Galaxy redshift surveys selected by neutral hydrogen using the Five-hundred metre Aperture Spherical Telescope

Alan R. Duffy,[1,2]⋆ Richard A. Battye,[1] Rod D. Davies,[1] Adam Moss[3] and Peter N. Wilkinson[1]

[1] *Jodrell Bank Observatory, School of Physics and Astronomy, University of Manchester, Macclesfield, Cheshire SK11 9DL*
[2] *Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*
[3] *Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*

ABSTRACT

We discuss the possibility of performing a substantial spectroscopic galaxy redshift survey selected via the 21-cm emission from neutral hydrogen using the Five-hundred metre Aperture Spherical Telescope (FAST) to be built in China. We consider issues related to the estimation of the source counts and optimizations of the survey, and discuss the constraints on cosmological models that such a survey could provide. We find that a survey taking around two years could detect $\sim 10^7$ galaxies with an average redshift of ~ 0.15 making the survey complementary to those already carried out at optical wavelengths. These conservative estimates have used the $z = 0$ H I mass function and have ignored the possibility of evolution. The results could be used to constrain $\Gamma = \Omega_{\mathrm{m}} h$ to 5 per cent and the spectral index, n_{s}, to 7 per cent independent of cosmic microwave background data. If we also use simulated power spectra from the *Planck* satellite, we can constrain w to be within 5 per cent of -1.

Key words: telescopes – surveys – cosmological parameters – radio lines: galaxies.

1 INTRODUCTION

Galaxy redshift surveys have played a significant role in constraining the cosmological model, with the largest surveys to date having been performed by the 2dFGRS[1] and SDSS[2] teams using optical techniques. Low-redshift samples have been used for sometime to derive constraints on cosmological parameter combinations Γ and $f_{\mathrm{b}} = \Omega_{\mathrm{b}}/\Omega_{\mathrm{m}}$, where Ω_{m} and Ω_{b} are the total matter and baryon densities defined relative to critical, and $h = H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})$, as well as the spectral index of the density fluctuations, n_{s}, and neutrino densities (Percival et al. 2001; Tegmark et al. 2004a,b; Cole et al. 2005). More recently, the SDSS team have used a more sparsely sampled, but larger volume survey of luminous red galaxies (LRGs) to measure the expected baryon acoustic peak on scales of $\sim 100 h^{-1}$ Mpc (Eisenstein et al. 2005). Since the acoustic scale is a standard ruler, this provides a distance measurement allowing further constraints to be placed on cosmological parameters and, in particular, those which parametrize the properties of the dark energy.

A large number of projects have been proposed recently to develop these ideas and to perform large photometric surveys in the optical (and possibly the infrared) waveband to find $\sim 10^8$ galax-ies out to $z \sim 1$ (Dark Energy Survey, hereafter DES[3]), as well as deeper spectroscopic surveys to find $\sim 10^5$ galaxies at $z \sim 3$ (Blake & Glazebrook 2003). Such surveys will have much larger search volumes than those presently available, allowing for significantly more accurate measurements of the power spectrum and hence of the acoustic scale at a variety of redshifts. This could lead to very tight constraints on the properties of the dark energy.

Another promising technique is to use the 21-cm emission from neutral hydrogen (H I) as a tracer of the galaxies allowing for a spectroscopic galaxy redshift survey (Abdalla & Rawlings 2004). This will require the development of telescopes with much higher survey speed, and hence a combination of larger collecting area and field of view (FoV), than those available at present. The current state-of-the-art is a shallow all-sky survey; the combined HIPASS (Zwaan et al. 2005) and HIJASS (Lang et al. 2003) survey, which has found $\sim 10^4$ galaxies. Fortunately, a worldwide technology development programme is underway with the ultimate aim of building a Square Kilometre Array (SKA[4]) with a collecting area $\sim 10^6$ m^2 and an FoV much greater than 1 deg^2, with these specifications the SKA surveys may detect $\sim 10^9$ galaxies.

Before the full SKA comes into operation (~ 2020) another large radio telescope will have been constructed which also has great potential for H I surveys. The Five-hundred metre Aperture Spherical

⋆E-mail: ad@jb.man.ac.uk
[1] 2dF homepage: www.aao.gov.au/AAO/2df
[2] SDSS homepage: www.sdss.org
[3] DES homepage: www.darkenergysurvey.org
[4] SKA homepage: www.skatelescope.org/documents

Telescope (FAST) will be an Arecibo-like telescope with a significantly larger aperture (~500 m) and FoV (Nan 2006). The aim of this paper is to discuss the scientific potential of a low-redshift H I survey of galaxies as a FAST key science programme. We will show that, if a focal plane array with 100 instantaneous beams can be developed, $\sim10^7$ galaxies can be detected within a realistic survey period (~2 yr) and that the low-redshift galaxy power spectrum can be measured much more accurately than is currently possible with optical surveys. Since hydrogen is the fundamental baryonic constituent of the Universe such a survey would provide an important view of the large-scale structure of the Universe with different biases and be an important stepping stone to the larger H I galaxy surveys which will become possible with the SKA.

In the next section we will briefly discuss some relevant observational parameters of FAST and then in subsequent sections, we will estimate the number counts, make forecasts for errors on the power spectrum and constraints on cosmological parameters that could be possible with such a survey.

2 FIVE-HUNDRED METRE APERTURE SPHERICAL TELESCOPE (FAST)

FAST is funded and when completed in around 2012–2013 it will be the largest single dish telescope in the world. The expected specifications of the telescope (Nan 2006) relevant to our discussion here are as follows.

(i) The prime-focus feed illuminates an area corresponding to a parabolic dish of diameter ~300 m with aperture efficiency of 70 per cent giving $A_{\text{eff}} = 50\,000$ m^2,

(ii) At 21 cm wavelength the beam width [(full width at half-maximum (FWHM)] $\theta \approx 3$ arcmin; the effective pixel (equating the volume of a cube of height unity to the integral under a gaussian beam with the same height) has side 1.064θ and area $1.133\theta^2$.

(iii) The maximum observable zenith angle is of the order of $40°$ allowing the observation of ≈50 per cent of the full sky without significant degradation of performance. Henceforth, the 'full sky' is defined to be the maximum amount of sky accessible to the telescope, that is, $\approx20\,000$ deg^2.

(iv) One of the principal receiver systems will be a focal plane array covering a frequency range around 1.4 GHz (the 21-cm H I line) which can produce n_{B} instantaneous beams each with a system temperature $T_{\text{sys}} = 25$ K (hence $A_{\text{eff}}/T_{\text{sys}} \approx 2000$ m^2 K^{-1} per beam) and dual polarization capability.

The expected thermal noise for a dual polarization single beam can be computed using

$$\sigma_{\text{noise}} = \sqrt{2}\frac{kT_{\text{sys}}}{A_{\text{eff}}}\frac{1}{\sqrt{\Delta\nu t}}, \tag{1}$$

for an observing time of t and a frequency bandwidth of $\Delta\nu$, where $k = 1380$ Jy m^2 K^{-1} is the Boltzmann constant. If we assume a bandwidth of $\Delta\nu = 1$ MHz which corresponds to a velocity linewidth of ≈200 km s^{-1} at 21 cm, then the instantaneous sensitivity of each beam of the FAST system will be $\sigma_{\text{inst}} \approx 1$ mJy s$^{1/2}$.

The focal ratio of the telescope ($f/D = 0.47$) restricts n_{B} to $\leqslant19$ for a conventional close-packed horn-based array, similar to that used for HIPASS and operated at the prime focus. Larger arrays require some of the horns to be far enough off-axis that aberration losses become unacceptable. A 19-beam array gives an instantaneous FoV of $\Omega_{\text{inst}} \approx 200$ arcmin^2. We will use this 19-beam system as the fiducial benchmark for our basic calculations. A receiver array with $n_{\text{B}} \sim 100$ may well be possible. This requirement could

only be met with close-packed phased arrays of small antenna elements; these would not take up much more area than the 19-beam horn system. The required phased array technology is actively being developed by several groups working within the international Square Kilometre Array R&D effort[5] and hence we can be confident that by the time FAST comes into operation in 2012–2013 the 100-beam receiver capability will be available. This would allow for $\Omega_{\text{inst}} \approx (n_{\text{B}}/19)200$ arcmin^2. Such a system will be at least twice as sensitive per beam as that on the Arecibo Telescope and have a survey speed $[\Omega_{\text{inst}}(A_{\text{eff}}/T_{\text{sys}})^2]$ which is at least $6(n_{\text{B}}/19)$ times faster than the seven-beam system presently available there. FAST will also have more frequency flexibility due to the excellent radio frequency interference (RFI) environment at the FAST site and will be able to observe twice the area of the sky than does the Arecibo Telescope.

