Research in Astron. Astrophys. **2009** Vol. **9** No. **6**, 653 – 664
http://www.raa-journal.org *http://www.iop.org/journals/raa*

***R**esearch in*
***A**stronomy and*
***A**strophysics*

Foreground removal of 21 cm fluctuation with multifrequency fitting

Li-Ping He[1,2]

[1] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China; *hlp@bao.ac.cn*
[2] Graduate University of Chinese Academy of Sciences, Beijing 100049, China

Received 2008 July 18; accepted 2008 November 17

Abstract The 21 centimeter (21 cm) line emission from neutral hydrogen in the intergalactic medium (IGM) at high redshifts is strongly contaminated by foreground sources such as the diffuse Galactic synchrotron emission and free-free emission from the Galaxy, as well as emission from extragalactic radio sources, thus making its observation very complicated. However, the 21 cm signal can be recovered through its structure in frequency space, as the power spectrum of the foreground contamination is expected to be smooth over a wide band in frequency space while the 21 cm fluctuations vary significantly. We use a simple polynomial fitting to reconstruct the 21 cm signal around four frequencies 50, 100, 150 and 200 MHz with an especially small channel width of 20 kHz. Our calculations show that this multifrequency fitting approach can effectively recover the 21 cm signal in the frequency range $100 \sim 200\,\mathrm{MHz}$. However, this method doesn't work well around $50\,\mathrm{MHz}$ because of the low intensity of the 21 cm signal at this frequency. We also show that the fluctuation of detector noise can be suppressed to a very low level by taking long integration times, which means that we can reach a sensitivity of $\approx 10\,\mathrm{mK}$ at $150\,\mathrm{MHz}$ with 40 antennas in 120 hours of observations.

Key words: cosmology: theory, observations, diffuse radiation — methods: data analysis

1 INTRODUCTION

Observations of the redshifted 21 cm line of neutral hydrogen are perceived as one of the most promising future probes of the Universe at high redshifts (Ali et al. 2008; Furlanetto et al. 2006; Morales et al. 2004; Santos et al. 2003; Holder et al. 2003). Fluctuations in the 21 cm line emission are induced both by inhomogeneities in the intergalactic neutral gas and in the ionized hydrogen fraction (Di Matteo et al. 2004; Madau et al. 1997; Tozzi et al. 2000; Ciardi et al. 2003; Furlanetto et al. 2004a) and by 'minihalos' with virial temperatures below 10^4 K (Iliev et al. 2002, 2003). Observations of fluctuations in redshifted 21 cm emission from neutral hydrogen will provide unprecedented information about the structure formation, the nature of the first luminous objects and the physical state of the intergalactic medium at high redshifts (Morales et al. 2006; Di Matteo et al. 2002; Cen 2003; Kumar et al. 1995; Gnedin & Ostriker 1997). Upcoming low-frequency radio facilities such as the 21 Centimeter Array (21 CMA) [1] (Pen et al. 2004), the Low Frequency Array (LOFAR) [2], the Mileura Wide-Field Array (MWA) [3] and the

[1] *http://21cma.bao.ac.cn*
[2] *http://www.lofar.org*
[3] *http://www.haystack.mit.edu/arrays/MWA*

Square Kilometer Array (SKA)[4] are currently under construction to explore 21 cm fluctuations associated with neutral hydrogen prior to and during reionization spanning the redshift range $5 \leq z \leq 20$. These will provide a powerful tool for probing the epoch, nature, and sources of reionization in the universe and their implications for cosmology (Di Matteo et al. 2002).

Unfortunately, the 21 cm signal is contaminated by strong foregrounds, which are five orders of magnitude above the 21 cm fluctuations, making its detection from the foreground a major challenge (Furlanetto et al. 2006; Santos et al. 2005). It is therefore crucial to study the property of the foregrounds in order to detect the 21 cm signal in future observations. Foregrounds are expected to be smooth as a function of frequency and the contribution at different frequencies are highly correlated, while the 21 cm signal varies rapidly in frequency space and is expected to be uncorrelated. This character may allow us to separate the 21 cm signal from foregrounds (Ali et al. 2008; Zaldarriaga et al. 2004). Di Matteo et al. (2002), Zaldarriaga et al. (2004) and Santos et al. (2005) suggested that the spectrum of the foregrounds can be well approximated by a power law in one pixel and the spectra can be fitted to individual spatial-frequency pixels in visibility space. The basic approach in Wang et al. (2006) is to subtract foregrounds separately in each angular direction in the sky by fitting their total intensity dependence on frequency by a log-log polynomial. They discussed their analysis with one-dimensional (1-D) simulated data over a small redshift range around $z \approx 8.09$.

