighted by the H I mass density squared and the bias squared, with DLA systems dominating both quantities.

Assuming that the relation between H I and halo mass is independent of large-scale overdensity (and hence ionizing background), the 21-cm PS due to DLAs is

$$P_{21}(k) = 400\,{\rm mK}\left(\frac{1+z}{7.5}\right) x_{\rm H\,I}^2 P(k)(\langle b\rangle + f\mu^2)^2, \quad (2)$$

where $x_{\rm H\,I} \equiv \Omega_{\rm H\,I}/\Omega_{\rm b}$ is the mass-weighted neutral fraction of hydrogen in the Universe. The PS is evaluated at a wavenumber $\boldsymbol{k}$ with modulus k and $\mu = \cos\theta$, where θ is the angle between the line-of-sight and wavenumber $\boldsymbol{k}$. The peculiar motion induced anisotropy is dependent on cosmology through the evolution of growth factor D via $f = {\rm d}\log D/{\rm d}\log(1+z)$, and follows from application of redshift space distortions (Kaiser 1987) to the 21-cm PS. We assume $f = 1$, which is appropriate at high redshift. From equation (2) we also find the spherically averaged PS:

$$P_{21}(k) = 400\,{\rm mK}\left(\frac{1+z}{7.5}\right) x_{\rm H\,I}^2 \left(\langle b\rangle^2 + \frac{2}{3}\langle b\rangle + \frac{1}{5}\right) P(k). \quad (3)$$

We note that in the case where $\langle b\rangle = 1$ (which corresponds to a uniformly ionized IGM), we have

$$P_{21}(k) = 400\,{\rm mK}\left(\frac{1+z}{7.5}\right) (1.87) x_{\rm H\,I}^2 P(k), \quad (4)$$

which includes the enhancement of fluctuations by the factor of $\langle(1+\mu^2)^2\rangle = 1.87$ owing to peculiar velocities (Barkana & Loeb 2005). Importantly, the PS is sensitive only to the first moment of $p(b)$. As a result, we do not need to calculate (or assume) a functional form for this distribution in order to estimate the constraints on $M_{\langle b\rangle}$.

Examples of spherically averaged PS are plotted in Fig. 1 (thick grey lines) assuming values of $\langle b\rangle$ corresponding to DLA halo masses of $M_{\langle b\rangle} = 10^{10}$, 10^{11} and $10^{12}\,{\rm M}_\odot$ (bottom to top), at each of the redshifts $z = 2.5$, 3.5, 4.5 and 5.5. We have assumed $x_{\rm H\,I} = 0.02$ in these models corresponding to observations of H I density in DLAs at high redshift (Prochaska et al. 2005). The wiggles at wavenumbers $k \sim 0.1\,{\rm Mpc}^{-1}$ are the baryonic oscillations that are familiar to galaxy redshift surveys (Eisenstein et al. 2005). The models with larger DLA host mass yield fluctuations with significantly more power, owing to the larger galaxy bias. It is this variation in the power that enables measurement of the DLA host mass.

3 SENSITIVITY TO FLUCTUATIONS IN 21-cm EMISSION FROM DLAS

As an example of the sensitivity of forthcoming low-frequency telescopes to the fluctuations in 21-cm emission after re-ionization, we have estimated the signal-to-noise ratio (S/N) for the MWA. Calculations of the sensitivity to the 21-cm PS for an interferometer of this sort have been presented by a number of authors. We follow the procedure outlined by McQuinn et al. (2006), drawing on results from Bowman, Morales & Hewitt (2006) for the dependence of the array antenna density on radius $\rho(r)$. The uncertainty in a measurement of the PS per mode in the survey volume has two separate components, due to the thermal noise of the instrument ($\delta P_{21,{\rm N}}$), and due to sample variance within the finite volume of the survey ($\delta P_{21,{\rm SV}}$). The former is dominated by the sky temperature

$T_{\rm sky}$ and so at frequency ν is proportional to $\nu^{-2.6}$. Since we are dealing with discrete systems rather than a diffuse IGM, the latter includes a Poisson component due to the finite sampling of each mode and equals $\delta P_{21,\rm SV} = P_{21}(k)\{1 + [\langle b\rangle^2 n_{\rm DLA} P(k)]^{-1}\}$, in which we approximate the number density of absorbing hosts by $n_{\rm DLA} \sim M{\rm d}n_{\rm ST}/{\rm d}M$, where ${\rm d}n_{\rm ST}/{\rm d}M$ is the Sheth–Tormen (Sheth & Torman 2002) mass function of dark matter haloes. The noise due to sample variance therefore depends both on the survey volume and on DLA host mass. We combine the above components to yield the uncertainty on the estimate of the PS within a k-space volume element d^3k:

$$\Delta P_{21} = (\delta P_{21,\rm SV} + \delta P_{21,\rm N})/\sqrt{N_{\rm c}}, \quad (5)$$

where the quantity $N_{\rm c} = 2\pi k^2 \sin\theta\, {\rm d}k\, {\rm d}\theta[\mathcal{V}/(2\pi)^3]$ denotes the number of modes observed within $d^3k = 2\pi k^2 \sin(\theta)\, {\rm d}k\, {\rm d}\theta$. In computing $N_{\rm c}$ we assume symmetry about the polar angle and express the wavevector $\boldsymbol{k}$ in components of its modulus k and angle θ relative to the line-of-sight.

The contamination of foregrounds provides an additional source of uncertainty in the estimate of the PS. McQuinn et al. (2006) have shown that it should be possible to remove the power due to foregrounds to a level below the cosmological signal, provided that the region of bandpass from which the PS is estimated is substantially smaller than the total bandpass available. Following the approximation suggested in McQuinn et al. (2006), we set $N_{\rm c} = 0$ if $2\pi/k\cos(\theta) > \Delta D$, where ΔD is the comoving length corresponding to the line-of-sight distance over which foregrounds can be removed. The number of modes observed depends on the volume of the survey, $\mathcal{V} = D^2 \Delta D(\lambda^2/A_{\rm tile})$, where $A_{\rm tile}$ is the total physical surface area of an antenna, and D is the comoving distance to the redshift of emission.