3 ESTIMATING GALAXY COUNTS

3.1 Properties of the H I galaxy distribution

The H I mass, M_{HI}, of a galaxy at redshift z is given in terms of the observed flux, S, and linewidth, ΔV_{o}, by (Roberts 1975):

$$\frac{M_{\text{HI}}}{M_{\odot}} = \frac{2.35 \times 10^5}{1+z}\left[\frac{d_{\text{L}}(z)}{\text{Mpc}}\right]^2\left(\frac{S}{\text{Jy}}\right)\left(\frac{\Delta V_{\text{o}}}{\text{km s}^{-1}}\right), \tag{2}$$

where $d_{\text{L}}(z)$ is the luminosity distance to the galaxy. We note that this formula is usually quoted in terms of the Euclidean distance and without the $(1 + z)^{-1}$; this is the correct formula in a Friedmann–Robertson–Walker (FRW) universe.

One can estimate the number of galaxies above some limiting mass $M_{\text{lim}}(z)$ by computing

$$N(M > M_{\text{lim}}, z) = \Delta\Omega\Delta z\frac{\text{d}V}{\text{d}z\text{d}\Omega}\int_{M_{\text{lim}}(z)}^{\infty}\frac{\text{d}N}{\text{d}V\text{d}M}\,\text{d}M, \tag{3}$$

where the sky area covered is $\Delta\Omega$ and the size of the redshift bin is Δz. $\text{d}V/\text{d}z\text{d}\Omega$ is the comoving volume element for the FRW universe and $\text{d}N/\text{d}V\text{d}M$ is the comoving number density of galaxies per unit mass which we shall assume can be approximated by a Schechter function

$$\frac{\text{d}N}{\text{d}V\text{d}M} = \frac{\theta^*}{M_{\text{HI}}^*}\left(\frac{M}{M_{\text{HI}}^*}\right)^{\alpha}\exp\left(-\frac{M}{M_{\text{HI}}^*}\right). \tag{4}$$

We will, for the most part, assume no evolution in the H I mass function out to the moderate redshifts appropriate for a FAST survey and, therefore, use values of the parameters derived from the most recent survey of the local Universe (Zwaan et al. 2005). In particular, we will choose $\alpha = -1.37$, $\theta^* = 1.42 \times 10^{-2}h^3$ Mpc^{-3} and $M_{\text{HI}}^* = 10^{9.8}$ M$_{\odot}$.

One can compute the average redshift of galaxies in the survey from $N(M > M_{\text{lim}}, z)$ by integrating appropriately over z, that is,

$$\langle z \rangle = \frac{\int_0^{\infty} z\,N(M > M_{\text{lim}}, z)\,\text{d}z}{\int_0^{\infty} N(M > M_{\text{lim}}, z)\text{d}z}. \tag{5}$$

In order to compute the limiting mass of the survey we will make a point-source approximation, that is, we will assume that each galaxy detected will be smaller than the telescope beam, and hence, all the flux it emits will contribute to a single beam. This is likely to be a good approximation for most objects in the survey, but will break
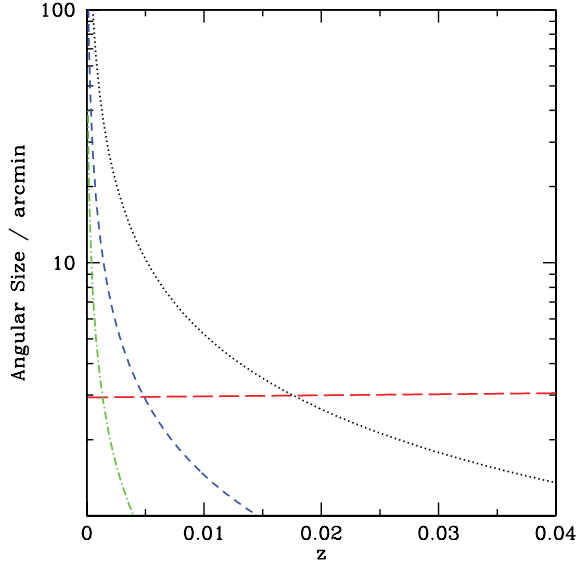
[5]SKA homepage: www.skatelescope.org/documents

Figure 1. The angular sizes (in arcminutes) for galaxies with H I masses $10^{10}\,M_\odot$ (dotted lines), $10^9\,M_\odot$ (dashed lines), and $10^8\,M_\odot$ (dot–dashed lines). The (approximately) horizontal long-dashed line represents the FWHM beamsize of FAST which scales as $\lambda \propto 1 + z$.

down for the closest, most massive objects. We expect those objects which violate this assumption to be the minority if we choose the survey strategy sensibly.

There is an empirically derived relation between the H I diameter, $D_{\rm HI}$, which is defined to be the region inside which the H I surface density is greater than $1\,M_\odot\,{\rm pc}^{-2}$, and the H I mass of a galaxy. This can be used to investigate the extent to which the point-source approximation is valid. We will use (Broeils & Rhee 1997; Verheijen & Sancisi 2001):

$$\frac{D_{\rm HI}}{\rm kpc} = \left(\frac{M_{\rm HI}}{10^{6.8}\,M_\odot}\right)^{0.55}, \tag{6}$$

which has been converted to an angular scale using the angular diameter distance $d_{\rm A}(z)$ for the range of relevant H I masses in Fig. 1. We note immediately that at $z \approx 0.02$ massive galaxies containing $10^{10}\,M_\odot$ of H I will be resolved by the FAST beam (see Fig. 1) but we postpone more detailed discussion of the validity of the point-source assumptions to later in Section 3.

The beam area increases like $\lambda^2 \propto (1 + z)^2$. As a result if an optimal survey strategy, in which one uniformly tiles the $z = 0$ slice, is chosen, then slices of the survey at higher redshift receive extra exposure due to the fact that these slices will overlap. The effect has been considered in ref. (Abdalla & Rawlings 2004) where it was shown that the flux limit relevant to a particular redshift slice is reduced by $(1 + z)^{-1}$. Hence, the flux limit for an observation, $S_{\rm lim}$, for a specific signal-to-noise ratio (S/N) is given by

$$S_{\rm lim} = ({\rm S/N})\frac{\sigma_{\rm noise}}{1 + z}. \tag{7}$$

A given object is detectable if its flux density in bandwidth $\Delta\nu = \nu\Delta V_{\rm o}/c$ is greater than $S_{\rm lim}$.

3.2 Simple estimates assuming a fixed observing bandwidth

A simple assumption, which should yield an order of magnitude estimate of the number of objects that one might expect to find in a survey, is that the linewidth of *all* the galaxies is $\Delta V_{\rm o} = 200\,{\rm km\,s}^{-1}$. Under this assumption, we have computed $M_{\rm lim}(z)$ and the number
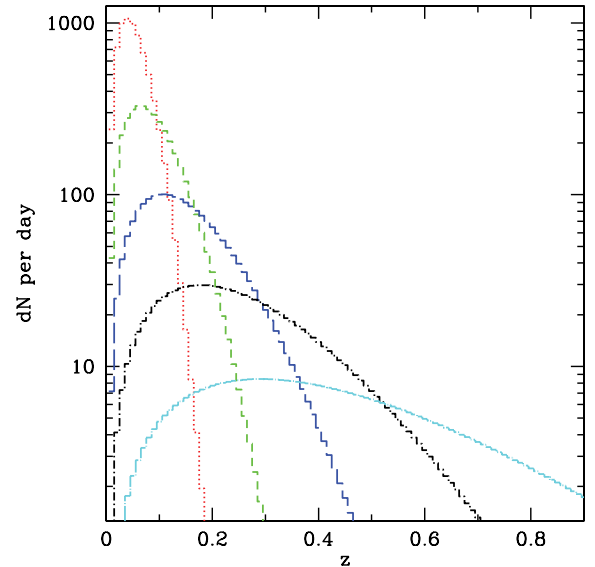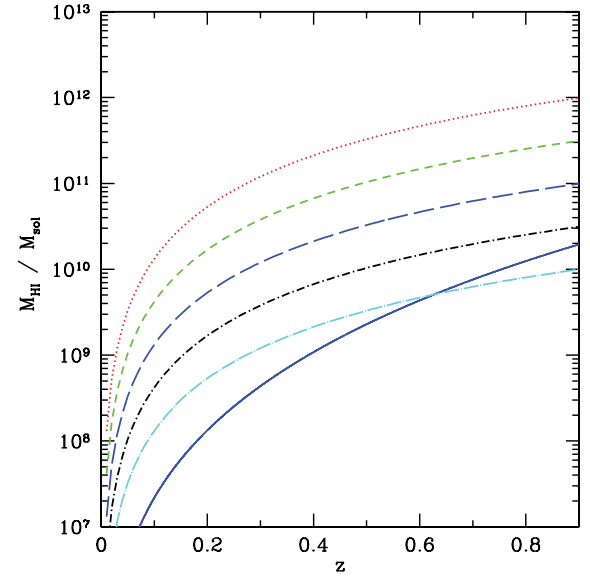