It is worthwhile expanding this analysis over the full relevant redshift range. Our approach is similar, but it emphasizes the removal of the foreground in two-dimensional (2-D) data. This paper complements the analysis in Wang et al. (2006), and discusses how well the multifrequency fitting can perform on 2-D data in the frequency range $50 \sim 200\,\mathrm{MHz}$. In Section 2, we present various models for simulation, such as the reionization model, foreground models and detector noise model. In Section 3, we describe the foreground removal method used in our calculation. We summarize our results in Section 4 and our conclusions in Section 5.

2 SIMULATION OF MODELS

2.1 Reionization Model

Observed 21 cm brightness temperature in the direction $\boldsymbol{n}$ is given by (Zaldarriaga et al. 2004; Furlanetto et al. 2006)

$$T(\boldsymbol{n},\nu) \;=\; T_0 \int dr W_{r_0}(r)\psi(\boldsymbol{n},r); \tag{1}$$

$$T_0 \;=\; 23\left(\frac{\Omega_{\rm b}h^2}{0.02}\right)\left(\frac{0.15}{\Omega_{\rm m}h^2}\frac{1+z}{10}\right)^{\frac{1}{2}}\ {\rm mK}; \tag{2}$$

$$\psi(\boldsymbol{n},r) \;=\; (1+\delta)\,x_{\rm H}\left(\frac{T_{\rm s}-T_{\rm CMB}}{T_{\rm s}}\right), \tag{3}$$

where T_0 is a normalization constant at r_0. The comoving radial distance corresponds to the observed frequency ν, and $\psi(\boldsymbol{n},r)$ is the dimensionless brightness temperature. $W_{r_0}(r)$ is the projection function which is peaked at r_0. δ, $x_{\rm H}$, and $T_{\rm s}$, respectively, correspond to the matter density contrast, neutral hydrogen fraction and the spin temperature of the IGM. $T_{\rm CMB} = 2.73(1+z)\,{\rm K}$ is the CMB temperature at redshift z.

The spin temperature $T_{\rm s}$ depends on the kinetic temperature $T_{\rm k}$ of the IGM (Zaldarriaga et al. 2004; Furlanetto et al. 2004a). The IGM is likely to be heated shortly after the first sources appear through X-ray, photoionization and shock heating (Chen et al. 2003; Furlanetto et al. 2004a; Zaldarriaga et al. 2004). Once heating has occurred, these processes will rapidly heat the IGM to $T_{\rm k} > T_{\rm CMB}$ (Zaldarriaga et al. 2004; Venkatesan et al. 2001; Chen et al. 2003). The situation most relevant to observations has $T_{\rm s} \sim T_{\rm k} \gg T_{\rm CMB}$, which should be reasonable even relatively early in reionization

[4] *http://www.skatelescope.org*

(Zaldarriaga et al. 2004; Furlanetto et al. 2004a,b). In this limit, the temperature ψ can be written by $\psi(\boldsymbol{n}, r) = (1+\delta)x_{\rm H}$. The correlation of the 21 cm temperature fluctuation $\Delta_\psi(\boldsymbol{n}, r) = \psi(\boldsymbol{n}, r) - \overline{\psi}$ can be represented by

$$\mu(x_{12}) = \langle \psi(\boldsymbol{x}_1)\psi(\boldsymbol{x}_2)\rangle - \langle\psi\rangle^2, \tag{4}$$

which depends on the parameters δ and $x_{\rm H}$. x_{12} is the distance between points x_1 and x_2. We can model the correlations of the neutral hydrogen fraction field as $\langle x_{\rm H}(x_1)x_{\rm H}(x_2)\rangle = \overline{x}_{\rm H}^2 - (\overline{x}_{\rm H} - \overline{x}_{\rm H}^2)f(x_{12}/R)$ under the simplifying assumption that the fluctuations in $x_{\rm H}$ were produced by a set of uncorrelated "bubbles" of typical size R, where $\overline{x}_{\rm H}$ is the average value of $x_{\rm H}$ and $f(x)$ is a function with the following limits: $f(x) \approx 1$ for $x \ll 1$ and $f(x) \approx 0$ for $x \gg 1$ (Zaldarriaga et al. 2004).

We can obtain the expression for the power spectrum of ψ by Fourier transforming Equation (4) (Zaldarriaga et al. 2004),

$$\Delta_\psi^2(k) = \overline{x}_{\rm H}^2\Delta_\rho^2(k) + (\overline{x}_{\rm H} - \overline{x}_{\rm H}^2)\Delta_{x\rho}^2(k) + (\overline{x}_{\rm H} - \overline{x}_{\rm H}^2)\Delta_x^2(k), \tag{5}$$

where

$$\Delta_x^2(k) = \frac{k^3\hat{f}(k)}{2\pi^2}, \tag{6}$$

$$\Delta_{x\rho}^2(k) = \frac{k^3}{2\pi^2}\int \frac{d\boldsymbol{k}'}{(2\pi^3)}P_\rho(\boldsymbol{k}-\boldsymbol{k}')\hat{f}(\boldsymbol{k}'). \tag{7}$$

$\hat{f}(k)$ is the Fourier transform of $f(x)$ and $P_\rho(k)$ is the power spectrum of the density fluctuations.

