HI Fluctuations at Large Redshifts: II – the Signal Expected for the GMRT

Somnath Bharadwaj[1] & Sanjay K. Pandey[2]

[1]*Department of Physics and Meteorology & Center for Theoretical Studies,
I.I.T. Kharagpur 721 302, India.   e-mail: somnath@phy.iitkgp.ernet.in*
[2]*Department of Mathematics, L.B.S. College, Gonda 271 001, India.*
e-mail: spandey@iucaa.ernet.in

Abstract. For the GMRT, we calculate the expected signal from red-shifted HI emission at two frequency bands centered at 610 and 325 MHz. The study focuses on the visibility-visibility cross-correlations, proposed earlier as the optimal statistical estimator for detecting and analyzing this signal. These correlations directly probe the power spectrum of density fluctuations at the redshift where the radiation originated, and thereby provide a method for studying the large scale structures at large redshifts. We present detailed estimates of the correlations expected between the visibilities measured at different baselines and frequencies. Analytic fitting formulas representing the salient features of the expected signal are also provided. These will be useful in planning observations and deciding an optimal strategy for detecting this signal.

Key words. Cosmology: theory, observations, large scale structures—diffuse radiation.

1. Introduction

In two earlier papers (Bharadwaj, Nath & Sethi 2001; hereafter referred to as paper I and Bharadwaj & Sethi 2001; hereafter referred to as paper II) we have studied the possibility of detecting redshifted 21 cm emission from neutral hydrogen (HI) at high redshift sand using these observations to probe large scale structures. Our approach has been based on the fact that the HI density in the redshift range $1 \leq z \leq 3.5$ is known from observations of Lyman-α absorption lines seen in quasar spectra. These observations currently indicate $\Omega_{gas}(z)$, the comoving density of neutral gas expressed as a fraction of the present critical density, to be nearly constant at a value $\Omega_{gas}(z) \sim 10^{-3}$ for $z \geq 1$ (Peroux *et al.* 2001). The bulk of the neutral gas is in clouds which have HI column densities greater than 2×10^{20} atoms/cm^2 (Peroux *et al.* 2001, Storrie-Lombardi, McMahon & Irwin 1996; Lanzetta, Wolfe & Turnshek 1995). These high column density clouds are responsible for the damped Lyman-α absorption lines observed along lines of sight to quasars. The flux of HI emission from individual clouds ($< 10 \, \mu$Jy) is too weak to be detected by existing radio telescopes unless the image of the cloud is

significantly magnified by an intervening cluster gravitational lens (Saini, Bharadwaj & Sethi 2001). Although we may not be able to detect individual clouds, the redshifted HI emission from the distribution of clouds will appear as a background radiation in low frequency radio observations. In the two earlier papers, and in the present paper we investigate issues related to calculating the expected signal and detecting it.

The hope of detecting this signal lies in the fact that the distribution of the HI clouds is expected to be clustered. The clustering pattern of the HI clouds will be reflected in the angular and spectral distribution of the redshifted 21 cm emission. It is the aim of this paper, and the two previous papers, to characterize the expected properties of this signal and to propose an optimum statistical estimator to extract this signal from the various foregrounds and the system noise which will also be present in any low frequency observation. In paper II we proposed the cross-correlations between the visibilities measured at different baselines and frequencies as a statistical estimator to optimally detect the HI signal. This has the advantage that the system noise contribution to visibilities measured at different baselines and frequencies are uncorrelated. Further, in paper II we showed that the visibility-visibility cross-correlations directly probes the power spectrum of density fluctuations at the epoch where the HI emission originated. In this paper we present results of a detailed investigation of the signal expected for this statistical estimator for observations with the GMRT.

We next present a brief outline of this paper. In section 2 we review the formalism used to calculate the visibility-visibility cross-correlations. In section 3 we present the values of a set of useful parameters to do with the GMRT (Swarup *et al.* 1991), the cosmological model chosen for these investigations and the HI distribution. In section 4 we present the results of our investigations, and in section 5 we discuss the results and present conclusions.

2. Formalism

We consider a situation where an interferometric array of N antennas is used to carry out low frequency radio observations. The observations are carried out at NC frequency channels $\{\nu_1, \nu_2, \nu_3, ..., \nu_{NC}\}$ covering a frequency band B centered at the frequency ν_c. For the purpose of this paper we assume that the antennas are distributed on a plane, and the position of each antenna is denoted by a two dimensional vector $\mathbf{d}_i$. The antennas all point in the same direction along the unit vector $\mathbf{n}$ which we take to be vertically upwards. The beam pattern $A(\vec{\theta})$ quantifies how the individual antenna, pointing upwards, responds to signals from different directions in the sky. This is assumed to be a Gaussian $A(\vec{\theta}) = e^{-\theta^2/\theta_0^2}$ with $\theta_0 \ll 1$ i.e., the beam width of the antennas is small, and the part of the sky which contributes to the signal can be well approximated by a plane.

The quantity measured in interferometric observations is the visibility $V(\mathbf{U}, \nu)$ which is recorded for every independent pair of antennas (baseline) at every frequency channel in the band. For any pair of antennas, the visibility depends on the vector $\mathbf{d} = \mathbf{d}_i - \mathbf{d}_j$ joining the position of the two antennas. It is convenient to express the visibility as a function of the variable $\mathbf{U}$ which is $\mathbf{d}$ expressed in units of the wavelength i.e., $\mathbf{U} = \mathbf{d}/\lambda$. Typically, the visibility is recorded for $\mathbf{U}$ in a range $U_{\min} \leq | \mathbf{U} | \leq U_{\max}$. The visibility corresponding to $\mathbf{U} = 0$ is the flux measured by a single antenna, and this is usually not recorded in an interferometric array.