Figure 2. (a) the S/N = 4 limiting mass for different integration times (in seconds) per pointing: $t_{\rm obs} = 6$ (dotted line), 60 (short-dashed line), 600 (long-dashed line), 6000 (dot–short-dashed line), 60000 (dot–long-dashed line). (b) the predicted number counts in a day of observation (defined as 18 h of on-source integration time) for each of the observation strategies for $n_{\rm B} = 19$. The binwidth is $\Delta z = 0.01$. The total number of galaxies predicted in each case is summarized in Table 1. The solid line (i.e. the lowest) in panel (a) represents the estimated H I mass in a typical volume enclosed by the beam area and the fiducial velocity width $200\,{\rm km\,s}^{-1}$. If this line is greater than a substantive fraction (say, 1/3) of the mass limit, one would expect a confusion related increase in the observed noise.

of galaxies one would expect to find in an average day of observation as a function of $t_{\rm obs}$, the amount of time spent observing each beam area. Allowing for calibration and telescope maintainance, we will use an on-source integration time of 18 h to represent 1 d. The mass limit and number counts as a function of redshift (with bin width $\Delta z = 0.01$) are presented in Fig. 2 for S/N = 4 for $n_{\rm B} = 19$. The probability of a spurious detection will be $\sim 6 \times 10^{-5}$, assuming a Gaussian distribution for the noise. Since an all sky survey will contain $\sim 4 \times 10^9$ independent volume elements (the number

Table 1. Predicted number counts with S/N = 4 in one day of observation using a receiver system with $n_B = 19$ for a range of observation strategies. Counts expected for longer observations, or for larger numbers of beams, can be computed using a simple linear scaling. The limiting fluxes are computed using equation (7) with $z = 0$. All numbers have been rounded to two significant figures.

t_{obs}/s	S_{lim}/μJy	$N_{4\sigma}(n_B = 19, t = 1\,d)$
6	1700	7000
60	520	3800
600	170	2000
6000	50	1000
60 000	17	510

of independent beam areas multiplied by the number of indepndent depth slices) this would mean $\sim 2 \times 10^5$ spurious detections. In such a survey, therefore, there will be around 1 per cent spurious detections (we will find that around 10^7 galaxies can be detected). This will not significantly affect the measurement of the power spectrum discussed in the later sections of this paper.

One can immediately scale the number counts discussed in the previous paragraph to longer integration times and larger numbers of beams. If t_{survey} is the total time of the survey and the focal plane array has n_B beams, then the number of galaxies detected at a particular S/N is given by

$$N(n_B, t_{survey}) = \left(\frac{n_B}{19}\right)\left(\frac{t_{survey}}{1\,d}\right) N(n_B = 19, t = 1\,d), \qquad (8)$$

although one has to be careful to make sure that $t_{survey} < t_{sky}$, the time taken to cover the entire sky available to FAST, which is given in terms of t_{obs} by

$$\left(\frac{t_{sky}}{1\,d}\right) = 5.5 \left(\frac{19}{n_B}\right)\left(\frac{t_{obs}}{1\,s}\right), \qquad (9)$$

that is, the full $20\,000\,\deg^2$ will be covered in 5.5 d for a system with 19 beams if one spends 1 s per pointing. We remind the reader that we have assumed that 1 d corresponds to 18 h of on-source integration time.

Under the assumptions of this calculation, we deduce, for example, that a survey of the whole sky available to FAST with a noise level of $\sigma_{noise} \approx 40\,\mu Jy$ per beam is possible by performing 600 s integrations in around 9 yr if $n_B = 19$, but that this would take less than 2 yr with $n_B \sim 100$. Such a survey would find around 7×10^6 galaxies with S/N = 4 with an average redshift $\langle z \rangle \approx 0.15$. In subsequent sections, we will improve the validity of the calculation by modelling additional important effects. However, we shall see that this simple calculation holds true as a first order estimate. It is already clear that, if such a cosmologically important redshift survey is to be completed in an acceptable amount of time, a focal plane array with $n_B \sim 100$ is essential.

We can check a posteriori that the point-source approximation that we have made is valid by examining the limiting mass and expected number counts at very low redshift. We note first that the number counts tend to zero as z tends to zero since the volume element is $\propto z^2$. Thus, even in the case of $t_{obs} = 6\,s$, there are very few objects with $z < 0.02$, the point at which $M_{HI} = 10^{10}\,M_\odot$ galaxies are resolved by FAST. The percentage of low-redshift objects will decrease as t_{obs} increases. We estimate that the percentage of galaxies that would be significantly resolved for $t_{obs} = 6\,s$ is less than 10 per cent and for $t_{obs} > 60\,s$ the effect will be negligible.

In the most extreme circumstances, one might also be sensitive to the effects of confusion even in a spectroscopic survey since the FAST beamsize is likely to be much larger than the size of the galaxies we are trying to detect. One can attempt to quantify the expected amount of neutral hydrogen in a volume of the Universe enclosed by a Gaussian beam with area $\Omega_B = \pi\theta_{FWHM}^2/(4 \log 2)$ and the fiducial linewidth $\Delta V_o = 200\,km\,s^{-1}$. This is given by

$$\langle M_{HI}\rangle = \rho_{HI}(z)(1+z)\frac{\Delta V_o}{c}\Omega_B \frac{dV}{dzd\Omega}, \qquad (10)$$

where conservatively we assume that $\rho_{HI}(z) = \rho_{HI}(0)(1 + z)^3$. We take the local H I density to be $\Omega_{HI} = 2.6 \times 10^{-4}h^{-1}$ as found in the HIPASS survey (Zwaan et al. 2005), that is, $\rho_{HI}(0) \approx 7.2 \times 10^7\,h\,M_\odot\,Mpc^{-3}$. The computed value of $\langle M_{HI}\rangle(z)$ is compared to the mass limits for various values of t_{obs} in Fig. 2. It is clear from this that the effects of confusion are only significant for large values of t_{obs}; one might imagine that if $\langle M_{HI}\rangle > M_{lim}/3$, then confusion could lead to an increase in the effective noise making it more difficult to find galaxies.

From the above discussion it seems sensible to consider values of $60 < t_{obs} < 6000\,s$, in order to avoid any possible corrections to the statistics due to resolving galaxies at low z and confusion at high z.

3.3 Mass and inclination dependent effects in the linewidth

In this section we attempt to improve our calculation of the number counts by taking into account the fact that not all galaxies have the same observed linewidth ΔV_o. This can happen for two reasons: first the intrinsic linewidth is expected to depend on the mass for virialized systems, and secondly not all galaxies will be observed edge-on. If they are observed face-on, then it is the motion of H I perpendicular to the disc which sets the linewidth and not rotation.

The intrinsic linewidth of a galaxy, corrected for broadening, has been shown empirically to be related to the H I mass by (Briggs & Rao 1993; Lang et al. 2003):

$$\frac{\Delta V_e}{420\,km\,s^{-1}} = \left(\frac{M_{HI}}{10^{10}\,M_\odot}\right)^{0.3}, \qquad (11)$$

although we note that this relation shows a large dispersion, especially for dwarf galaxies.

The observed linewidth of a galaxy, ΔV_θ, which subtends an angle θ between its spin axis and the line-of-sight can be computed using the Tully–Fouque rotation scheme (Tully & Fouque 1985):

$$\begin{aligned}(\Delta V_e \sin(\theta))^2 &= (\Delta V_\theta)^2 + (\Delta V_t)^2 \\ &\quad - 2\Delta V_\theta \Delta V_t[1 - e^{-(\Delta V_\theta/\Delta V_c)^2}] \\ &\quad - 2(\Delta V_t)^2 e^{-(\Delta V_\theta/\Delta V_c)^2}.\end{aligned} \qquad (12)$$

$\Delta V_c = 120\,km\,s^{-1}$ represents an intermediate transition between the small galaxies with Gaussian H I profiles in which the velocity contributions add quadratically and giant galaxies with a 'boxy' profile reproduced by the linear addition of the velocity terms. $\Delta V_t \approx 20\,km\,s^{-1}$ is the velocity width due to random motions in the disc (Rhee & van Albada 1996; Verheijen & Sancisi 2001).

With this definition of θ, $\theta = 0$ corresponds to face-on and $\theta = \pi/2$ to edge-on. In cases where $\Delta V_\theta \gg \Delta V_c$, one can see that $\Delta V_\theta = \Delta V_t + \Delta V_e \sin\theta$. For $\theta = 0$, one finds that $\Delta V_\theta = \Delta V_t$, whereas for $\theta = \pi/2$, one finds that $\Delta V_\theta = \Delta V_t + \Delta V_e$ as expected.