The corresponding three-dimensional power spectrum of the 21 cm temperature fluctuation is given by

$$P_\psi(k) = \frac{2\pi^2}{k^3}\Delta_\psi^2(k). \tag{8}$$

We assume a ΛCDM concordance cosmology with $\Omega_{\rm m} = 0.27, \Omega_\Lambda = 0.73, \Omega_{\rm b} = 0.04, h = 0.71, \sigma_8 = 0.84$ throughout our calculations.

The average properties of the universe evolve on a timescale of $\Delta z \sim 1$, while the local properties of the universe change on a much smaller scale of $\Delta z < 1$. Across such a small redshift range, we can approximate parameters by their values at z_0 (Wang et al. 2006).

The importance of the redshift space distortions is determined by the ratio $\mathcal{R} = (\Delta\nu/\nu)/(v/c)$, where v is the typical random bulk velocity of the gas. Redshift space distortions will be unimportant if $\mathcal{R} > 1$ (Zaldarriaga et al. 2004).

If we make the above-mentioned approximation and neglect the redshift distortions, the 21 cm signal near a given z_0 has an isotropic 3-D power spectrum $P_{\rm 3D}$ which can be projected onto a 2-D power spectrum $P_{\rm 2D}(k, z_0)$ (Wang et al. 2006; Peacock 1999):

$$P_{\rm 2D}(k, z \approx z_0) = \frac{1}{2\pi}\int_k^\infty P_{\rm 3D}(k', z \approx z_0)\frac{k'}{\sqrt{k'^2 - k^2}}dk'. \tag{9}$$

The above approximation would require us to observe the 21 cm signal in a sufficiently narrow redshift range $\Delta z < 1$, corresponding to a bandwidth $\lesssim 10\,{\rm MHz}$. In this paper, we simulated the 21 cm fluctuation at four frequencies 50, 100, 150 and 200 MHz with a total bandwidth of 2 MHz with each frequency split into 100 channels.

The brightness temperature of the cosmological signal used in our simulation is produced from fast Fourier transforms (FFTs). The simulated signal in real space in the region $0 < x < L, 0 < y < L$ is

$$f(x,y) = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\Big[A(i,j)\cos\Big(\frac{2\pi\sqrt{x^2+y^2}}{L}\sqrt{i^2+j^2}\Big) + B(i,j)\sin\Big(\frac{2\pi\sqrt{x^2+y^2}}{L}\sqrt{i^2+j^2}\Big)\Big], \tag{10}$$

where $A(i,j)$ and $B(i,j)$ are Gaussian random variables with zero mean and standard deviations $\Delta A = \Delta B = [P_{\rm 2D}(k,z_0)/2]^{1/2} = [P_{\rm 2D}(2\pi\sqrt{i^2+j^2}/L,z_0)/2]^{1/2}$ (Wang et al. 2006). L is a box size which should be small enough to satisfy $\Delta z \ll 1$, and N is an integer which should be large enough to include all the information from the $P_{\rm 2D}(k)$.

Although we simulate the 21 cm signal as a Gaussian random field, in reality, the 21 cm signal is non-Gaussian in nature, because the intensity of the 21 cm signal is a function of the spin temperature of the gas and the local ionization fraction, which may have a very complex distribution, in addition to the density fluctuations (Morales et al. 2004). However, the key quantity in which we are interested depends on the power spectra of the signal, foregrounds and detector noise, instead of whether the statistics are Gaussian or not. Therefore, we can assume the 21 cm signal to be a Gaussian random field and the signals between the different frequency channels are completely uncorrelated.

2.2 Foreground Model

The main foreground components which dominate the sky in frequency range $50 \sim 200\,{\rm MHz}$ are: Galactic synchrotron emission, Galactic free-free emission and emission from extragalactic sources.

Galactic synchrotron emission originates from the interaction between the free electrons in the interstellar medium and the Galactic magnetic field. It is the dominant contamination, and comprises a fraction of about 70% at $150\,{\rm MHz}$ (Shaver et al. 1999). Its intensity can be given by a running power law in frequency (Wang et al. 2006):

$$I_{\rm syn} = A_{\rm syn}\Big(\frac{v}{v_*}\Big)^{-\alpha_{\rm syn}-\Delta\alpha_{\rm syn}\log(\frac{\nu}{\nu_*})}. \tag{11}$$