In paper II we calculated the HI contribution to the visibility $V(\mathbf{U}, \nu)$. This can be expressed as

$$V(\mathbf{U}, \nu) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \Delta^s_{\mathrm{HI}}(\mathbf{k}, \nu) e^{-ik_\parallel r_\nu} a_\nu \left(\mathbf{U} - \frac{\mathbf{k}_\perp r_\nu}{2\pi}\right). \tag{1}$$

The frequency ν appears in many of the terms on the right hand side of equation (1) to indicate that these are to be evaluated at the redshift $z = \nu_e/\nu - 1$ where the radiation originated. The first term $\bar{I}_\nu$ is the specific intensity of the redshifted HI emission expected if the HI were uniformly distributed. This is given to be (from equation (3) of paper II)

$$\bar{I}_\nu = \frac{2.7\,\mathrm{Jy}}{\mathrm{degree}^2}\, h\, \Omega_{\mathrm{gas}}(z)\, \left[\frac{H_0}{H(z)}\right], \tag{2}$$

where $H(z)$ is the Hubble parameter, and H_0 is its present value.

It should be noted that equation (14) of paper II is incorrect. Equation 3 of paper II is correct. The error came about when expressing $\bar{n}_{\mathrm{HI}}(z)$, the comoving number density of HI atoms in terms of Ω_{HI}. As a result of this and an error in the computer programs that were used, the predictions for 610 MHz in paper II are also incorrect.

The HI emission originates from the gas located at a distance r_ν in the direction $\mathbf{n}$. The term $\Delta^s_{\mathrm{HI}}(\mathbf{k}, \nu)$ in equation (1) is the Fourier transform of the density contrast of the HI distribution in redshift space. Any Fourier mode $\mathbf{k}$ can be decomposed into two parts, a component $k_\parallel$ parallel to $\mathbf{n}$ and a component $\mathbf{k}_\perp$ perpendicular $\mathbf{n}$. Assuming that the HI density fluctuations evolve according to linear theory, and that these may be related to the fluctuations in the underlying dark matter distribution through a linear bias parameter $b(z)$, we have

$$\Delta^s_{\mathrm{HI}}(\mathbf{k}, \nu) = b(z)D(z)\left[1 + \beta(z)\frac{k_\parallel^2}{k^2}\right]\Delta(\mathbf{k}), \tag{3}$$

where $D(z)$ is the growing mode of density perturbations (Peebles 1980), $\beta(z) = f(\Omega_m, \Omega\Lambda)/b(z)$ (Lahav *et al.* 1991) is the linear redshift distortion parameter, and $\Delta(\mathbf{k})$ is the Fourier transform of the density fluctuations of the dark matter distribution at the present epoch.

We also have

$$\langle\Delta(\mathbf{k})\Delta^*(\mathbf{k}')\rangle = (2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}')P(k), \tag{4}$$

where $\langle\rangle$ denotes ensemble average, and $P(k)$ is the power-spectrum of dark matter fluctuations at the present epoch.

The quantity $a_\nu(\mathbf{U})$ in equation (1) is the Fourier transform of $A(\theta)$, the beam pattern. For a Gaussian beam, the Fourier transform also is a Gaussian and we have

$$a_\nu \left(\mathbf{U} - \frac{\mathbf{k}_\perp r_\nu}{2\pi}\right) = \pi\theta_{0\nu}^2 \exp\left[-\pi^2\theta_{0\nu}^2\left(\mathbf{U} - \frac{\mathbf{k}_\perp r_\nu}{2\pi}\right)^2\right]. \tag{5}$$

This function is sharply peaked around $\mathbf{k}_\perp = 2\pi \mathbf{U}/r_\nu$, and only values of $\mathbf{k}_\perp$ in the interval $| \Delta\mathbf{k}_\perp | < 1/(r_\nu \theta_{0\nu})$ centered around this value will contribute to the integral in equation (1). It follows that the visibility $V(\mathbf{U}, \nu)$ picks up contributions only from those fluctuations $\Delta^s_{\mathrm{HI}}(\mathbf{k}, \nu)$ for which $\mathbf{k}_\perp \approx 2\pi \mathbf{U}/r_\nu$.

A point to be noted is that in the Gaussian model for the beam pattern its Fourier transform is also a Gaussian which is a non-compact function. In reality, the Fourier transform of the beam pattern is compact and it is zero outside a radius D/λ i.e., $a_\nu(\mathbf{U} - \frac{\mathbf{k}_\perp r_\nu}{2\pi}) = 0$ for $| \mathbf{U} - \frac{\mathbf{k}_\perp r_\nu}{2\pi} | > D/\lambda$. This fact plays a role in the later discussion and we shall refer back to it at the appropriate place.

In this paper we shall study in some detail the cross-correlation $\langle V(\mathbf{U}_1, \nu_1) V^*(\mathbf{U}_2, \nu_2) \rangle$ between the visibilities measured at two different baselines and frequencies. Using equations (1), (3), (4), this turns out to be

$$\langle V(\mathbf{U}_1, \nu_1) V^*(\mathbf{U}_2, \nu_2) \rangle = \left[\bar{I} D b\right]_1 \left[\bar{I} D b\right]_2 \int \frac{d^3 k}{(2\pi)^3} P(k) \left[1 + \beta_1 \frac{k_\parallel^2}{k^2}\right]$$

$$\times \left[1 + \beta_2 \frac{k_\parallel^2}{k^2}\right] \cos[k_\parallel(r_1 - r_2)] \, a_1 \left(\mathbf{U}_1 - \frac{\mathbf{k}_\perp r_1}{2\pi}\right) a_2^* \left(\mathbf{U}_2 - \frac{\mathbf{k}_\perp r_2}{2\pi}\right), \quad (6)$$

the imaginary part being zero. We shall use equation (6) to calculate the correlations between the visibilities at different baselines and frequencies. We next discuss a few approximations which can be used to simplify equation (6) and which help in interpreting this equation. We test the range of validity of these approximations and comment on this later in the paper.