3.1 PS sensitivity of the MWA

Estimates of the noise for detection of the spherically averaged PS for the MWA are plotted in Fig. 1 (thin solid lines). Note that only redshifts $z \gtrsim 3.5$ are accessible with the MWA antennas design. When complete, the MWA will comprise a phased array of 512 tiles (each tile will contain 16 cross-dipoles) distributed over an area with diameter 1.5 km. To compute the noise on the PS we model the antennas distribution as having $\rho(r) \propto r^{-2}$ with a maximum radius of 750 m and a finite density core of radius 18 m. We assume a 1000 h integration on each of three fields, and a foreground removed bandpass of $B = 8$ MHz within a total processed bandpass of 32 MHz.

The sensitivity to the 21-cm PS is dependent on both the sensitivity of the telescope to a particular mode, and to the number of such modes in the survey. The former is set by the effective collecting area ($A_{\rm e}$) of each antenna element (as well as the total number of antennas), while the latter is sensitive to the total physical area covered by each antenna (which we refer to as $A_{\rm tile}$). This issue is discussed in more detail in Wyithe et al. (2008). In computing the sensitivity we have assumed $A_{\rm e} \sim 16(\lambda^2/4)\,{\rm m}^2$ and $A_{\rm tile} = 16\,{\rm m}^2$ (corresponding to the design of the MWA). The combined uncertainties include the minimum k cut-off due to foreground subtraction. Comparison of the noise estimate with the expected 21-cm signal shows that the

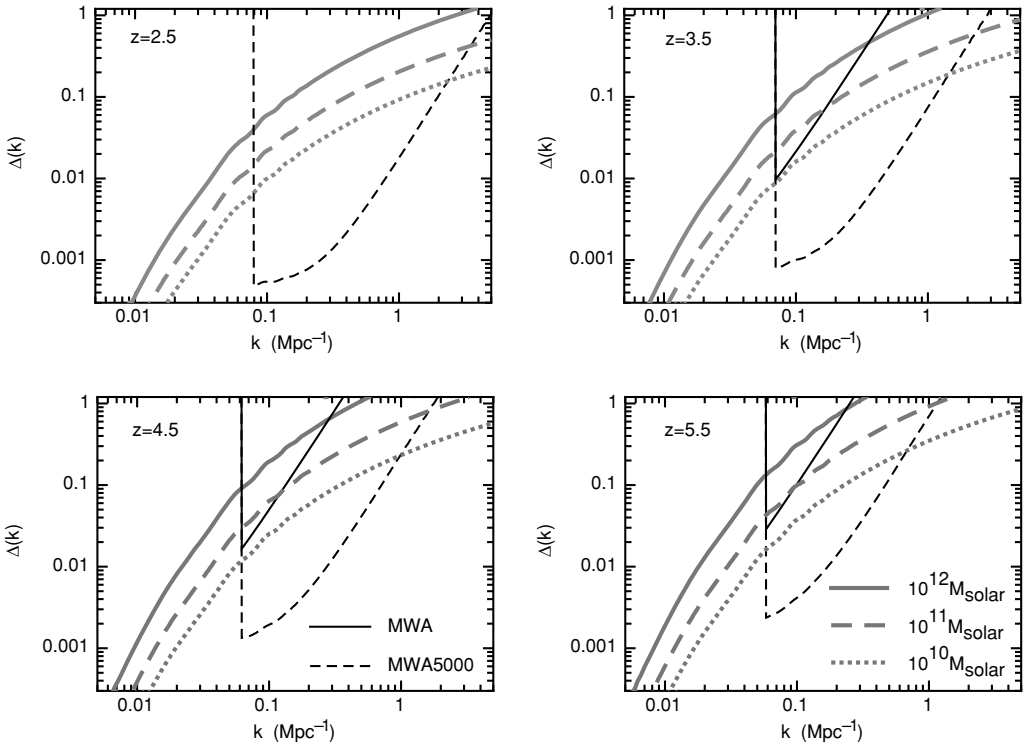

Figure 1. The spherically averaged PS of 21-cm fluctuations due to DLAs after re-ionization (thick grey lines). We have plotted the PS in dimensionless form $\Delta \equiv k^3 P(k)/(2\pi^2)$. The model PS shown was computed assuming three values for the DLA mass, $M_{\langle b\rangle} = 10^{10}$, 10^{11} and $10^{12}\,{\rm M}_\odot$. Also shown for comparison are estimates of the noise for MWA (thin solid lines) and a future extension MWA5000 (thin dashed lines). In each case we plot the uncertainty within k-space bins of size $\Delta k = k/10$, and have assumed observations for 1000 h in each of three different fields. The uncertainty includes a minimum k cut-off due to foreground subtraction. At $z < 3.5$ only the MWA5000 is considered since the MWA antennas are not effective at frequencies greater than ∼300 MHz. The noise curves, which vary at large k owing to Poisson noise, are shown only for $M_{\langle b\rangle} = 10^{12}\,{\rm M}_\odot$.

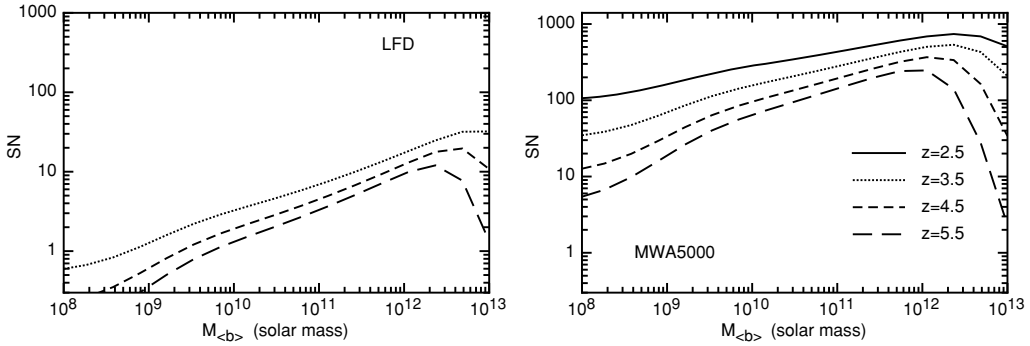

Figure 2. The integrated S/N for detection of the spherically averaged 21-cm PS as a function of DLA mass. The cases of four different redshifts are shown for each of the MWA (left-hand panel) and MWA5000 (right-hand panel).