Amplitude $A_{\rm syn}$ and the spectral index $\alpha_{\rm syn}$ gradually vary with position and frequency (Shaver et al. 1999; Lawson et al. 1987). Shaver et al. (1999) and Jelić et al. (2008) suggested that the difference of the amplitude $A_{\rm syn}$ and spectral index $\alpha_{\rm syn}$ between different frequency directions and lines of sight is very slight. The intensity $I_{\rm syn}$ is still a very smooth function of frequency. Furthermore, the main purpose of this paper is to test the accuracy of foreground removal over every line of sight. Therefore, the slight variation of $A_{\rm syn}$ and $\alpha_{\rm syn}$ can be neglected, and the same $A_{\rm syn}$ and $\alpha_{\rm syn}$ can be used throughout our field of view. Tegmark et al. (2000) suggested $\alpha_{\rm syn} = 2.8$ based on the data from Platania et al. (1998). The spectral index varies across the sky with $\Delta\alpha_{\rm syn} = 0.15$. Here, we can take a conservative estimate of 0.1 (Santos et al. 2005; Wang et al. 2006). $\nu_* = 150\,{\rm MHz}$ is the reference frequency at which the normalization is done. The amplitude $A_{\rm syn}$ can be extrapolated from Tegmark et al. (2000), so we get $A_{\rm syn} = 335.4\,{\rm K}$ (Wang et al. 2006).

The free-free emission is due to the very diffuse ionized gas. It contributes only $\sim 1\%$ of the total foreground within the frequency range we considered (Shaver et al. 1999). It follows a similar spatial power spectrum as that of the synchrotron emission, but with a lower amplitude. Its emission intensity can also be written as:

$$I_{\rm ff} = A_{\rm ff}\Big(\frac{v}{v_*}\Big)^{-\alpha_{\rm ff}-\Delta\alpha_{\rm ff}\log(\frac{\nu}{\nu_*})}, \tag{12}$$

where $\alpha_{\rm ff} = 2.15$, $\Delta\alpha_{\rm ff} = 0.01$ (Tegmark et al. 2000; Shaver et al. 1999) and $A_{\rm ff} = 33.5\,{\rm K}$ (Wang et al. 2006).

The extragalactic radio sources also provide a serious contamination of the brightness temperature fluctuations of 21 cm emission, with the contribution of the order 27% at $150\,{\rm MHz}$ (Shaver et al. 1999). Power law spectra of the extragalactic sources are much more complicated than that of the Galactic synchrotron radiation and free-free emission. They are overlapped spectra with different spectral indices.

We adopt a more complicated extragalactic source model in our simulation, with a huge number of sources randomly positioned in the sky. The intensity in every pixel of each source is (Wang et al. 2006)

$$I_{\rm ps}(x,y) = \left(\frac{dB}{dT}\right)^{-1}\Omega_{\rm sky}^{-1}S_i^*\left(\frac{\nu}{\nu_*}\right)^{-\alpha_i}\exp\left(-\frac{(x-x_i)^2}{2\sigma_x^2}-\frac{(y-y_i)^2}{2\sigma_y^2}\right), \tag{13}$$

where (x_i, y_i) is the central position of the source, and σ_x, σ_y represent the shape of the source. The conversion factor dB/dT can be deduced from the Rayleigh-Jeans law $B = 2kT/\lambda^2$, and we attain $dB/dT = 6.9 \times 10^5\,\mathrm{mJy\ K^{-1}}$ at frequency $150\,\mathrm{MHz}$. Ω_{sky} is the solid angle per pixel, $\Omega_{\mathrm{sky}} = 10.8\,\mathrm{arcmin}^2$ at $150\,\mathrm{MHz}$ when the assumed longest baseline is 3 km. S_i^* is the flux of the ith source, randomly generated from the source count distribution $dN/dS = 4(S/1\,\mathrm{Jy})^{-1.75}$ (Di Matteo et al. 2004; Wang et al. 2006; Santos et al. 2005; Pierpaoli 2003; Ali et al. 2008). Di Matteo et al. (2004) suggested that sources brighter than $S_{\mathrm{c}} = 0.1\,\mathrm{mJy}$ can be detected by potential low-frequency interferometers and directly removed from the images. Di Matteo et al. (2002) said that the appearance of the radio sky below $S_{\min} = 1\ \mu\mathrm{Jy}$ is not well known. Hence, here we take the maximum flux $S_{\max} = 0.1\,\mathrm{mJy}$, and the minimum flux $S_{\min} = 0.001\,\mathrm{mJy}$. α_i is the spectral index of the ith source, randomly generated from the Gaussian distribution

$$f(\alpha) = \frac{1}{\sqrt{(2\pi)}\sigma_\alpha} \exp\left[-\frac{(\alpha-\alpha_0)^2}{2\sigma_\alpha^2}\right]. \tag{14}$$

Wang et al. (2006) showed that the residuals of the subtraction are insensitive to spectral indices as long as $\sigma_\alpha \leq 60$. Here, we allow $\sigma_\alpha = 2$ throughout our calculations.