Typically, the bandwidth B is small compared to the central frequency ν_c, and the quantities I_ν, $b(z)$, $D(z)$, $\beta(z)$, $\theta_{0\nu}$ and r_ν in equation (6) do not vary substantially across the band. We can evaluate these at ν_c instead of calculating them separately at ν_1 and ν_2, and at some instances we do not explicitly show the subscript ν_1 or ν_2 for these quantities. The term $\cos[k_\parallel(r_1 - r_2)]$ is an exception because $k_\parallel(r_1 - r_2)$ may be large even if $\Delta\nu = \nu 2 - \nu_1 \ll \nu_c$. This term may be approximated as $\cos[k_\parallel r_\nu' \Delta\nu]$, where $r_\nu' = dr_\nu/d\nu$.

The term $a_1(\mathbf{U}_1 - \frac{\mathbf{k}_\perp r}{2\pi}) a_2^*(\mathbf{U}_2 - \frac{\mathbf{k}_\perp r}{2\pi})$ in equation (6) is a product of two functions, one sharply peaked at $\mathbf{k}_\perp = 2\pi \mathbf{U}_1/r$ and another at $\mathbf{k}_\perp = 2\pi \mathbf{U}_2/r$. The two peaks have very little overlap if $| \mathbf{U}_1 - \mathbf{U}_2 | \gg 1/\theta_0$, and the visibilities measured at two such baselines will be uncorrelated. There will be correlations only between visibilities measured at different frequencies for nearby baselines, i.e., $| \mathbf{U}_1 - \mathbf{U}_2 | \leq 1/\theta_0$. For the situation where $\mathbf{U}_1 = \mathbf{U}_2 \gg 1/\theta_0$, we may use the approximation (paper II)

$$| a_1 \left(\mathbf{U}_1 - \frac{\mathbf{k}_\perp r}{2\pi}\right) |^2 \approx \left(\frac{2\pi^3 \theta_0^2}{r^2}\right) \delta^2 \left(\mathbf{k}_\perp - \frac{2\pi}{r} \mathbf{U}_1\right) \quad (7)$$

whereby equation (6) is considerably simplified and we have

$$\langle V(\mathbf{U}, \nu) V^*(\mathbf{U}, \nu + \Delta\nu) \rangle = \frac{[\bar{I} b D \theta_0]^2}{2r^2} \int_0^\infty dk_\parallel P(k) \left[1 + \beta \frac{k_\parallel^2}{k^2}\right]^2 \cos(k_\parallel r' \Delta\nu) \quad (8)$$

where $k = \sqrt{(2\pi U/r)^2 + k_\parallel^2}$.

Table 1.

Frequencies ν_c (MHz)	610 325
Possible bandwidths - B (MHz)	16 8 4 2 1 0.5
No. of channels - NC	128
Antenna separations - d (m)	$d_{min} \sim 60$ $d_{max} \sim 25 \times 10^3$

Table 2.

ν_c	U_{min}	U_{max}	$\theta_0 \approx 0.6 \times \theta_{FWHM}$
610	123	51×10^3	$0.54°$
325	65	27×10^3	$1.08°$

Equation (8) provides considerable insight on the visibility-visibility cross-correlation in the situation where $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$. The point to note that the visibility signal measured at a baseline $\mathbf{U}$ receives contributions only from the fluctuations $\Delta(\mathbf{k})$ for which $\mathbf{k}_\perp = (2\pi \mathbf{U})/r$, i.e., $\mathbf{k}_\perp$ is fixed but $k_\parallel$ can have any value. As a consequence the correlation $\langle V(\mathbf{U}, \nu) V^*(\mathbf{U}, \nu + \Delta\nu) \rangle$ directly probes the power spectrum $P(k)$ at all Fourier modes $k > k_{min} = 2\pi U/r$.

In the next sections of this paper we use the formulas presented here to investigate the HI signal expected for the GMRT and study how this is related to the large-scale structures at the redshift where the HI emission originated. We also compare the predictions of equation (6) with those of (8) and test the range of validity of the approximations discussed above.

3. Some useful parameters

3.1 *GMRT*

The Giant Meterwave Radio Telescope (GMRT) has 30 radio antennas of 45 m diameter each. We present below, in tabular form (Tables 1 and 2), some of the parameters relevant for the proposed HI observations. We have restricted our analysis to only two of the GMRT frequencies which correspond to HI in the range $1 \leq z \leq 3.5$.

3.2 *Cosmological model*

For the background cosmological model we have used $H_0 = 100\,h$ km/s/Mpc with $h = 0.7$, $\Omega_{m0} = 0.3$, $\Omega_{\Lambda 0} = 0.7$, $\Omega_{Baryon\,0} = 0.015\,h^{-2}$ and $\Omega_{gas} = 1. \times 10^{-3}$. The power spectrum of dark matter density fluctuations is normalized to COBE (Bunn & White 1996), and its shape is determined using the analytic fitting form for the CDM power spectrum given by Efstathiou, Bond & White (1992). The value of the shape parameter turns out to be $\Gamma = 0.2$ for the set of cosmological parameters used here. The power spectrum is shown in Fig. 1.