MWA could detect the PS at $3.5 \lesssim z \lesssim 5.5$ with high significance provided that the DLA hosts are sufficiently massive (and therefore biased).

3.2 PS sensitivity of the MWA5000

At values of $k \sim$ a few $\times 10^{-1}\,\mathrm{Mpc}^{-1}$, the measurement of the PS using the MWA will be limited by the thermal sensitivity of the array, and so the S/N achievable in this regime will be greatly enhanced by a subsequent generation of telescopes with larger collecting area. As an example we consider a hypothetical follow-up telescope to the MWA which would comprise 10 times the total collecting area. We refer to this follow-up telescope as the MWA5000. The design philosophy for the MWA5000 would be similar to the MWA, and we therefore assume antennas distributed as $\rho(r) \propto r^{-2}$ with a diameter of 2 km and a flat density core of radius 80 m (see McQuinn et al. 2006). For MWA5000 we assume the antennas design to be optimized at the redshift of observation (in which case we assume $A_\mathrm{e} = A_\mathrm{tile}$). In Fig. 1 we present estimates for measurement uncertainty of the 21-cm PS using MWA5000 (thin dashed lines). An MWA5000 would achieve a high S/N detection of the PS, even in cases where the DLA hosts have a small value of galaxy bias.

3.3 Dependence of S/N on DLA host mass

In Fig. 2 we show the integrated S/N for detection of the spherically averaged PS as a function of DLA host mass. The S/N improves with increasing host mass owing to the increase in bias. Four different redshift cases are shown for each of the MWA (left-hand panel) and MWA5000 (right-hand panel). Fig. 2 quantifies the results described in Fig. 1. The MWA can detect the spherically averaged PS at a S/N ≥ 3 for DLA masses $M_{\langle b\rangle} \gtrsim 10^{11}\,\mathrm{M}_\odot$, while the MWA5000 could detect the PS at a S/N ≥ 3 for DLA masses as small as $M_{\langle b\rangle} \sim 10^{8}\,\mathrm{M}_\odot$. In each case S/N is limited at the highest masses by the Poisson noise introduced into the PS by the discreteness of the DLA emission.

As part of this procedure, the model PS which assumes a linear extrapolation of the primordial PS is compared to an observed PS. Since the observed PS will include non-linear evolution at sufficiently large k, a restriction is placed on the maximum value of k used in estimation of the S/N via equation (6). Through studying the PS derived from N-body simulations it has been shown (Seo & Eisenstein 2005) that with respect to the baryonic acoustic oscillations, the PS can be treated as linear on scales corresponding to $k < 0.4\,\mathrm{Mpc}^{-1}$ at $z = 3$, with this limiting scale increasing towards higher redshifts.[6] We therefore restrict our fitting to wavenumbers $k < 0.4\,\mathrm{Mpc}^{-1}$. The same restriction is placed for the calculations in subsequent sections.

4 MEASUREMENT OF DLA HOST MASS USING SPHERICALLY AVERAGED 21-cm PS

In this section we consider the spherically averaged PS, which would be the observable of choice for 21-cm PS observations that achieve only modest S/N. To estimate the potential for constraints on the DLA host mass properties from the spherically averaged PS, we first calculate the regions of parameter space $\boldsymbol{p} = (x_{\mathrm{H\,I}}, M_{\langle b\rangle})$ that are allowed around true models with $\boldsymbol{p}_\mathrm{o}$. We assume there is no uncertainty in the PS of matter fluctuations, which would increase the uncertainty somewhat by introducing distortions on to the PS via assumption of an incorrect cosmology. As part of this procedure, we construct likelihoods

$$\ln \mathcal{L}(\boldsymbol{p}) = -\frac{1}{2}\sum_k \left[\frac{P_{\Delta T}(k,\,\boldsymbol{p}) - P^\mathrm{t}_{\Delta T}(k,\,\boldsymbol{p}_\mathrm{o})}{\Delta P^\mathrm{sph}_{\Delta T}(k)}\right]^2, \qquad (6)$$

where the sum is over bins of k, and $\Delta P^\mathrm{sph}_{\Delta T}$ is the spherically averaged uncertainty on measurement of the 21-cm PS.

In Fig. 3 we show the likelihood contours for combinations of $x_{\mathrm{H\,I}}$ and $M_{\langle b\rangle} = M(\langle b\rangle)$ given input models with $M_{\langle b\rangle} = 10^{12}\,\mathrm{M}_\odot$ (upper panels) and $M_{\langle b\rangle} = 10^{10}\,\mathrm{M}_\odot$ (lower panels), combined with $x_{\mathrm{H\,I}} = 0.02$ and observation with the MWA (diagonal sets of contours). The contours illustrate the degeneracy between $x_{\mathrm{H\,I}}$ and $M_{\langle b\rangle}$ in the resulting PS. Also shown is the existing constraint on the neutral fraction from quasar absorption line studies (vertical sets of contours showing the current uncertainty within $\Delta z = 0.5$ bins; Prochaska et al. 2005). The figure illustrates how the neutral fraction constraint can be used to break the degeneracy. It is clear from Fig. 3 that the constraint on $x_{\mathrm{H\,I}}$ will be the limiting factor for measurement of the DLA host mass via the spherically averaged PS (in cases where the host bias is large).