2.3 Detector Noise Model

The sensitivity of a receiving system is determined by the system noise. For an array of n antennas, there are $N = n(n-1)/2$ independent baselines. If the signals are integrated over time interval τ, the theoretical rms noise level in terms of temperature is

$$\Delta T = \frac{K\lambda^2 T_{\mathrm{sys}}}{A_{\mathrm{e}}\Omega_{\mathrm{b}}\sqrt{2N\Delta\upsilon\tau}}, \tag{15}$$

where K is a loss factor depending on the receiver system. For total power at the receiver, K is close to unity, and we will set $K = 1$ (Rohlfs 1986; Furlanetto et al. 2006). A_{e} is the effective area of the antenna, like 21 CMA, $A_{\mathrm{e}} = 216\,\mathrm{m}^2$. T_{sys} is the system temperature. It has two major contributions. The first comes from the sky and is frequency and elevation dependent, and the second comes from the receivers and is frequency independent. For a Gaussian beam, $\Omega_{\mathrm{b}} = \Omega_{\mathrm{MA}} = 1.133\theta_{\mathrm{b}}^2$, where θ_{b} is the full width to half power. $\Delta\nu$ is the channel width. Here, we take $\Delta\nu = 2\,\mathrm{MHz}/100 = 20\,\mathrm{kHz}$.

In this paper, we plan to measure the brightness fluctuations in the frequency range $50 \sim 200\,\mathrm{MHz}$. The angular extent we use is $20^\circ \times 20^\circ$ at $50\,\mathrm{MHz}$, $10^\circ \times 10^\circ$ at $100\,\mathrm{MHz}$, $7^\circ \times 7^\circ$ at $150\,\mathrm{MHz}$, and $5^\circ \times 5^\circ$ at $200\,\mathrm{MHz}$. The corresponding angular resolutions are $8.4' \times 8.4'$, $4.2' \times 4.2'$, $2.8' \times 2.8'$, $2.0' \times 2.0'$. The number of pixels in each map is $N_{\mathrm{pixel}} = 128^2$. The lowest noise level we can reach is $\Delta T^{\mathrm{N}} = \Delta T/\sqrt{N_{\mathrm{pixel}}}$, which is at the level of mK.

We can reach such low noise levels by taking sufficiently long integration times using the synthesis array with forty antennas. We only selected the 21 cm signal when its fluctuation exceeded 3σ of rms noise. Table 1 shows the results of the total integration time at every frequency.

Table 1 Integration Time at Different Frequencies, with 40 Antennas, 20 kHz Channel Width

ν [MHz]	$\Delta T_{21\mathrm{cm}}$ [mK]	Time [d]	ν [MHz]	$\Delta \mathrm{T}_{21\mathrm{cm}}$ [mK]	Time [d]
100	3.32	3189	160	23.67	11
110	23.87	39	170	12.83	32
120	63.44	4	180	6.8	102
130	79.18	2	190	3.56	342
140	61.38	3	200	1.95	1066
150	39.59	5			

3 FOREGROUND-REMOVAL METHOD

Consider a pixelized sky map consisting of M numbers $y_f^1, y_f^2, \cdots, y_f^{\rm M}$, where y_f^i is the temperature of the ith pixel at frequency $\nu_f (f = 1, 2, \cdots, F)$. F is the frequency channel number. In this paper, we take $F = 100$. The maps are simulated on 128^2 grids, so we have $M = 128^2$. If we remove the foreground pixel-by-pixel, this means we need to do 128^2 fittings and removals. Taking one pixel for example, $y_f = \log(I^f_{21\rm cm} + I^f_{\rm fg} + I^f_{\rm n})$ is the total intensity on the pixel, including the 21 cm signal $I^f_{21\,\rm cm}$, foreground contamination $I^f_{\rm fg}$, and detector noise $I^f_{\rm n}$.

The log-frequencies $x_f = \log(\nu_f)$ have some linear combination with the sky temperature y_f. Their relationship can be represented by

$$\begin{aligned} y_f &= \sum_{j=0} x_f^i a_i = x_f^0 a_0 + x_f^1 a_1 + x_f^2 a_2 + \cdots \\ &= a_0 + a_1 \log(\nu_f) + a_2 [\log(\nu_f)]^2 + \cdots \end{aligned} \tag{16}$$

Grouping x_f, y_f into an $F \times m$ matrix $\boldsymbol{X}$ and F-dimensional vector $\boldsymbol{y}$, respectively, we can generally write

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{a} + \boldsymbol{n}, \tag{17}$$

where m is the number of fitting parameters and $\boldsymbol{a}$ is the m-dimensional parameter vector, $\boldsymbol{n}$ is the vector that cannot be represented by parameter $\boldsymbol{a}$, including the contributions from the 21 cm signal, detector noise and residual foregrounds.