In addition there is a broadening effect, ΔV_{inst}, of the H I profile due to the frequency resolution of the instrument, R. For a range of galaxy profiles, this broadening is found to be $\Delta V_{inst} \approx 0.55R$ (Bottinelli et al. 1990). An appropriate value for the present discussion is $\Delta V_{inst} \approx 16\,km\,s^{-1}$. It has been shown (Lang et al. 2003)

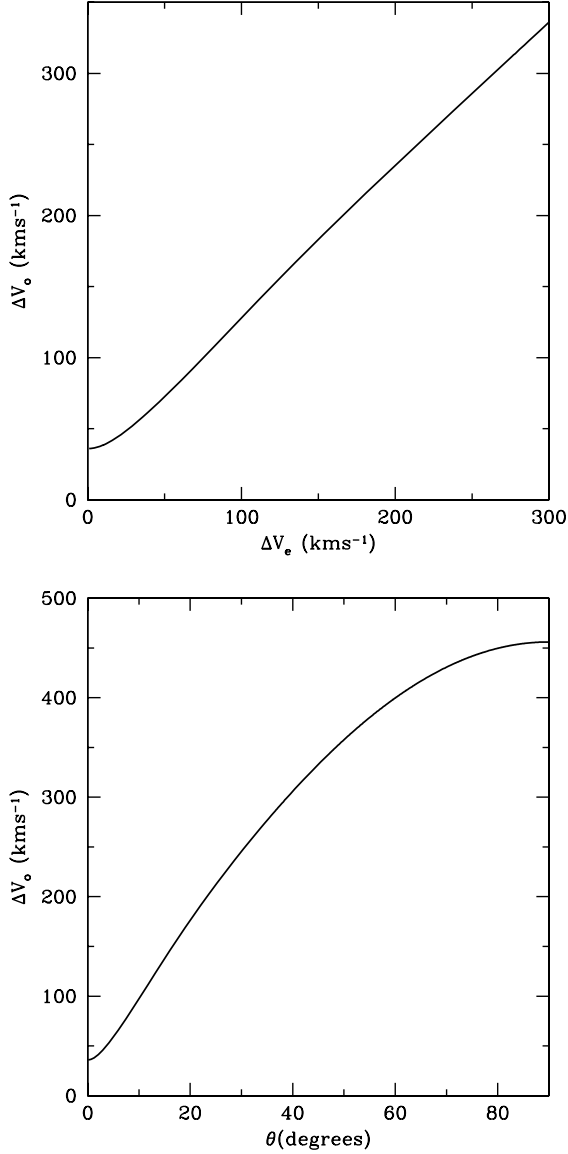**Figure 3.** The effective observed linewidth ΔV_o as a function of (a) ΔV_e for $\theta = \pi/2$, that is, edge on; (b) θ for $\Delta V_e = 420 \, \text{km s}^{-1}$.

that ΔV_{inst} should be added linearly to ΔV_θ to give the effective observed linewidth,

$$\Delta V_o(\theta) = \Delta V_\theta + \Delta V_{\text{inst}}. \tag{13}$$

ΔV_θ can be computed by using the Newton–Raphson method to find the root of equation (12) for a given $\Delta V_e(M_{\text{HI}})$. The effective observed linewidth is illustrated in Fig. 3 as a function of ΔV_e for an edge on galaxy and as function of θ for $\Delta V_e = 420 \, \text{km s}^{-1}$.

The limiting mass will now depend on θ as well as z, $M_{\text{lim}}(z, \theta)$, since it will be easier to find objects which have $\theta = 0$. In order to incorporate this we have take into account the probability of a given angle θ, $\mathcal{P}(\theta)$, when computing the number counts

$$N\,(M > M_{\text{lim}}) = \int_0^{\pi/2} \mathrm{d}\theta \, \mathcal{P}(\theta) \, \frac{\mathrm{d}N}{\mathrm{d}\theta}, \tag{14}$$

where

$$\frac{\mathrm{d}N}{\mathrm{d}\theta} = \Delta\Omega\Delta z \frac{\mathrm{d}V}{\mathrm{d}z\mathrm{d}\Omega} \int_{M_{\text{lim}}(z,\theta)}^{\infty} \frac{\mathrm{d}N}{\mathrm{d}V\mathrm{d}M} \mathrm{d}M. \tag{15}$$
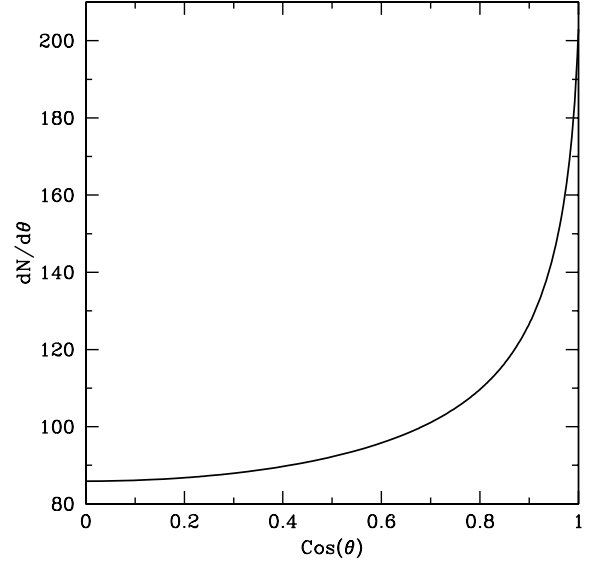

Figure 4. The effect of the angle of inclination on the detectability of galaxies. We have used $z = 0.1$, $\Delta z = 0.01$, $t_{\text{obs}} = 600 \, \text{s}$ and S/N $= 4$.

As an illustration, we have plotted $\mathrm{d}N/\mathrm{d}\theta$ against $\cos\theta$ at $z = 0.1$ with $\Delta z = 0.01$ for $t_{\text{obs}} = 600 \, \text{s}$ and S/N $= 4$ in Fig. 4 which shows that objects with θ close to zero are preferentially selected.

A simple assumption would be that the galaxies are randomly distributed on the celestial sphere and that the spin axes are also randomly distributed. This is not the case for satellite galaxies around their central galaxy, where a small anisotropy of the order of a few percent has been observed in SDSS data (Yang et al. 2006), but it is likely to be a good approximation for the FAST surveys discussed here which will preferentially select large, well-separated galaxies with an average separation of $\approx 5 \, \text{Mpc}$. In this case, $\mathcal{P}(\theta) = \sin\theta$ and therefore one can think of the value of $\cos\theta$ being uniformly distributed. Hence, one can deduce that

$$N\,(M > M_{\text{lim}}) = \Delta\Omega\Delta z \frac{\mathrm{d}V}{\mathrm{d}z\mathrm{d}\Omega} \int_0^1 \mathrm{d}(\cos\theta) \int_{M_{\text{lim}}(z,\theta)}^{\infty} \frac{\mathrm{d}N}{\mathrm{d}V\mathrm{d}M}. \tag{16}$$

The results of taking account these effects are presented in Table 2 and Fig. 5. It is clear that the number counts are slightly larger, by around 30 per cent, than under the assumption that $\Delta V_o = 200 \, \text{km s}^{-1}$. The scaling formulae discussed in Section 3.2 apply again here. For a survey with $n_B \approx 100$ taking around 2 yr with $t_{\text{obs}} = 600 \, \text{s}$, one would be able to cover all the sky available to FAST and find $\sim 10^7$ galaxies.

Table 2. The equivalent of Table 1 when the effects of mass and inclination on the observed linewidth are taken into account. The value of S_{lim} is dependent both on the mass and on the angle of inclination; in this case, we have therefore not attempted to include it. All numbers have been rounded to two significant figures.

$t_{\text{obs}}(s)$	$N_{4\sigma}(n_B = 19, t = 1 \, \text{d})$
6	9500
60	5200
600	2800
6000	1400
60 000	690

a realistic portrayal of the intricate physical processes acting on the H I mass function. It is designed solely as an illustrative example.

If the H I mass function is well-represented by a Schechter function, then the total density of H I is given by

$$\rho_{\text{H\,I}} = \Gamma(\alpha + 2)\theta^* M_{\text{H\,I}}^*. \quad (17)$$

If we now make the assumption that the faint-end slope α does not evolve with redshift, and that $\theta^* \neq \theta^*(z)$, we can model the effective change in the hydrogen mass due to passive star formation as a net shift of the break $M_{\text{H\,I}}^* = M_{\text{H\,I}}^*(z)$ with redshift. We assume that a simple form of this evolution

$$M_{\text{H\,I}}^*(z) = M_{\text{H\,I}}^*(0)(1 + \beta z), \quad (18)$$

where β is some constant of proportionality that determines the rate at which the hydrogen mass is consumed by star formation. This should be valid at low redshifts. In the subsequent discussion we will assume that $0 \leqslant \beta \leqslant 3$ in this model.