From equation (3) we see that measurement of the spherically averaged PS with a S/N yields a fractional error on the combination $b_\mathrm{DLA}\,x_{\mathrm{H\,I}}$ with value

$$\frac{\Delta(b_\mathrm{DLA}x_{\mathrm{H\,I}})}{b_\mathrm{DLA}x_{\mathrm{H\,I}}} = \frac{1}{2\,\mathrm{S/N}}, \qquad (7)$$

[6] Additional non-linear effects may arise that are not present in galaxy redshift surveys owing to internal rotation curve of the neutral hydrogen. However, the corresponding wavenumbers are very large, $k \gg 1\,\mathrm{Mpc}^{-1}$.

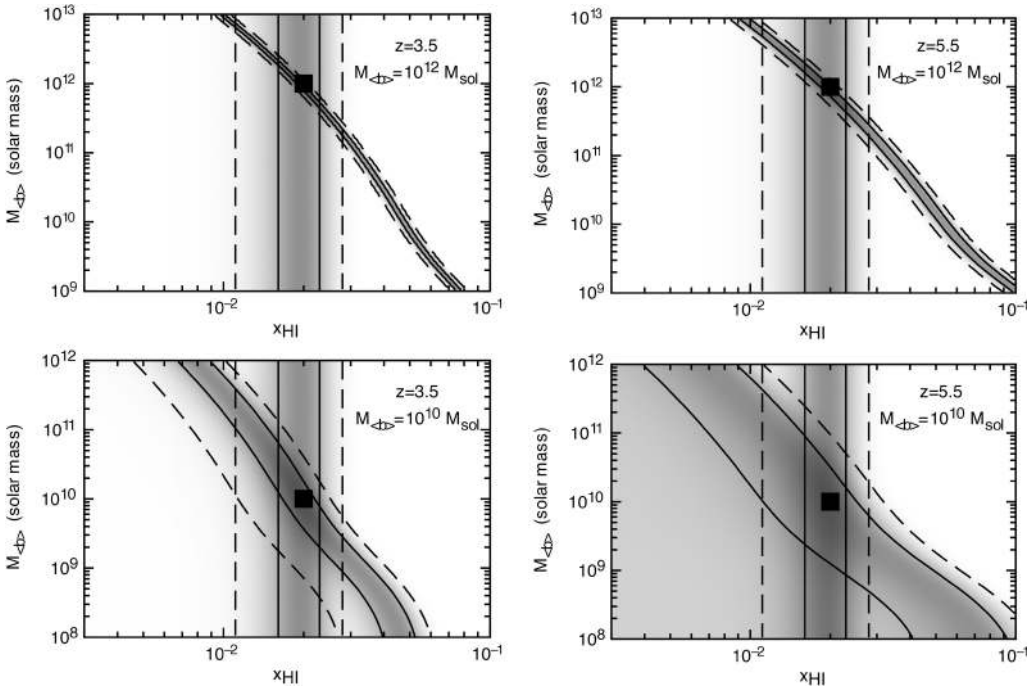

Figure 3. Constraints on DLA host mass and neutral fraction from a spherically averaged 21-cm PS. The diagonal contours show the loci of points with likelihoods corresponding to 64 per cent (solid lines) and 10 per cent (dashed lines) of the maximum, given a true model with DLA masses of $10^{12}\,\mathrm{M}_\odot$ (upper panels) and $10^{10}\,\mathrm{M}_\odot$ (lower panels) combined with a neutral fraction $x_{\mathrm{H\,I}} \sim 0.02$ and observation with the MWA. The vertical lines show the 64 and 10 per cent likelihood contours for neutral fraction based on existing observations of DLAs within a redshift bin of width $\delta z = 0.5$.

where we have defined the quantity $b_{\mathrm{DLA}} = \sqrt{\langle b\rangle^2 + 2/3\,\langle b\rangle + 1/5}$. In deriving equation (7) we have assumed the mass PS to have negligible uncertainty. The measurement of neutral fraction with uncertainty $\Delta x_{\mathrm{H\,I}}$ is independent of the PS (having been obtained via the column density in DLAs along the line-of-sight towards high-redshift quasars). The uncertainty in the characteristic bias (Δb_{DLA}) is therefore

$$\Delta b_{\mathrm{DLA}} = b_{\mathrm{DLA}} \sqrt{\left[\frac{\Delta(b_{\mathrm{DLA}} x_{\mathrm{H\,I}})}{b_{\mathrm{DLA}} x_{\mathrm{H\,I}}}\right]^2 + \left(\frac{\Delta x_{\mathrm{H\,I}}}{x_{\mathrm{H\,I}}}\right)^2}. \quad (8)$$

The corresponding uncertainty in $\langle b\rangle$ is

$$\Delta\langle b\rangle = \frac{\mathrm{d}\langle b\rangle}{\mathrm{d}b_{\mathrm{DLA}}}\Delta b_{\mathrm{DLA}} = \frac{b_{\mathrm{DLA}}}{\langle b\rangle + 1/3}\Delta b_{\mathrm{DLA}}, \quad (9)$$

yielding the uncertainty in host mass (in dex):

$$\Delta(\log M_{\langle b\rangle}) = \frac{\mathrm{d}\log M_{\langle b\rangle}}{\mathrm{d}\langle b\rangle}\Delta\langle b\rangle, \quad (10)$$

where $\mathrm{d}(\log M_{\langle b\rangle})/\mathrm{d}\langle b\rangle$ is determined via equation (1). Before presenting examples of the level of uncertainty that could be achieved by upcoming experiments, we reiterate that the measured quantity will be $\langle b\rangle$ while the quantity of physical interest is the host mass $M_{\langle b\rangle}$. Since these quantities are connected through a theoretical expression (equation 1) we note that there is a component of systematic theoretical uncertainty associated with its accuracy. However, we do not pursue the magnitude of this uncertainty in the illustrative examples shown in this paper.