Very little is known about the bias b which relates the HI fluctuations to the dark matter fluctuations, and we have used a fixed value $b = 1$ throughout.
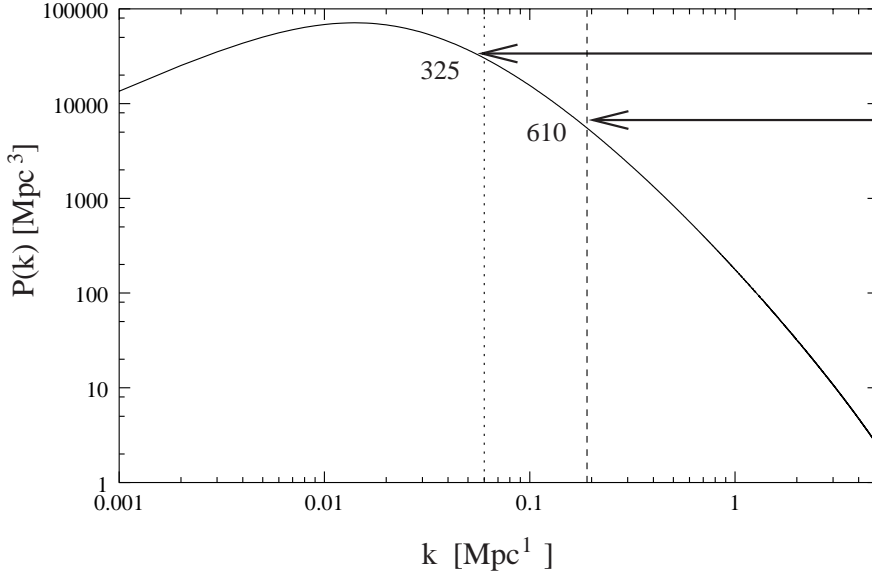
Figure 1. This shows the part of the power spectrum probed by the smallest baselines in GMRT for the two frequencies indicated in the figure. The power spectrum is normalised to COBE and is shown for $z = 0$.

Table 3.

ν_c	z	r (Mpc)	$\bar{I}$ (Jy/deg^2)	$r^{'}$ (Mpc/MHz)	k_{min} (Mpc^{-1})
610	1.33	4030	9.0×10^{-4}	7.7	0.16
325	3.37	6686	3.7×10^{-4}	11.3	0.09

3.3 *HI*

We first present the values of the redshift, comoving distance and the expected specific intensity of the HI distribution for the two GMRT frequencies of our interest (Table 3). We also show values of $r^{'} = dr/d\nu$ which are useful in calculating the comoving separation corresponding to a given frequency interval i.e., $\Delta r = r^{'} \Delta \nu$. The last column in Table 3 shows the value of k_{min} for the baseline $U = 100$. This represents the smallest Fourier mode probed by this baseline.

In Fig. 1 we show the part of the power spectrum which will be probed by U_{min} the smallest baseline available on GMRT. The value of k_{min} will be higher for the larger baselines and these will probe a smaller part of the power spectrum.

4. Results

4.1 *Single dish observations*

It is possible to use the GMRT as a set of 30 single dish antennas instead of an interferometric array. In such observations the flux incident on each antenna is recorded
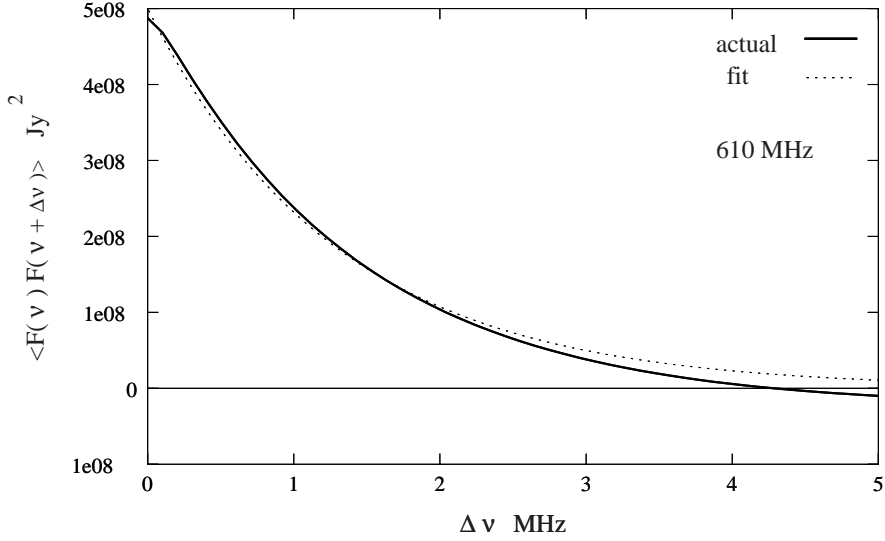
Figure 2. This shows the contribution from redshifted HI emission to the correlation between the flux expected at two different frequencies in the 610 MHz band for single dish observations using GMRT. The analytic fit (equation 9) is also shown.
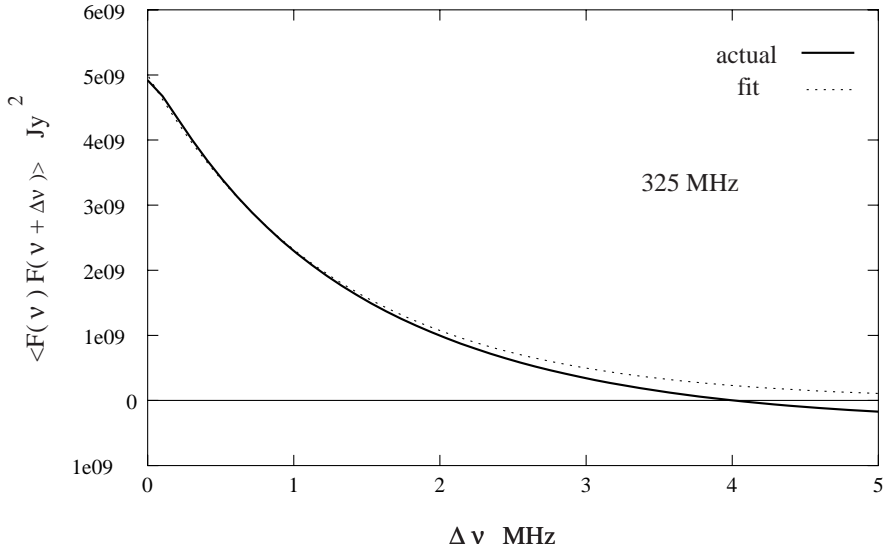
Figure 3. This shows the contribution from redshifted HI emission to the correlation between the flux expected at two different frequencies in the 325 MHz band for single dish observations using GMRT. The analytic fit (equation 9) is also shown.