One can determine values compatible with present knowledge of the evolution of $\Omega_{\text{H\,I}}$. Under the assumption of equation (18) the total neutral hydrogen density of equation (17) is given by

$$\rho_{\text{H\,I}} = \Gamma(\alpha + 2)\theta^* M_{\text{H\,I}}^*(0)(1 + \beta z) = \rho_{\text{H\,I}}(z = 0)(1 + \beta z). \quad (19)$$

Using the cosmic density of neutral gas $\Omega_g(z)$ (measured relative to the present day critical density), one can estimate the value of β by assuming that the fraction of H I and He is constant between $z = 0$ and $z = 1$. By taking the ratio of measured neutral gas cosmic densities in the local Universe, $\Omega_g(z \approx 0) = 3.5 \times 10^{-4}$, and from damped Lyman α systems at higher-redshift (Prochaska, Herbert-Fort & Wolfe 2005; Rao & Turnshek 2000), $\Omega_g(z \approx 1) = 1.0 \times 10^{-3}$, we can estimate $\beta \sim 2$, which is also compatible with other works (Lah et al. 2007).

The effects of including this mass evolution are shown in Fig. 6 for the 100-beam FAST survey with $t_{\text{obs}} = 600$ s. It is clear that evolution leads to an increase in the number of galaxies found per day which are 14 420, 19 070, 25 190 and 32 780 for $\beta = 0, 1, 2$ and 3, respectively, and the median redshift increases with β. Essentially,

Figure 5. The equivalent of Fig. 2 for the case of FAST with 100 beams when the effects of mass and inclination on the observed linewidth are taken into account.

3.4 Effects of evolution

The calculations of the previous section are based on the assumption that the H I mass function does not evolve. In this section we discuss the possible effects of evolution on the number counts. It should be noted that none of these evolving models are used in the cosmological parameter constraints in Section 4. The basic picture which we will put forward is that the amount of H I is likely to increase as we go back in cosmic time and that the estimate of number counts based on the H I mass function measured at $z = 0$ is likely to be a lower limit of the number counts found by a particular survey. We will make a simple assumption that the relevant galaxies are formed by the epochs probed by FAST, and that H I is being used as the fuel for star formation, and hence is reducing in each of the galaxies as a function of time. We note that this passive star formation assumption ignores the possibility of mergers. We further note that this is only a toy model and is of course far too simplistic to be considered
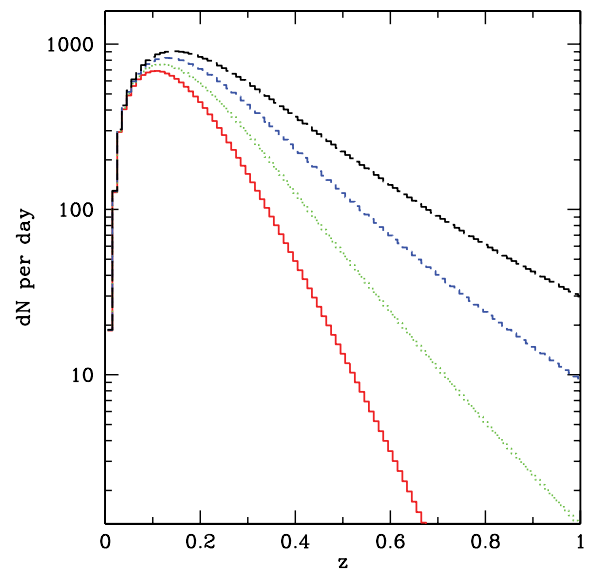

Figure 6. The predicted daily number counts with S/N = 4, $\Delta z = 0.01$ binwidth, for FAST operating with $n_B = 100$ and $t_{\text{obs}} = 600$ s as a function of redshift with a simple redshift-evolving mass break in the Schechter function of equation (4). This is shown as a factor of $\beta = [0, 1, 2, 3]$ for the solid lines, dotted lines, dashed lines, and dot–dashed lines.
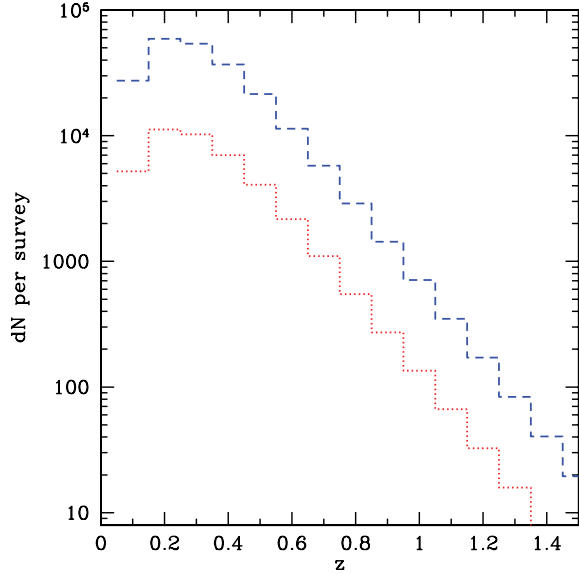
Figure 7. The number of galaxies that would be found in a survey lasting 30 d with $t_{obs} = 6000$ s. In the case of $n_B = 19$, one expects to find around 42 000 galaxies with $\langle z \rangle \approx 0.3$ in 17 deg^2, whereas one might find as many as 222 000 in 87 deg^2 for $n_B = 100$. We have used $\Delta z = 0.1$ and the $z = 0$ H I mass function in both cases.

we have shown that one can take the current prediction of $\sim 10^7$ galaxies found by the 100-beam FAST in two years of observation is a conservative lower bound. If we take seriously the value of $\beta = 2$, then one would expect to find nearly twice as many galaxies as in the case of no evolution.

Prior to embarking on the full, all-sky galaxy redshift survey, it would be sensible to perform an exploratory, deeper survey of a much smaller area, with the objective of measuring the evolution of the H I mass function. As well as being of legitimate scientific interest in its own right such a survey would allow one to investigate the optimal depth of the main redshift survey and pin-down some of the questions raised in this section. In Fig. 7, we show the redshift distribution of galaxies that one would expect for an H I survey with $t_{obs} = 6000$ s lasting 30 d using the $z = 0$ mass function and bins$> \Delta z = 0.1$. One would expected to find around 42 000 galaxies with $\langle z \rangle \approx 0.3$ using $n_B = 19$ and 222 000 for $n_B = 100$. In order to estimate accurately the parameters of the mass function, we estimate that one would require ~ 1000 galaxies per bin. Therefore, we believe that it should be possible to determine the H I mass function out to $z \approx 0.6$ for $n_B = 19$ and to $z \approx 1$ $n_B = 100$ with such a month-long survey.

4 EXPECTED ERRORS ON THE MATTER POWER SPECTRUM

We now use the predicted galaxy number counts, assuming no evolution of the kind described in the last section, to estimate the errors of the galaxy power spectrum at $z = \langle z \rangle \approx 0.15$. $P(k, z)$ is related to the power spectrum $P(k, 0)$ by

$$P(k, z) = [D(z)]^2 P(k),\qquad(20)$$

where $D(z)$ is the growth factor computed from

$$D(z) = \frac{5\Omega_m}{2} E(z) \int_z^\infty \frac{(1+z')\mathrm{d}z'}{[E(z')]^3},\qquad(21)$$

and $E(z) = H(z)/H_0$.

Errors on the power spectrum are due to two factors: sample variance, that is, the fact that not all k modes are measured, and shot-noise which is the effective noise on the measurement of an individual mode. The total error σ_P on the measurement of the power spectrum, $P(k, z)$, for a given k with logarithmic bin width $\Delta(\log k)$ can be expressed as (Feldman, Kaiser & Peacock 1994; Tegmark 1997):

$$\left(\frac{\sigma_P}{P}\right)^2 = 2\frac{1}{4\pi k^3 \Delta(\log k)}\frac{(2\pi)^3}{V_{eff}(k)}\left(\frac{1+nP}{nP}\right)^2,\qquad(22)$$

where $P = P(k, z)$ and $n = n(z)$ is the number density of galaxies which are detected (making nP dimensionless):

$$n(z) = \int_{M_{lim}(z)}^\infty \frac{\mathrm{d}N}{\mathrm{d}V\mathrm{d}M}\,\mathrm{d}M,\qquad(23)$$

and $V_{eff}(k)$ is the effective survey volume probed for a particular k mode:

$$V_{eff}(k) = \Delta\Omega \int_0^\infty \left(\frac{nP}{1+nP}\right)^2 \frac{\mathrm{d}V}{\mathrm{d}z\mathrm{d}\Omega}(z)\mathrm{d}z.\qquad(24)$$

For definiteness, we will consider a survey taking two years performed by the FAST telescope with a focal plane array with $n_B = 100$ and $t_{obs} = 600$ s per instantaneous FoV, $\Omega_{inst} \approx 1000$ arcmin^2. The choice of $t_{obs} = 600$ s is simply the required survey time for scanning the full sky in 2 yr and hence is the most efficient use of the telescopes resources. The justification for such an observation strategy is the simple fact that the area covered is inversely linear in integration time per field while the depth attained goes as the square root of the integration time; the most cosmic volume is attained from full sky surveys. We have also calculated the effect of probing different baryonic acoustic modes in the sky with varying observation strategies and determined that the simple maxim of 'wide-shallow' is always preferable to 'narrow-deep' in large-scale galaxy surveys. Using the methods of Section 3.3 we have estimated that such a survey will find $\sim 10^7$ galaxies with S/N = 4, the vast majority of which are spatially unresolved.