The values of $\Delta(\log M_{\langle b\rangle})$ that are obtained via equation (10) represent the dispersion of a likelihood function $\mathcal{L}(\log M_{\mathrm{obs}}|\log M_{\langle b\rangle})$ for observation of $\log M_{\mathrm{obs}}$ given a true value $\log M_{\langle b\rangle}$. Following the observation of a 21-cm DLA PS the likelihood function could then be used to obtain an a posteriori measurement of the DLA host mass. Thus, in the instance of a flat prior probability distribution for the logarithm of host mass, the dispersion approximates the error (in dex) achievable for the mass of the DLA host given the true mass $M_{\langle b\rangle}$. The dispersion from equation (10) is plotted in Fig. 4 as a function of host mass, at a range of redshifts, and for different assumptions regarding the uncertainty in neutral fraction and properties of the 21-cm telescope used [the MWA (left-hand panels) and MWA5000 (right-hand panels)]. The fractional uncertainty in $x_{\mathrm{H\,I}}$ was assumed to be $\Delta x_{\mathrm{H\,I}}/x_{\mathrm{H\,I}} = 0.2$ (upper panels; corresponding to one $\delta z = 0.5$ bin in the SDSS DR3 study of Prochaska et al. 2005) and $\Delta x_{\mathrm{H\,I}}/x_{\mathrm{H\,I}} = 0.09$ (central panels; corresponding to a combination of all five $\delta z = 0.5$ bins in the Prochaska et al. 2005 study). With current observations of DLA absorption systems, the host mass could be determined to within a factor as small as 2–3, provided that the hosts are massive.

It is apparent from equation (8) that the measurement precision of the DLA clustering bias (and therefore the DLA host mass) is limited by the uncertainty in neutral fraction once $\mathrm{S/N} \gg (x_{\mathrm{H\,I}}/\Delta x_{\mathrm{H\,I}})/2 \gtrsim 3$–5. Thus, with current observations of quasar absorption systems, the DLA host mass estimate using an spherically averaged PS would not be limited by the PS measurement over a range of masses for the MWA. Moreover, the mass estimate would not be limited by the PS measurement at any host mass using the MWA5000. We therefore show additional results for the case where the error in measurement of $\langle b\rangle$ is dominated by the uncertainty in determination of the 21-cm PS rather than the neutral fraction. The resulting values of $\Delta(\log M_{\langle b\rangle})$ are shown in the lower panels of Fig. 4. The curves suggest that the achievable accuracy would be limited by the determination of $\Omega_{\mathrm{H\,I}}$ for host masses larger than $\sim 10^{10}\,\mathrm{M}_\odot$ using the MWA. In this most optimistic case the MWA5000 could measure the characteristic DLA host mass to the level of a few per cent over the entire plausible mass range.