and the signal from the 30 antennas is combined. The flux $F(\nu)$ is the visibility $V(\mathbf{U}, \nu)$ at $\mathbf{U} = 0$, and the flux expected from the redshifted HI emission can be calculated using equation (1). We use equation (6) to calculate the correlation between the flux measured at different frequencies $\langle F(\nu)F(\nu + \Delta\nu)\rangle$. The results for the two GMRT bands of interest are shown in Figs. 2 and 3. The approximate formula (equation 8) is valid only when $|\mathbf{U}| \gg 1/\theta_0$, and this cannot be used here.

We find that the HI flux at neighbouring frequencies are correlated. The correlation signal in the 610 MHz band is nearly an order of magnitude higher than the correlation signal in the 325 MHz band. In both the bands the correlation falls exponentially as $\Delta\nu$ is increased. For $\Delta\nu \leq 2\,\text{MHz}$, the correlation can be very well approximated by the fitting formula

$$\langle F(\nu)F(\nu + \Delta\nu)\rangle = C_\nu \exp\left(\frac{-\Delta\nu}{1.3\text{MHz}}\right) \qquad \left\{ \begin{array}{l} C_{610} = 5 \times 10^{-8}\text{Jy}^2 \\ C_{325} = 5 \times 10^{-9}\text{Jy}^2. \end{array} \right. \qquad (9)$$

The fitting formula (equation 9) fails at larger $\Delta\nu$ where $\langle F(\nu)F(\nu + \Delta\nu)\rangle$ becomes negative indicating that the fluxes have a weak anti-correlation for $\Delta\nu > 4\,\text{MHz}$.

4.2 *Interferometric observations*

We next consider interferometric observations with GMRT. The correlation $\langle V(\mathbf{U}_1, \nu_1)V^*(\mathbf{U}_2, \nu_2)\rangle$ is maximum when $\mathbf{U}_1 = \mathbf{U}_2$, and we consider this situation first. In addition, it is a good approximation to represent the correlation as a function of $\Delta\nu = \nu_2 - \nu_1$, provided ν_1 and ν_2 are in the same GMRT band, Our results are shown in Figs. 4 and 5.
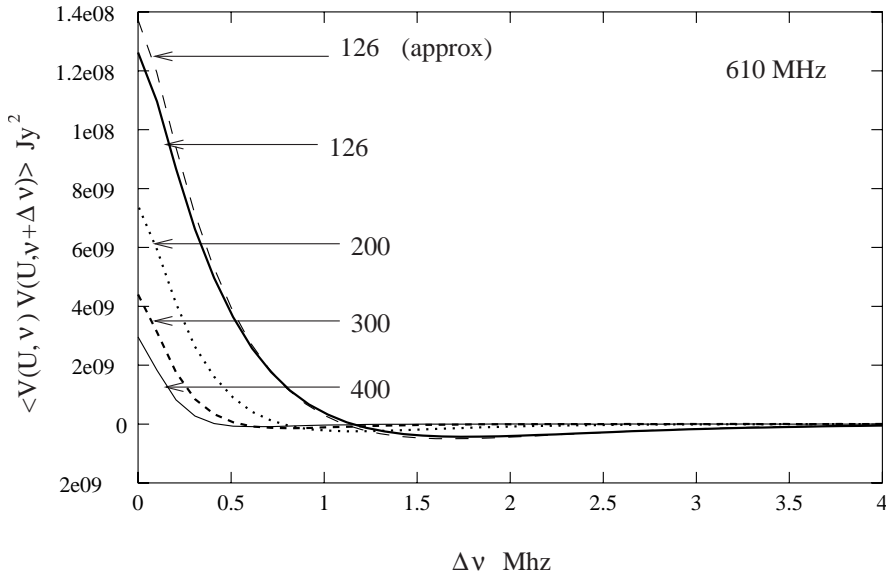
Figure 4. This shows the correlation between the visibilities $V(\mathbf{U}, \nu)$ and $V(\mathbf{U}, \nu+\Delta\nu)$ expected for the same baseline $\mathbf{U}$ at two different frequencies. The results are shown for different values of U (shown in the figure) starting from $U_{\min}$ for the 610MHz band. The results have been calculated using equation (6), the predictions of the approximate formula (equation (8) are also shown for $U_{\min}$.