In our computation, we fit the log intensity of the foreground as $\log(I) = a_1 + a_2 \log(\nu) + a_3 \log(\nu)^3 + a_4 \log(\nu)^5$. The number of fitting parameters is $m = 4$. Now we can rewrite Equation (17) as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_F \end{bmatrix} = \begin{bmatrix} x_1^0 & x_1^1 & x_1^3 & x_1^5 \\ x_2^0 & x_2^1 & x_2^3 & x_2^5 \\ \vdots & \vdots & \vdots & \vdots \\ x_F^0 & x_F^1 & x_F^3 & x_F^5 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_F \end{bmatrix}.$$

We estimate $\widetilde{\boldsymbol{a}}$ of the sky map $\boldsymbol{a}$ by minimizing χ^2

$$\chi^2 = \sum_{f=1}^{F} \frac{\left(y_f - \sum_{j=1}^{m} x_i^i a_i\right)^2}{\sigma_i^2}.$$

Tegmark et al. (1997, 2000) showed that the minimum-variance choice is

$$\widetilde{\boldsymbol{a}} = (\boldsymbol{X^t N^{-1} X})^{-1} \boldsymbol{X^t N^{-1} y},$$

where $\boldsymbol{N}$ is the covariance matrix of the contributions from the 21 cm signal and the detector noise. We assume that both the 21 cm signal and detector noise are uncorrelated between different frequencies, so $\boldsymbol{N}$ becomes

$$N = \langle n n^{\rm t} \rangle = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{\rm n} \end{bmatrix} \begin{bmatrix} n_1 & n_2 & \ldots & n_{\rm n} \end{bmatrix} = \begin{bmatrix} n_1^2 & 0 & \ldots & 0 \\ 0 & n_2^2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & \vdots & \ldots & n_n^2 \end{bmatrix}.$$

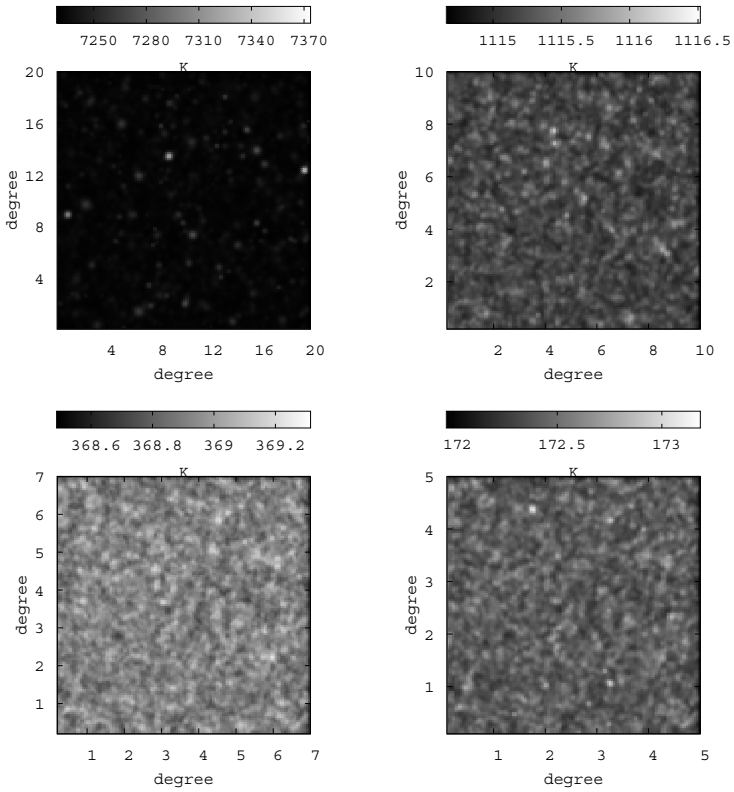

Fig. 1 Simulated maps of total intensity emission of synchrotron radiation, free-free emission foreground, extragalactic point sources, detector noise and the 21 cm signal, at $50\,\mathrm{MHz}$ (top-left), $100\,\mathrm{MHz}$ (top-right), $150\,\mathrm{MHz}$ (bottom-left) and $200\,\mathrm{MHz}$ (bottom-right).

4 RESULTS

In this section, we present the results of the detection of the 21 cm signal from the total intensity maps. We simulated the foregrounds and 21 cm signal around the four frequencies mentioned in Section 2 in a $2\,\mathrm{MHz}$ frequency width. The variation, σ, over the map in the frequency range $100 \sim 200\,\mathrm{MHz}$ is $\sim \mathrm{mK}$ for the detector noise which is the fiducial value for a future-generation experiment. Note that the mean values of the 21 cm signal and noise are set to zero.

The detection of the 21 cm signal is done in intensity maps by analysis along the frequency direction. Figure 1 shows the maps of the total intensity including the Galactic synchrotron emission, Galactic free-free emission, emission from unresolved extragalactic sources, the 21 cm signal and the detector noise. The maps are simulated around 50, 100, 150 and 200MHz in steps of $20\,\mathrm{kHz}$. The fields of view at these frequencies are $20^\circ \times 20^\circ, 10^\circ \times 10^\circ, 7^\circ \times 7^\circ$, and $5^\circ \times 5^\circ$, with each divided into 128^2 pixels.