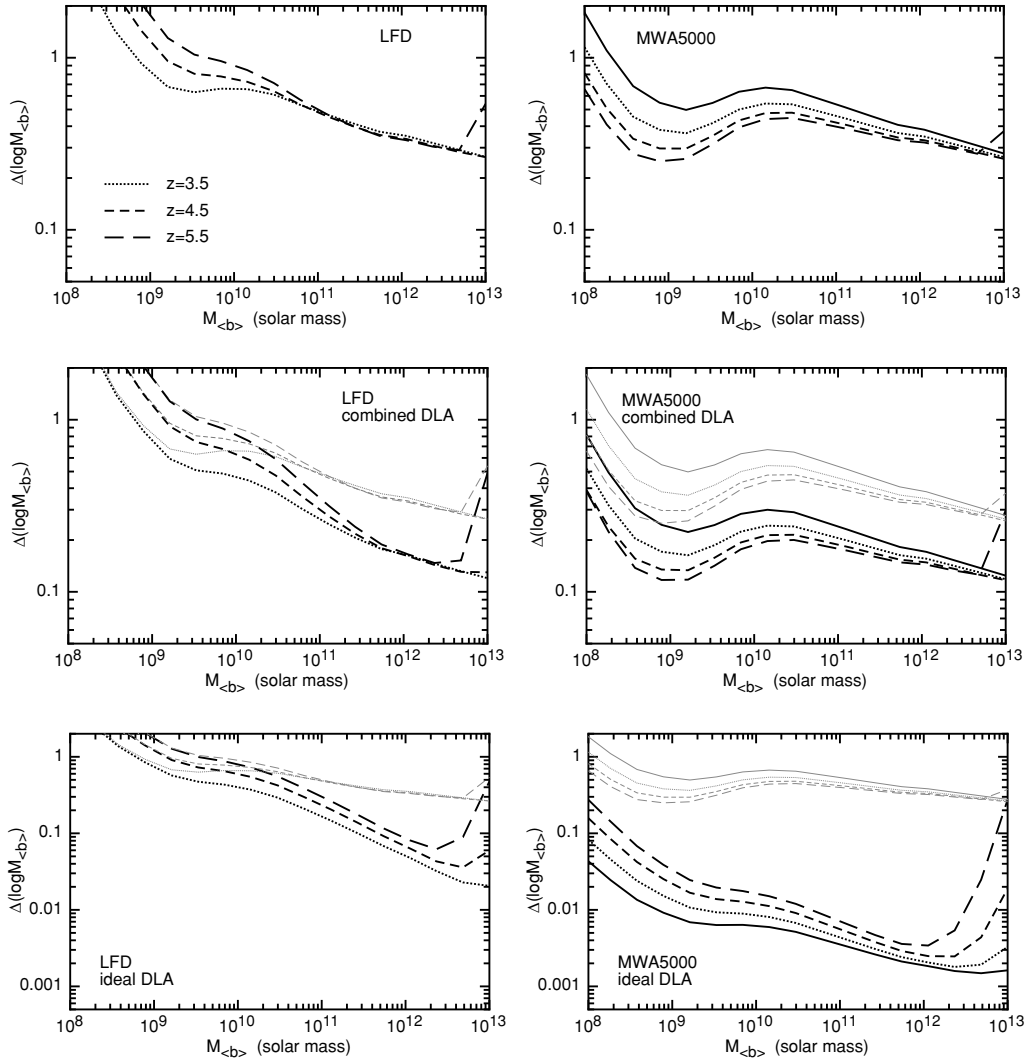

Figure 4. The uncertainty (in dex) on the mass of the DLA host achievable using an spherically averaged 21-cm PS. The cases of different redshifts are shown for each of the MWA (left-hand panels) and MWA5000 (right-hand panels). Upper panels: the fractional uncertainty in $x_{\rm H\,I}$ was assumed to be $\Delta x_{\rm H\,I}/x_{\rm H\,I} = 0.2$ (corresponding to one $\delta z = 0.5$ bin in the SDSS study of Prochaska et al. 2005). Central panels: the fractional uncertainty in $x_{\rm H\,I}$ was assumed to be $\Delta x_{\rm H\,I}/x_{\rm H\,I} = 0.09$ (corresponding to a combination of all five $\delta z = 0.5$ bins in the SDSS study, or a future survey with smaller error bars). Lower panels: the fractional uncertainty in $x_{\rm H\,I}$ was assumed to be negligible. In the central and lower panels, the curves for $\Delta x_{\rm H\,I}/x_{\rm H\,I} = 0.2$ are reproduced for comparison (thin grey curves).

5 MEASUREMENT OF THE DLA HOST MASS USING ANISOTROPY OF THE OBSERVED 21-cm PS

In the previous section we discussed the constraints that could be placed on the DLA host mass via an spherically averaged 21-cm PS. That analysis broke the degeneracy between mass averaged neutral fraction and the galaxy bias of DLA hosts using an independent measurement of neutral fraction from the H I column density in quasar absorption systems. In this section we use the peculiar velocity induced anisotropy of the observed PS to break the degeneracy between bias and neutral fraction using the fact that these quantities contribute differently to the PS viewed at different angles relative to the line-of-sight (equation 2).

To estimate the potential for constraints on the DLA host mass, we again calculate the region of parameter space $\boldsymbol{p} = (x_{\rm H\,I}, M_{\langle b\rangle})$ that is allowed around a true solution with $\boldsymbol{p}_{\rm o}$. We assume there is no uncertainty in the PS of matter fluctuations, which would increase the uncertainty somewhat by introducing distortions on to the PS via assumption of an incorrect cosmology. As part of this procedure, we construct likelihoods

$$\ln \mathcal{L}(\boldsymbol{p}) = -\frac{1}{2}\sum_{k,\mu}\left[\frac{P_{\Delta T}(k,\mu,\boldsymbol{p}) - P^{\rm t}_{\Delta T}(k,\mu,\boldsymbol{p}_{\rm o})}{\Delta P_{\Delta T}(k,\mu)}\right]^2, \qquad (11)$$

where the sum is over bins of k and μ.

The results are shown in Figs 5 and 6 for the MWA and MWA5000, respectively, using the observing strategy outlined in Section 4. In each case likelihood contours are shown for the parameter set $(x_{\rm H\,I}, M_{\langle b\rangle})$ at two redshifts, $z = 2.5$ and 3.5, assuming a true value of $x_{\rm H\,I} = 0.02$ and three true values of host mass $M_{\langle b\rangle} = 10^{10}$, 10^{11} and $10^{12}\,{\rm M}_\odot$. Also shown are the corresponding likelihoods for $M_{\langle b\rangle}$. We have assumed a flat prior probability for $M_{\langle b\rangle}$. The S/N of the MWA is not sufficiently large to break the degeneracy between bias and neutral fraction. Therefore, as in the previous section we impose a Gaussian prior probability on $x_{\rm H\,I}$,