We have calculated the correlations using two different formulas i.e., equations (6) and (8). The latter can only be used when $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{U}$, and it is based on an approximation which is valid only when $\mathbf{U} \gg 1/\theta_0$. We find that the approximate
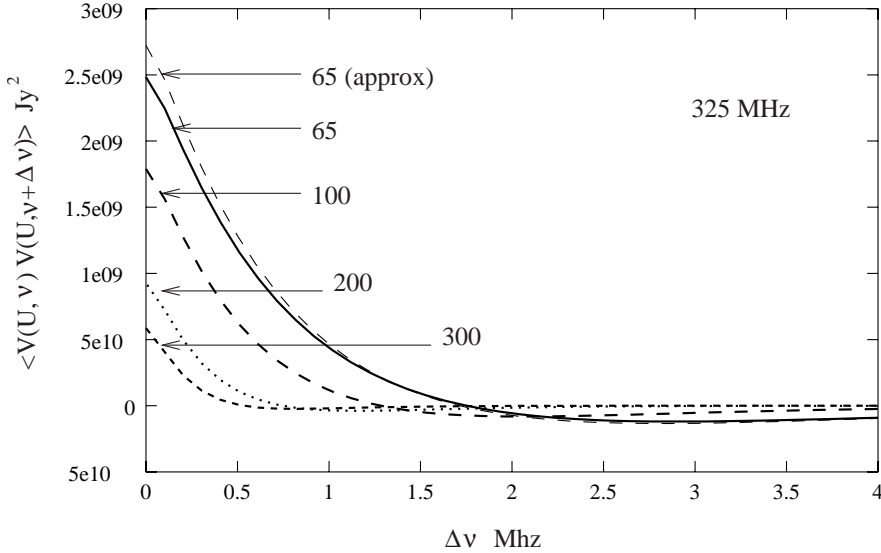
Figure 5. This shows the same thing as figure 4 but for the 325MHz band.

formula (equation 8) matches equation (6) quite well ($\sim 10\%$) at the smallest baseline $U_{\min}$ for both 610 MHz and 325 MHz (Figs. 4 and 5). The agreement is better than 10% for the larger baselines.

The correlation signal is the strongest at the smallest baseline $U_{\min}$ and falls quickly for the larger values of U. The larger baselines receive contributions from a smaller part of the power spectrum (Fig. 1) and the fall-off with U reflects the shape of the power spectrum at the mode $k \sim (2\pi U/r)$. At a fixed U, the visibilities at adjacent frequencies are correlated, the correlation falls approximately exponentially with increasing $\Delta\nu$. At the larger baselines the correlation decreases faster with increasing $\Delta\nu$. At all baselines, the correlation crosses zero and there are anti-correlations for large $\Delta\nu$. We find that in the range of U and $\Delta\nu$ where the visibility-visibility correlation signal is large, it can be approximated by the fitting formula

$$\langle V(U, \nu)V^*(U, \nu + \Delta\nu)\rangle = C_\nu \exp\left(\frac{-\Delta\nu\, u^{0.8}}{0.7}\right) \frac{\sin\left(2\,\Delta\nu\, u^{1.2}\right)}{2\,\Delta\nu\, u^{1.2}}, \qquad (10)$$

$$C_{610} = 1.6 \times 10^{-8} \mathrm{Jy}^2 u^{-1.2} \qquad\qquad C_{325} = 1.8 \times 10^{-9} \mathrm{Jy}^2 u^{-0.9}$$

where $u = (U/100)$ and $\Delta\nu$ is in MHz.