The simulation results in a randomly chosen line of sight are shown in Figure 2. The top panel shows the total foreground contamination in a pixel, the middle panel includes the foregrounds and the 21 cm signal. We can see the small fluctuation in the middle panel, which is caused by the fluctuation of the 21 cm signal. The bottom panel shows the comparison between the simulated 21 cm signal (solid line) and the recovered 21 cm signal (dashed line). The residual (dotted curve) represents the error in the foreground fitting. The recovered 21 cm signals are in good agreement with the simulated ones at 100, 150, and 200 MHz. The disagreement at $50\,\mathrm{MHz}$ is due to the low intensity of the 21 cm signal

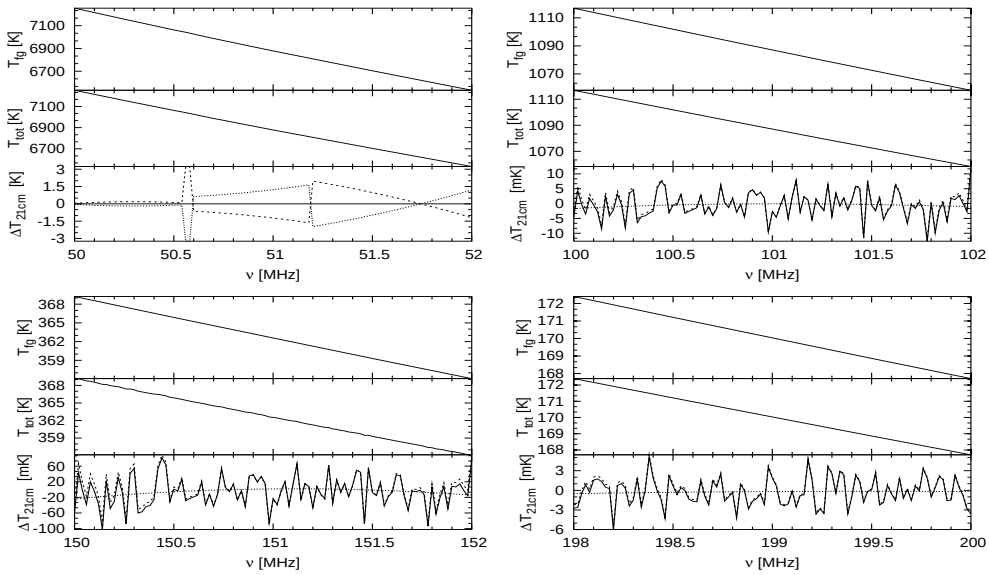

Fig. 2 Signal at one pixel. The top panel shows the smooth component of the foreground, the middle panel has the 21 cm signal added. The bottom is the comparison between the simulated 21 cm signal (solid line) and the recovered 21 cm signal (dashed line). The dots in the bottom panel represents the residual.

level, which is about $10^{-20}\,\mathrm{mK}$. The fitting precision is at the level of $0.1\mathrm{mK}$. Therefore, the recovered signals are dominated by the fitting error, and much higher than the simulated signals.

We remove the smooth foreground components from the total intensity along the line of sight for each individual pixel by polynomial fitting in a logarithmic scale. The solid line in Figure 3 represents the simulated total intensity of foregrounds along one line of sight, and the dashed line is the fitted intensity. As we can see, there is good agreement between the simulated and fitted foregrounds. However, due to the presence of many sources per pixel, which can lead to a complex cumulative spectrum, the spectrum cannot be exactly fitted by a simple power-law model. The fitting error is presented in the bottom panel of the map. The errors are very small which means that the foreground can be well approximated by our fitting function.

Figure 4 shows the comparison between the simulated and recovered 21 cm signal in terms of standard deviation. The top panel of the map at $50\,\mathrm{MHz}$ is the standard deviation σ of the simulated 21 cm signal over all lines of sight, and the bottom panel is the σ of the recovered 21 cm signal. Because the fitting error is 19 orders of magnitude above the 21 cm signal intensity at 50 MHz, the σ of the recovered 21 cm signal shown in the bottom panel is dominated and corrupted by the fitting error. The curve shape actually represents the trace of the fitting error. The solid line in the top panel of other maps represents σ of the simulated 21 cm signal, and the dashed line the σ of the recovered 21 cm signal. The difference between them is explained by the relative error in the bottom panel. All of them are very small, which show that this polynomial fitting in logarithmic scale can effectively reconstruct the 21 cm signal in the frequency range $100 \sim 200\,\mathrm{MHz}$. After removing the fitted foregrounds from the simulated total signals, the residuals include the 21 cm signal, detector noise and foreground fitting residuals. Figure 5 shows the intensity maps of the residuals after foreground removal. As we can see in Figure 2, the foreground residual is much lower than the 21 cm signal, which is about one order of magnitude below the 21 cm signal, and as mentioned in Section 2.3, the 21 cm fluctuation exceeds 3σ of the variation of detector noise, so the residual intensity is dominated by the 21 cm signal. From these maps, we can see that the results are really good except at $50\,\mathrm{MHz}$.