The expected errors on the power spectrum are presented in Fig. 8 for $\Delta(\log k) = 0.04$, which should be sufficiently large to ensure the errors are uncorrelated, according to a simple prescription detailed in Meiskin, White & Peacock (1999). The errors increase at low k due to sample variance and would also increase at high k due to shot noise were it not for our choice of a logarithmic binning; that is, the bins do not contain an equal number of k-modes. We have also divided the power spectrum by a model with no baryons which illustrates the acoustic features in the spectrum. The average redshift $\langle z \rangle \approx 0.15$ will prevent much cosmological information being gleaned from the acoustic scale and therefore we will not use it explicitly in our estimates of the cosmological parameters. It is clear, however, from Fig. 8 that the detection of the baryonic features would be possible with these data.

It is now interesting to compare the characteristics of the proposed FAST survey with those presently available. The effective volume of the proposed FAST survey evaluated at the $k = k_A = 0.075 h$ Mpc^{-1} scale of the acoustic peak in the power spectrum, is $V_{eff} = 0.82 \times 10^9 h^{-3}$ Mpc^3. This is shown in Fig. 9 where the effective volume probed, as well as the galaxy number count, of a variety of optical survey schemes are given, namely the SDSS main and SDSS LRG results as well as 2dFGRS, in comparison with FAST $n_B = 19$ for both $t_{obs} = 120$ s and $t_{obs} = 600$ s survey modes as well as our proposed version with $n_B = 100$ and $t_{obs} = 600$ s.
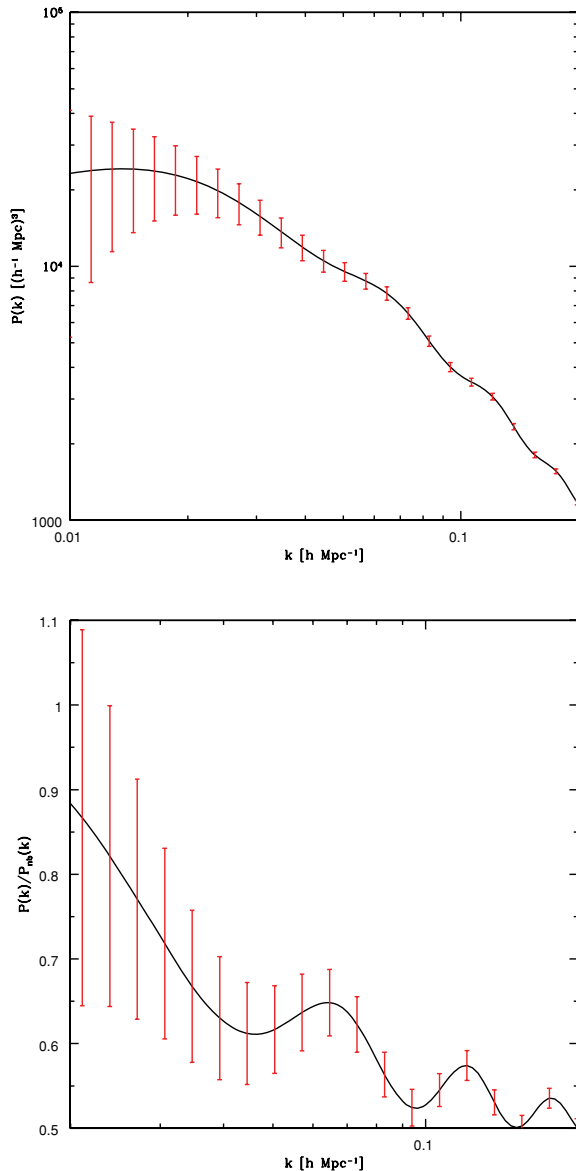
Figure 9. The effective V_{eff} at $k = 0.075h\,\text{Mpc}^{-1}$ for various completed and proposed surveys plotted against the approximate number of galaxies they have or might find. The volumes are quoted in terms of that of the SDSS main galaxy sample which has $V_{\text{SDSS}} = 1.5 \times 10^8 h^{-3}\,\text{Mpc}^3$. The open square is for the SDSS main sample, the cross is for the SDSS LRG sample, the hexagon is for the 2dFGRS sample, the pentagon is FAST with $n_{\text{B}} = 19$ with $t_{\text{obs}} = 600\,\text{s}$, while the circle is with $t_{\text{obs}} = 120\,\text{s}$ and the open triangle is for the proposed FAST survey taking 2 yr with $n_{\text{B}} = 100$.

Figure 8. (a) The expected errors on the power spectrum for the proposed FAST survey; (b) the same errors, but on a scale where the power spectrum has been divided by the equivalent no-baryon model.

5 ERROR FORECASTS FOR COSMOLOGICAL PARAMETERS

In the previous section we have shown that a survey using the FAST with $n_{\text{B}} = 100$ would find $\approx 10^7$ galaxies. It is interesting to determine the improvement that this would bring to cosmological parameter estimation through the accurate sampling of the galactic power spectrum relative to current large-scale optical surveys; 2dFGRS and SDSS. Since the FAST survey will have similar redshift coverage, but in a different waveband, this would bring more than just an improvement in the S/N.

For the 2dFGRS survey we use the data presented in Cole et al. (2005). We select 32 bandpowers in the range $0.022 < k/h(\text{Mpc}^{-1}) < 0.147$, where non-linear effects are expected to be small in this region. Even so, in Cole et al. (2005) it was proposed
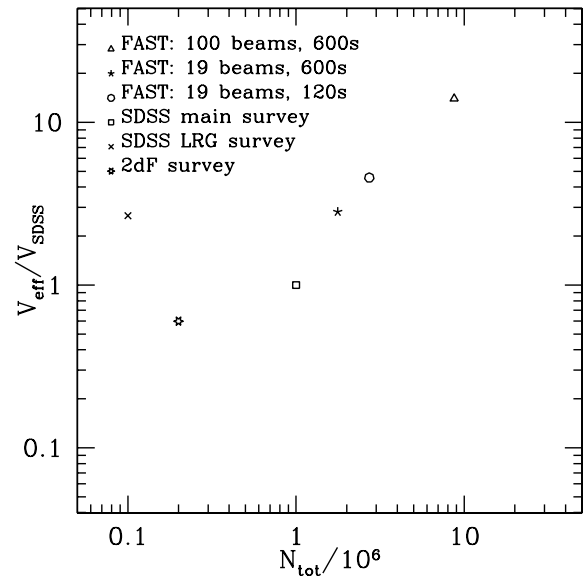
to use a fitting function of the form

$$P_{\text{obs}}(k) = b^2 \frac{1 + Qk^2}{1 + Ak} P_{\text{lin}}(k), \qquad (25)$$

to model the residual non-linear effects on the original, linear, power spectrum $P_{\text{lin}}(k)$. For the 2dFGRS data, the authors advocate using the parameters $A = 1.4\,\text{Mpc}^{-1}$ and $Q = 4\,\text{Mpc}^{-2}$. We follow this treatment, and analytically marginalize over the bias parameter b in the subsequent Markov-Chain Monte Carlo (MCMC) analysis. For the SDSS data, we use data presented in Tegmark et al. (2004a), using 19 bandpowers in the range $0.016 < k/(h\,\text{Mpc}^{-1}) < 0.205$. We take the power spectrum of galaxies to be of the form $P_{\text{obs}}(k) = b^2 P_{\text{lin}}(k)$, and analytically marginalize over the bias. For each of these datasets we use the window functions provided by Cole et al. (2005) and Tegmark et al. (2004a) respectively.

For the FAST data, we generated a fiducial power spectrum using the CAMB software (Lewis, Challinor & Lasenby 2000) using cosmological parameters that best-fit the ΛCDM model (Spergel et al. 2007), in this case $[\Omega_{\text{c}}h^2, \Omega_{\text{b}}h^2, h, n_{\text{s}}, A_{\text{s}}]$ are given by [0.104, 0.0223, 0.734, 0.951, 2.02×10^{-9}], where A_{s} is the initial amplitude of fluctuations. We then sample $P_{\text{lin}}(k)$ at closely spaced values of k/h, from which we interpolate to obtain $P_{\text{lin}}(k)$ at the appropriate FAST survey k/h values, along with the simulated error bars.