We next calculate the correlation $\langle V(\mathbf{U}, \nu)V^*(\mathbf{U} + \Delta\mathbf{U}, \nu + \Delta\nu)\rangle$ between the visibilities at two different baselines and frequencies. We use equation (6) to calculate this. In general, the behaviour of the correlation will depend on both, the magnitude and the direction of $\Delta\mathbf{U}$. We have separately considered two situations (1) $\Delta\mathbf{U}$ parallel to $\mathbf{U}$ and (2) $\Delta\mathbf{U}$ perpendicular to $\mathbf{U}$. The results are presented in Figs. 6 and 7 for $U = 200$. The results are similar for other values of U.
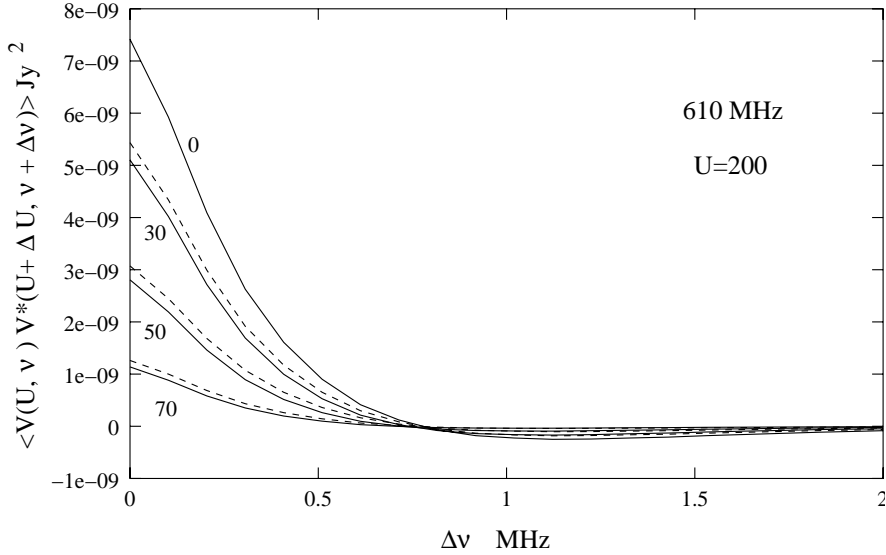
Figure 6. This shows the expected correlation between the visibilities $V(\mathbf{U}, \nu)$ and $V(\mathbf{U}+\Delta\mathbf{U}, \nu+\Delta\nu)$ at two different baselines and frequencies. The results are shown at $U = 200$ for different values of ΔU (shown in the figure), when $\Delta\mathbf{U}$ is parallel to $\mathbf{U}$ (solid line) and $\Delta\mathbf{U}$ is perpendicular to $\mathbf{U}$ (dashed line). These results are for the 610MHz band.
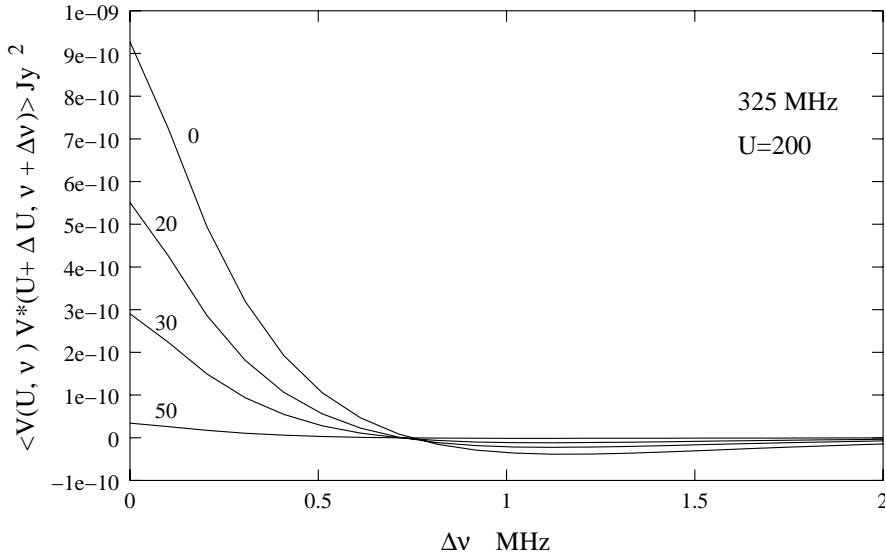
Figure 7. This is the same as figure 6 except that it is for the 325MHz band. Here the results do not depend on whether $\Delta\mathbf{U}$ is parallel or perpendicular to $\mathbf{U}$.

We find that the correlation falls rapidly as $\Delta\mathbf{U}$ increases. The $\Delta\mathbf{U}$ dependence is found to be anisotropic for the 610 MHz band. The correlation falls faster when $\Delta\mathbf{U}$ is parallel to $\mathbf{U}$ than when it is perpendicular to $\mathbf{U}$. The $\Delta\mathbf{U}$ dependence is found to

be isotropic for the 325 MHz band, in both the bands the $\Delta \mathbf{U}$ dependence can be well approximated by a Gaussian

$$\langle V(\mathbf{U}, \nu) V^*(\mathbf{U} + \Delta \mathbf{U}, \nu + \Delta \nu) \rangle = \exp \left[-\left(\frac{\Delta U_\parallel}{a_\nu} \right)^2 - \left(\frac{\Delta U_\perp}{b_\nu} \right)^2 \right]$$
$$\times \langle V(U, \nu) V^*(U, \nu + \Delta \nu) \rangle \qquad (11)$$
$$a_{610} = 50, \ b_{610} = 53 \qquad\qquad a_{325} = b_{325} = 28,$$

where $\Delta U_\parallel$ is the component of $\Delta \mathbf{U}$ parallel to $\mathbf{U}$, and $\Delta U_\perp$ is the component perpendicular to $\mathbf{U}$. This fitting function is found to be a very good approximation over a large range of of $\mathbf{U}$, $\Delta \mathbf{U}$ and $\Delta \nu$.

It should be noted that contrary to the predictions of our fitting formula (equation 11), the cross-correlations between the different baselines is *exactly zero* i.e., $\langle V(\mathbf{U}, \nu) V^*(\mathbf{U} + \Delta \mathbf{U}, \nu + \Delta \nu) \rangle = 0$ if $| \Delta \mathbf{U} | > D/\lambda$. This is a consequence of the fact that we have modeled the Fourier transform of the beam pattern as a Gaussian whereas it actually has compact support (discussed in section 2). The cross-correlations are zero for $| \Delta \mathbf{U} | > 90$ at 610 MHz and $| \Delta \mathbf{U} | > 50$ at 325 MHz. This discrepancy does not significantly change our results as the Gaussian function in equation (11) is extremely small ($< 1\%$) at the values of $| \Delta \mathbf{U} |$ where the cross-correlations vanish.

5. Discussion and conclusions

We have investigated in detail the redshifted HI emission signal expected in GMRT observations in the 325MHz and the 610MHz frequency bands. The properties of the visibility-visibility cross-correlation $\langle V(\mathbf{U}, \nu) V^*(\mathbf{U} + \Delta \mathbf{U}, \nu + \Delta \nu) \rangle$ which we propose as a statistical estimator for detecting and analyzing the signal have been studied. The signal is maximum at $U_{\min}$, and we find that $\langle V(\mathbf{U}_{\min}, \nu) V^*(\mathbf{U}_{\min}, \nu) \rangle$ is $\sim 1.3 \times 10^{-8}$Jy^2 at 610 MHz and $\sim 2.5 \times 10^{-9}$Jy^2 at 325 MHz. The signal falls at larger baselines, and it drops by an order of magnitude by $U \sim 10 \times U_{\min}$. Fourteen of the thirty GMRT antennas are located within 1km. These baselines could be used for detecting the signal whereas the larger baselines where there is very little signal could be used as a control to test that what we detect is actually the HI signal and not an artifact.