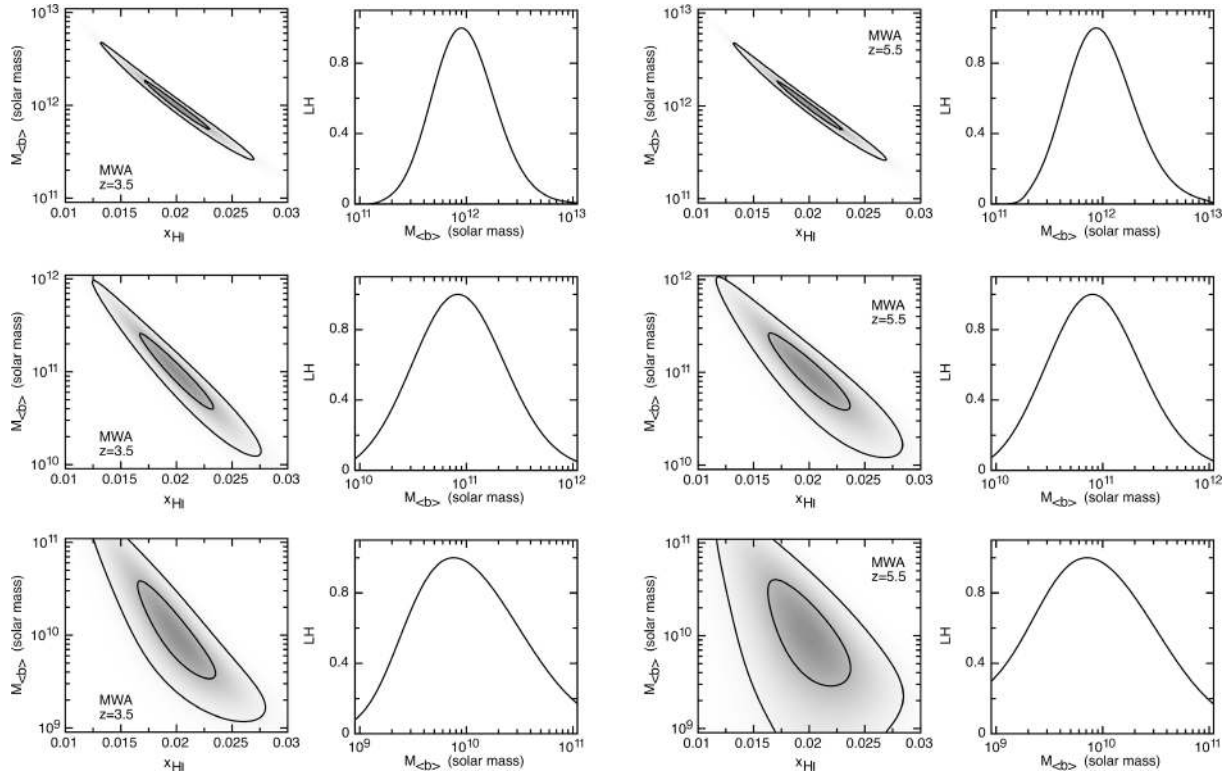

Figure 5. Constraints on the DLA host masses derived from measurement of the anisotropy of the 21-cm PS. We assume observations using the MWA and three fields integrated for 1000 h each. Results are shown at two redshifts, and for three values of the true host mass ($M_{\langle b\rangle} = 10^{10}$, 10^{11} and $10^{12}\,\mathrm{M}_{\odot}$; from bottom to top). We assume a true value of $x_{\mathrm{H\,I}} = 0.02$. In each case likelihood contours are shown at 10 and 64 per cent of the maximum for the parameter set ($x_{\mathrm{H\,I}}$, $M_{\langle b\rangle}$). These contours represent the ellipses containing 35 and 90 per cent of normally distributed data, respectively. Also shown are the corresponding likelihoods for $M_{\langle b\rangle}$. We have assumed a flat prior probability for $M_{\langle b\rangle}$, and a Gaussian prior probability for the neutral fraction with a variance of $\Delta x_{\mathrm{H\,I}} = 0.2 x_{\mathrm{H\,I}}$.

with a variance of $\Delta x_{\mathrm{H\,I}} = 0.2\,x_{\mathrm{H\,I}}$, corresponding to current uncertainty from quasar absorption line studies (Prochaska et al. 2005). Fig. 5 shows that at the sensitivity of the MWA, the constraints on the DLA host mass are very similar to those obtained from the spherically averaged PS. As a result nothing is gained by considering the angular dependence. On the other hand, the degeneracy is broken at the S/N obtained by the MWA5000. As a result, we have not imposed a prior probability on the neutral fraction from quasar absorption studies in this case, and the allowed values of $x_{\mathrm{H\,I}}$ in Fig. 6 are derived directly from the 21-cm PS, in addition to $M_{\langle b\rangle}$. Fig. 6 shows that the MWA5000 would obtain accuracies of a few per cent on the neutral fraction, and a few 10s of per cent in the DLA host mass.

Our discussion throughout this paper has assumed that the underlying PS is perfectly known, so that the uncertainties on measured quantities like $M_{\langle b\rangle}$ are dominated by the 21-cm observations, rather than by cosmology. However, it should be noted that uncertainty in the PS amplitude (which is proportional to the normalization of the primordial PS, σ_8) is degenerate[7] with x_{H}. As a result, in cases where the fractional error in the measured PS amplitude is smaller than the fractional error in σ_8, the precision that can be obtained on $x_{\mathrm{H\,I}}$ and hence on $M_{\langle b\rangle}$ are limited by cosmological constraints. On the other hand, measurements of the underlying cosmology do not add additional uncertainty relative to the constraints shown in Figs 5 and 6 in cases where the fractional uncertainty in the PS amplitude is larger than the fractional uncertainty in σ_8. Fig. 6 shows that under the assumption of a known primordial PS, the value of $x_{\mathrm{H\,I}}$ will be constrained at the 10 per cent level with MWA5000. Since the precision of current constraints on σ_8 are already comparable to this level (e.g. Spergel et al. 2007), and that future experiments will substantially improve on this precision, we have not included uncertainty on σ_8 into the illustrative constraints in this analysis.

[7] The direct degeneracy is between $x_{\mathrm{H\,I}}$ rather than $\langle b\rangle$ because $\langle b\rangle$ appears in a factor which also depends on angle relative to the line-of-sight.

6 CONCLUSIONS

Observations of DLA systems out to redshift $z \sim 5$ show the density parameter of H I to be $\Omega_{\mathrm{H\,I}} \sim 10^{-3}$, indicating that the mass averaged neutral hydrogen fraction remains at the level of a few per cent through most of cosmic history. These DLA systems will produce 21-cm intensity fluctuations, whose PS has an amplitude that depends both on the total mass of neutral hydrogen within the DLA systems ($\Omega_{\mathrm{H\,I}}$), and the masses of the DLA host haloes (through galaxy bias). Since $\Omega_{\mathrm{H\,I}}$ is measured via quasar absorption line studies, we show that measurement of the spherically averaged 21-cm PS amplitude at $2 \lesssim z \lesssim 5$ could be used to determine the DLA host mass. Using a telescope of collecting area equal to the MWA, the DLA host mass could be determined to within a factor as small as $\sim$2 (provided the hosts were massive galaxies), with the accuracy limited by the determination of $\Omega_{\mathrm{H\,I}}$ in this case.

In observations of the 21-cm PS with high S/N, the limitation of the quasar absorption line derived neutral fraction could be removed