For completeness we have investigated three FAST survey schemes, namely the fiducial 19-beam FAST, with integration times per pixel of $t_{\text{obs}} = 120\,\text{s}$ and $600\,\text{s}$, and the 100-beam system with 600 seconds of integration per pixel each of which would take ~ 2 years to complete. For each scheme we use 28 bandpowers in the range $0.005 < k/(h\,\text{Mpc}^{-1}) < 0.155$, and take the galaxy power spectrum to have the form $P_{\text{obs}}(k) = b^2 P_{\text{lin}}(k)$. We also assume that the errors on each of the bins are uncorrelated due to the large binwidths taken in k as mentioned in Section 4. The 19-beam survey with $t_{\text{obs}} = 120\,\text{s}$, which could find $\approx 3 \times 10^6$, galaxies would cover the

Table 3. Summary of parameter values which come from the SDSS main sample and 2dFGRS, and the simulated constraints expected from the proposed FAST surveys with : (i) $n_B = 100$, $t_{obs} = 600$ s; (ii) $n_B = 19$, $t_{obs} = 120$ s; and (iii) $n_B = 19$, $t_{obs} = 600$ s. Values presented are the fractional errors for each parameter.

Parameter	SDSS	2dFGRS	FAST (19/120)	FAST (19/600)	FAST (100/600)
$\delta\Omega_m/\Omega_m$	0.416	0.221	0.230	0.310	0.078
$\delta\Gamma/\Gamma$	0.328	0.161	0.166	0.231	0.049
$\delta f_b/f_b$	0.378	0.253	0.468	0.486	0.283
$\delta n_s/n_s$	0.251	0.141	0.148	0.178	0.071

full 20 000 deg², albeit to a shallower depth than the 100-beam case, whereas the 19-beam $t_{obs} = 600$ s survey would only cover 4000 deg² and find $\approx 2 \times 10^6$ galaxies.

For the MCMC analysis, we use COSMOMC (Lewis & Bridle 2002) to create chains to estimate the confidence limits on the cosmological parameters. Since large parameter degeneracies exist when only including large-scale structure data in the fit, we impose the consistency relation that the angle $\theta_{acoustic}$ subtended by the first acoustic peak is 1.040, which is strongly constrained by the *WMAP3* data (Spergel et al. 2007). This essentially leaves three remaining free parameters in the fit: $\Omega_c h^2$; $\Omega_b h^2$ and n_s, which are used to compute the parameter $\Gamma = \Omega_m h$ and $f_b = \Omega_b/\Omega_m$. This is because we marginalize over bias, which absorbs the initial amplitude of fluctuations, and h can then be derived from these parameters and $\theta_{acoustic}$.

In Table 3, we present the fractional error on each of the marginalized cosmological parameters for the 2dFGRS and SDSS datasets, along with the three FAST observing schemes. The errorbars on f_b are of a similarly poor level to those possible with 2dFGRS and SDSS. This is probably since dependence of the power spectrum on f_b is weak and the present surveys have reached the ceiling on how well this can be measured from the matter power spectrum. However, the errors on Γ and n_s are significantly improved for the case of $n_B = 100$ and $t_{obs} = 600$ s, the area of the error ellipse in the $\Gamma - n_s$ direction is reduced by about a factor of 6 from that possible with 2dFGRS. It is clear, however, that the FAST surveys with $n_B = 19$ do not significantly improve on the constraints already available. Out of the two, that with $t_{obs} = 120$ s which covers the whole sky does much better than $t_{obs} = 600$ s making it clear, yet again, that surveys which find the largest number of objects (and which cover all the available sky) will typically constrain cosmological parameters the best.

In Fig. 10, we present marginalized 2D likelihoods of Γ versus f_b. Here, we impose a prior of $n_s = 0.95 \pm 0.02$, as preferred by the *WMAP3* data (Spergel et al. 2007). The first thing that is apparent is that the SDSS and 2dFGRS do not agree on the central value; something which has been noted in the literature and is thought to be associated with the different selection criteria for the two surveys (Cole Sanchez & Wilkins 2006). What is also evident is that FAST with $n_B = 100$ performs considerably better than the two presently available surveys.

Low-redshift galaxy redshift surveys cannot constrain the properties of dark energy directly, but they can in combination with measurements of the cosmic microwave background (CMB). Essentially, the measurement of Γ by the redshift survey breaks the angular diameter degeneracy, allowing simultaneous measurements of h and the equation-of-state parameter of the dark energy $w = $
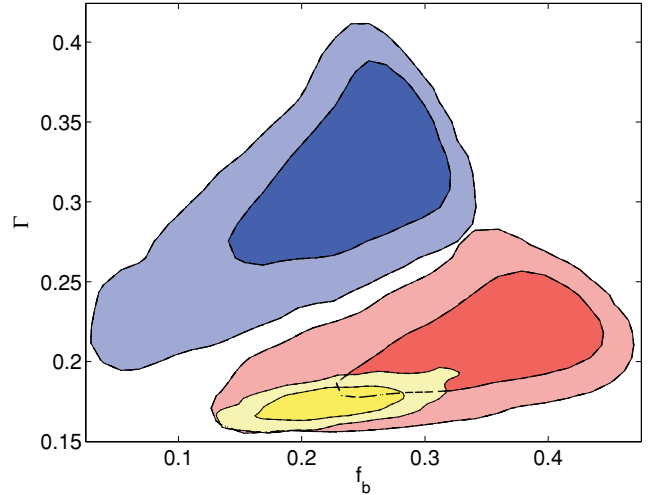

Figure 10. Cosmological parameter estimation for three surveys: the SDSS (top panel, blue), 2dFGRS (middle panel, red) and the proposed 100-beam FAST (bottom panel, yellow). Contours are 1 and 2σ, respectively, representing the 68 and 95 per cent probability estimates. Due to the tension between the 2dFGRS and SDSS central values, the fiducial FAST cosmology was based on *WMAP3*. Note the large improvement in constraining this parameter space, thanks to the order of magnitude increase in the number counts as well as in the volume probed.

Table 4. Expected errors for a seven-parameter fit using *Planck* alone and *Planck* + FAST 100/600. There is a degeneracy between w and h for *Planck* only, which is broken by the inclusion of FAST. τ is the optical depth to reionization, assumed to be instantaneous, and A_s is the standard power spectrum amplitude.

Parameter	*Planck*	*Planck* + FAST 100/600
$\Omega_b h^2$	0.0223 ± 0.0002	0.0223 ± 0.0002
$\Omega_c h^2$	0.104 ± 0.002	0.104 ± 0.002
n_s	0.952 ± 0.005	0.952 ± 0.005
$\log(10^{10}A_s)$	3.01 ± 0.01	3.01 ± 0.01
h	0.751 ± 0.131	0.731 ± 0.015
τ	0.090 ± 0.006	0.090 ± 0.0006
w	-1.02 ± 0.328	-0.99 ± 0.05

P/ρ. In order to see how FAST could impact on this we have performed a joint analysis using a simulated *Planck* data set using a noise power spectrum for the temperature anisotropies out to $\ell_{max} = 2000$ of $N_\ell^{TT} = 9 \times 10^{-5}$ μK², and that for the E-mode polarization given by $N_\ell^{EE} = 4N_\ell^{TT}$. We have assumed full sky observations and that optimal uncorrelated errors can be reconstructed from the data. The results are presented in Table 4 which show that w can be measured to within 5 per cent of -1, and h to within ≈ 2 per cent.