The correlation $\langle V(\mathbf{U}, \nu) V^*(\mathbf{U}, \nu + \Delta \nu) \rangle$ fall exponentially with increasing $\Delta \nu$, the width being around 0.5 MHz at $U_{\min}$. The correlations fall faster at the larger baselines. The $\Delta \nu$ dependence of the correlation will be crucial in detecting the HI signal, for neither the system noise nor the galactic/extragalactic continuum radiation are expected to have such a feature.

The correlation $\langle V(\mathbf{U}, \nu) V^*(\mathbf{U} + \Delta \mathbf{U}, \nu + \Delta \nu) \rangle$ falls rapidly with increasing $\Delta \mathbf{U}$, the $\Delta \mathbf{U}$ dependence being well described by a Gaussian. The correlation falls by 50% when $| \Delta \mathbf{U} | \sim 42$ at 610MHz and $| \Delta \mathbf{U} | \sim 23$ at 325MHz. The correlation falls by 10% when $| \Delta \mathbf{U} | \sim 16$ at 610MHz and $| \Delta \mathbf{U} | \sim 9$ at 325MHz. The $\Delta \mathbf{U}$ dependence plays an important role in three different contexts in this discussion.

- For a pair of antennas at a separation $\mathbf{d}$, the value $\mathbf{U} = \mathbf{d}/\lambda$ changes as ν varies across the frequency band. The correlation between the visibilities measured by

this pair of antennas at two different frequencies will be $\langle V(\mathbf{U}, \nu)V^*(\mathbf{U}+\Delta\mathbf{U}, \nu+\Delta\nu)\rangle$, where $\Delta\mathbf{U} = -(\Delta\nu/\nu)\,\mathbf{U}$ arises due to the change in wavelength. For $U \leq 500$ and $\Delta\nu \leq 1\mathrm{MHz}$, which is the range where there is significant correlation, we find $|\Delta\mathbf{U}| < 2$. Our investigations show that this effect will cause the correlation to fall by a factor less than 10% from $\langle V(\mathbf{U}, \nu)V^*(\mathbf{U}, \nu+\Delta\nu)\rangle$, and we can ignore the change in $\mathbf{U}$ due to the change in frequency.

- The rotation of the earth causes all the baselines to change in time and for a time interval Δt we have $\Delta\mathbf{U} = \Delta t\,\vec{\omega}\times\mathbf{U}$. This has to be taken into account when deciding on the time interval Δt over which the visibility can be integrated without its value changing significantly. Our analysis shows that the correlation in the visibility falls by 10% or less if $\Delta U \leq 10$ and we use this as a criteria to determine Δt. In the range $U \leq 500$ where we expect most of the signal we obtain $\Delta t \sim 5\,\mathrm{min}$.

- As the earth rotates, GMRT observations will produce visibility data for a very large number of baselines $\mathbf{U}$. The values of $\mathbf{U}$ will not be uniformly distributed in the plane of the array, and there will be regions which are more densely sampled than others. It will be useful to reduce the volume of data by combining the data in the densely sampled regions and thereby produce a set of values of the visibility on a regular grid of $\mathbf{U}$ values. Our analysis shows that the grid has to be at an interval $\Delta U \sim 40$ at 610MHz and ~ 25 at 325MHz, the visibility signal gets uncorrelated after this interval.

Two issues, the system noise and the contribution of the galactic and extragalactic continuum sources have not been addressed here. Further, we have modeled the HI gas as having a continuous distribution, whereas in reality the HI gas is in discrete, small clouds. Investigations are currently underway on these issues and the results will be communicated in forthcoming publications.

Finally, the visibility-visibility correlation signal depends on the equation of state of the dark matter and the dark energy at the epoch when the HI emission originated. This dependence comes in through the redshift space distortion parameter, the comoving distance and its derivative, all of which depend on the equation of state of the universe. We shall address the possibility of using HI observations to probe the equation of state of the universe at high redshifts in a future paper.

Acknowledgement

References

Bharadwaj, S., Nath, B., Sethi, S. K. 2001, *JApA*. **22,** 21.
Bharadwaj, S., Sethi, S. K. 2001, *JApA*, **22,** 293.
Bunn, E. F., White, M. 1996, *Astrophys. J.*, **460,** 1071.
Efstathiou, G., Bond, J. R., White, S. D. M. 1992, *MNRAS*, **250,** 1p.

Lahav, O., Lilje P. B., Primack J. R., Rees M., 1991, *MNRAS*, **251,** 128.

Lanzetta, K. M., Wolfe, A. M., Turnshek, D. A. 1995, *Astrophys. J.*, **430,** 435.

Peebles, P. J. E. 1980, *The Large-Scale Structure of the Universe* (Princeton: Princeton University Press)

Péroux, C., McMahon, R. G., Storrie-Lombardi, L. J., Irwin, M. J. 2001, *MNRAS* astro-ph/0107045

Saini, T., Bharadwaj, S., Sethi, K. S. 2001, *Astrophys. J.*, **557,** 421.

Storrie–Lombardi, L. J., McMahon, R. G., Irwin, M. J. 1996, *MNRAS*, **283,** L79.

Swarup, G., Ananthakrishan, S., Kapahi, V. K., Rao, A. P., Subrahmanya, C. R., Kulkarni, V. K. 1991, *Curr. Sci.*, **60,** 95.