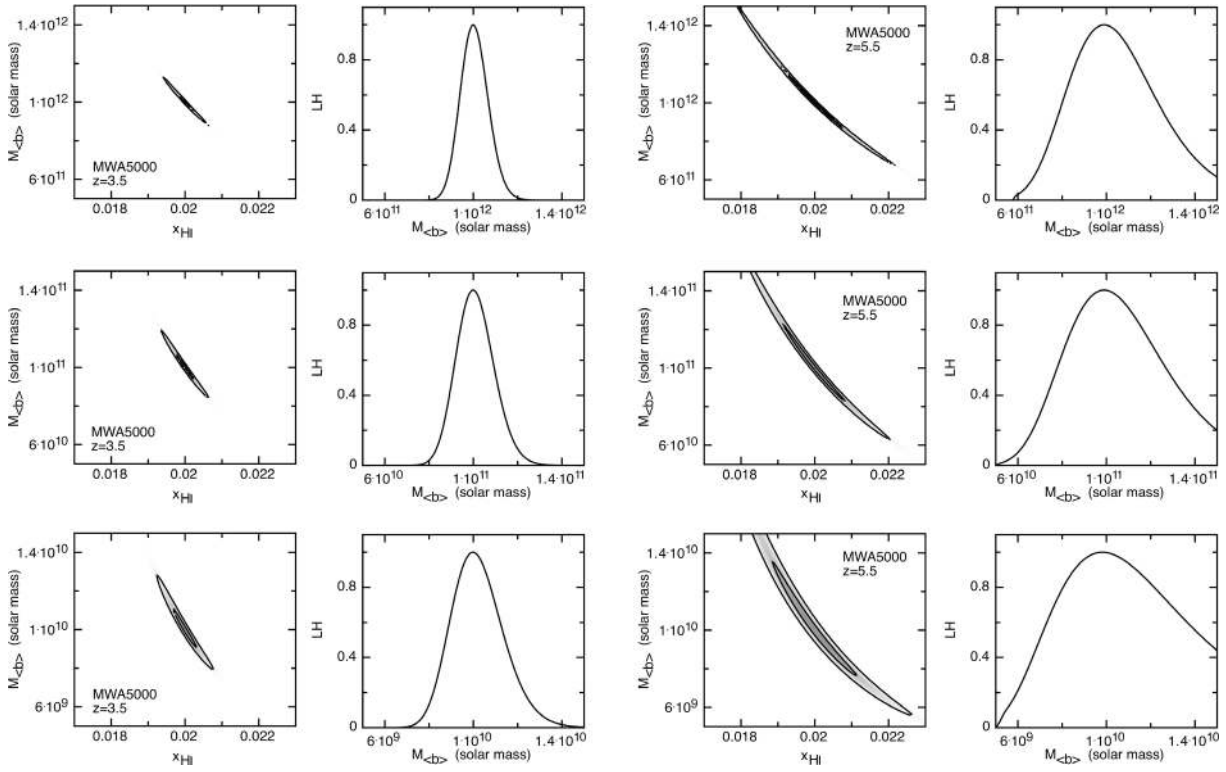

Figure 6. As per Fig. 5, but assuming observation using the MWA5000, and no prior constraint on $x_{\rm H\,I}$.

by including the observed angular dependence due to peculiar motions in the analysis. In this case the measurement of DLA host mass would be limited by the sensitivity to the 21-cm PS, and future telescopes with larger collecting area would significantly increase the precision (to within a few 10s of per cent) with which the DLA host mass could be determined.

The DLA host mass is currently estimated via clustering analyses of absorption systems which are very rare, only being found in one of every few quasar lines-of-sight. Thus measurement of the 21-cm PS, which measures the clustering of all DLAs in a large volume, has the potential to greatly increase the precision with which the DLA host mass is known. Since DLAs are thought to host the majority of gas available for star formation, measurement of the DLA host mass will be a valuable contribution to our understanding of galaxy formation and the star formation history.

ACKNOWLEDGMENTS

The author would like to thank Avi Loeb for pointing out an error, and Michael Murphy for helpful comments on the draft manuscript. This research was supported by the Australian Research Council.

REFERENCES

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, ApJ, 619, 697
Barkana R., Loeb A., 2005, ApJ, 624, L65
Barnes D. G. et al., 2001, MNRAS, 322, 486
Bouché N., Murphy M. T., Péroux C., Csabai I., Wild V., 2006, MNRAS, 371, 495
Bouché N., Murphy M. T., Péroux C., Davies R., Eisenhauer F., Förster Schreiber N. M., Tacconi L., 2007, ApJ, 669, L5
Bowman J. D., Morales M. F., Hewitt J. N., 2006, ApJ, 638, 20
Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett., 100, 091303
Colbert J. W., Malkan M. A., 2002, ApJ, 566, 51
Cooke J., Wolfe A. M., Gawiser E., Prochaska J. X., 2006, ApJ, 652, 994
Curran S. J., Tzanavaris P., Murphy M. T., Webb J. K., Pihlström Y. M., 2007, MNRAS, 381, L6
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Furlanetto S. R., Loeb A., 2002, ApJ, 579, 1
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Jones D. H. et al., 2004, MNRAS, 355, 747
Jones D. H., Saunders W., Read M., Colless M., 2005, Publ. Astron. Soc. Aust., 22, 277
Kaiser N., 1987, MNRAS, 227, 1
Kanekar N., Chengalur J. N., 2003, A&A, 399, 857
Loeb A., Wyithe S., 2008, Phys. Rev. Lett., 100, 161301
McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
Mao Y., Tegmark M., McQuinn M., Zaldarriaga M., Zahn O., 2008, preprint (arXiv:0802.1710)
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Nagamine K., Wolfe A. M., Hernquist L., Springel V., 2007, ApJ, 660, 945
Pen U.-L., Staveley-Smith L., Peterson J., Chang T.-C., 2008, preprint (arXiv:0802.3239)
Press W., Schechter P., 1974, ApJ, 187, 425
Pritchard J. R., Loeb A., 2008, preprint (arXiv:0802.2102)
Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123
Rao S. M., Turnshek D. A., Nestor D. B., 2006, ApJ, 636, 610
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Seo H.-J., Eisenstein D. J., 2005, ApJ, 633, 575
Shen Y. et al., 2007, AJ, 133, 2222
Sheth R., Torman G., 2002, MNRAS, 321, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Spergel D. N. et al., 2007, ApJS, 170, 377
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728
Warren S. J., Møller P., Fall S. M., Jakobsen P., 2001, MNRAS, 326, 759
White M., Martini P., Cohn J. D., 2007, preprint (arXiv:0711.4109)
Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861
Wyithe J. S. B., Loeb A., 2008, MNRAS, 383, 606
Wyithe J. S. B., Morales M. F., 2007, MNRAS, 379, 1647
Wyithe J. S. B., Loeb A., Geil P. M., 2008, MNRAS, 383, 1195

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