6 COMPARISON WITH EARLY-PHASE SKA

An early phase of the SKA with, say, 10 per cent of the total collecting area (i.e. an effective collecting area of $\sim 70\,000$ m² and thus ~ 40 per cent larger than FAST) and an FoV of at least $\Omega_{inst} = 1$ deg² (that is more than three times larger than the 100-beam case) is likely to come into operation during the first five years of FAST operation. It is interesting to compare and contrast the likely performance of such an early-stage SKA relative to FAST. We will assume
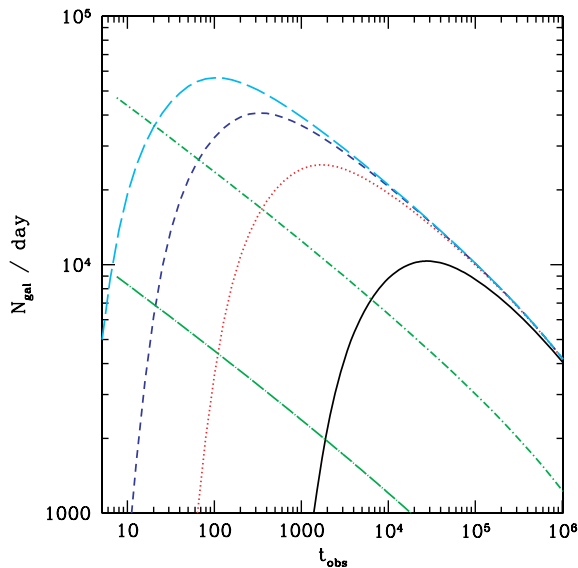
Figure 11. 10 per cent SKA number counts (6σ) for a variety of surveying strategies defined by t_{obs}, with a range of proposed angular resolutions. The resolutions considered here are 6 arcsec (solid lines), 12 arcsec (dotted lines), 18 arcsec (dashed lines) and 24 arcsec (long-dashed lines). The 19-beam (dotted–long-dashed lines) and 100-beam (dotted–short-dashed lines) surveys possible with FAST are presented for comparison. As explained in the text, the 10 per cent SKA will be significantly limited by the effect of resolving out galaxies, hence the trend seen here of better resolution demanding a survey scheme that favours deeper exposures to probe higher redshifts and observe more galaxies as point-like.

that, for H I observations, the value of T_{sys} for the 10 per cent SKA is the same as for FAST (we note that it is likely to be slightly higher, but not appreciably so; our results can be scaled by T_{sys}^{-2} where appropriate). The angular resolution will, however, be much higher since the baselines on the SKA 'core' will extend out to $\sim$5 km and hence the synthesized beam will be $\sim$10 arcsec. This introduces an extra complication into predicting the number counts since many galaxies may be resolved and hence use of the point-source approximation will overestimate the number of objects detected, particularly for shallow surveys.

In order to model this effect we have introduced an upper limit in the integral in equation (3) which is defined as the mass of object at a specific redshift which is resolved by the telescope beam. Of course, this then underestimates the number of objects found since some resolved objects are always found in any survey. However, the number which are missed by this procedure are likely to be small since the vast majority of detections are just above mass threshold.

The number of detections expected in 1 d are presented in Fig. 11 for a range of angular resolutions between 6 and 24 arcsec as a function of t_{obs}. Included also for comparison are predictions for FAST with 19 and 100 beams. The first thing to notice is that the FAST curves are almost straight lines which peak at the lowest possible value of t_{obs} confirming our earlier assertion that the best observing strategy for FAST, in terms of the number of objects detected, is to choose the value of t_{obs} so that one would cover the whole sky available to FAST in a defined survey time. The curves for the 10 per cent SKA have a very different shape, following a line similar to that for the FAST case for large values of t_{obs} before eventually turning over and plummeting to zero. This is the effect of the finite resolution, with the maximum being where the majority of objects found in a survey fill the beam exactly.

It is clear to see that for $\theta_{FWHM} > 15$ arcsec the 10 per cent SKA would make most of its detections in the point-source regime for $t_{obs} > 100$ s. A higher resolution is more likely and if we consider the case of 12 arcsec resolution, then the optimum value of $t_{obs} \approx$ 1000 s, but in terms of the number of objects found per day it is only marginally more powerful than for FAST with $n_B = 100$. For higher resolution still the optimal value of t_{obs} increases to around 2×10^4 s $\approx$ 6 h at 6 arcsec. In this case, one would be performing a much deeper survey which might capable of probing power spectrum at much higher redshifts and constraining the dark energy using the baryonic acoustic oscillations. It would certainly be very different in nature to that which we have proposed for FAST (Abdalla & Rawlings 2004).

To summarize, if the 10 per cent stage of the SKA has relatively low resolution and an FoV of $\sim$1 deg^2, then it will be a factor of a few times more powerful than FAST, in terms of the number of galaxies found at relatively low redshifts. However, if it has much higher angular resolution, then the survey which it will yield will be dominated by objects at much higher redshifts. If this were to be the case, the FAST and 10 per cent SKA would be complementary.

7 CONCLUSIONS

In this paper, we have shown that the upcoming FAST is an extremely capable instrument for large-scale neutral hydrogen surveys over the whole sky with the potential for large numbers of galaxies being discovered daily in the redshift range 0–0.2 and tens of galaxies out to mid-redshifts of the order of 0.2–0.4. This would provide complementary information to the redshift surveys already performed at optical wavelengths and would be considerably better than those presently available (or likely to be available in the near future) selected using neutral hydrogen as the tracer.

In calculating the number of galaxies expected, we have taken into account the effect of different angles of inclination for the galaxies to the observer and shown that an increase of $\sim$30 per cent is expected over simplistic approaches. Essentially the increase arises from the detection of face-on galaxies which are otherwise neglected. A first-order estimate has also been made of the effect of the evolution of the hydrogen mass of a galaxy by means of a toy model based on a linear increase in the mass of M^* with redshift. Although simplistic this nevertheless indicates that quite conservative increases in the neutral hydrogen content of a galaxy with redshift can lead to a factor of 2 increase in the expected number counts. This would lead to reductions in errors of the power spectrum due to shot noise, although we have not used this evolving H I mass function for the cosmological parameter constraints of Section 5. Prior to the main FAST survey, we have suggested a deep, month-long survey to investigate these evolution effects by estimating the H I mass function out to $z \approx 0.6$.

For the FAST H I galaxy survey to be a major improvement over current optical surveys it will be necessary to detect $\sim$10^7 galaxies over $\sim$20 000 deg^2, most of which will lie within redshifts of a few tenths. A uniformly sampled $\sim$10^7 H I galaxy survey will allow nearly an order of magnitude improvement of cosmological parameter constraints than currently provided by SDSS. For example the FAST survey will, independent of CMB observations, constrain n_s to 7 per cent and Γ to 5 per cent. Used in conjunction with observations of the CMB made by *Planck*, this could reduce the volume of the overall cosmological parameter space.

We have shown, however, that to achieve this performance in a reasonable time ($\sim$2 yr) a receiving system providing of the order of 100 beams is necessary. Conventional horn-based systems with

physical limits to their packing density are restricted to ~19 beams at the FAST prime focus and hence the 100-beam requirement can only be met with close-packed phased arrays of small antenna elements; these would not take up much more area than the 19-beam horn system. The required phased array technology is actively being developed by several groups working on technology for the SKA and hence we can be confident that by the time FAST comes into operation in 2012–2013 the 100-beam receiver capability will be available.

We have briefly looked ahead to the first phase of the SKA and compared the shallower FAST surveys with those from the much higher resolution interferometric surveys which will come somewhat later than that from FAST. The FAST and early-SKA surveys are complementary with the higher resolution SKA surveys preferentially finding higher redshift objects.

ACKNOWLEDGMENTS

REFERENCES

Abdalla F., Rawlings S., 2004, MNRAS, 360, 27
Blake C., Glazebrook K., 2003, ApJ, 594, 665
Bottinelli L., Gouguenheim L., Fouque P., Paturel G., 1990, A&AS, 82, 391
Briggs F. H., Rao S., 1993, ApJ, 417, 494
Broeils A., Rhee M.-H., 1997, A&A, 324, 877
Cole S. et al., 2005, MNRAS, 362, 505
Cole S., Sanchez A. G., Wilkins S., 2006, preprint (astro-ph/0611178)
Eisenstein D. et al., 2005, ApJ, 633, 560
Feldman H., Kaiser N., Peacock J., 1994, AJ, 426, 23
Lah P. et al., 2007, MNRAS, 376, 1357
Lang R. et al., 2003, MNRAS, 342, 738
Lewis A., Bridle S., 2002, Phys. Rev., D66, 103511
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Meiskin A., White M., Peacock J., 1999, MNRAS, 304, 851
Nan R., 2006, Sci. China G, 49, 129
Percival W. et al., 2001, MNRAS, 327, 1297
Prochaska J., Herbert-Fort S., Wolfe A., 2005, ApJ, 635, 123
Rao S. M., Turnshek D. A., 2000, ApJ, 130, 1
Rhee M.-H., van Albada T., 1996, A&AS, 115, 407
Roberts M. S., 1975, in Sandage M., Kristan J., eds, Galaxies and the Universe. University of Chicago Press, Chicago
Spergel D. et al., 2007, ApJS, 170, 377
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Tegmark M. et al., 2004a, ApJ, 606, 702
Tegmark M. et al., 2004b, Phys. Rev. D, 69, 103501
Tully R., Fouque P., 1985, ApJS, 58, 67
Verheijen M., Sancisi R., 2001, A&A, 370, 765
Yang X., van den Bosch F. C., Mo H. J., Mao S., Kang X., Weinmann S. M., Guo Y., Jing Y. P., 2006, MNRAS, 369
Zwaan M., Meyer M., Staveley-Smith L., Webster R., 2005, MNRAS, 359, L30

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