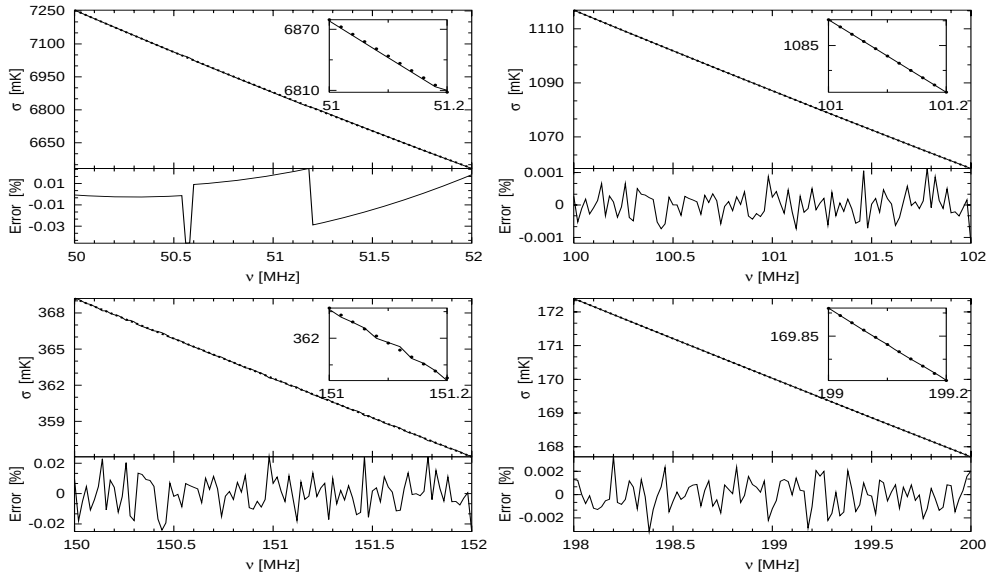

Fig. 3 Simulated total intensity and fit total intensity in a single pixel. The solid line is the simulated signal and the dashed line is the fit signal. The bottom represents the relative fitting error.

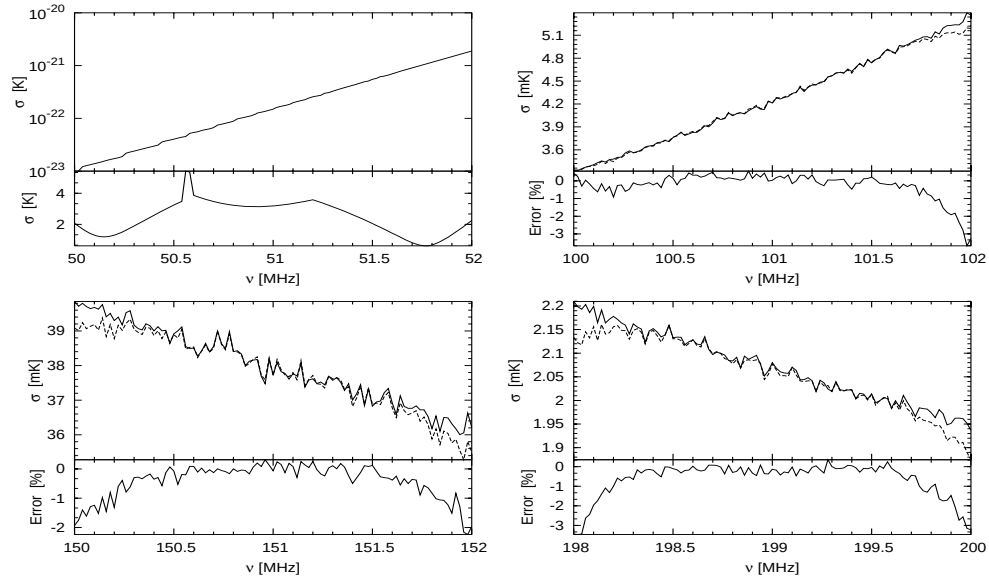

Fig. 4 Detection of 21 cm signal from maps with 128^2 grids around 50, 100, 150 and 200 MHz. The top panel at $50\,\mathrm{MHz}$ shows the standard deviation σ of the simulated 21 cm signal over all lines of sight, and the bottom panel shows σ of the recovered 21 cm signal. For the other frequencies, the σ derived from the simulated and recovered 21 cm signals are represented by the solid line and the dashed line of the top panel, respectively. The relative errors are shown in the bottom panel.

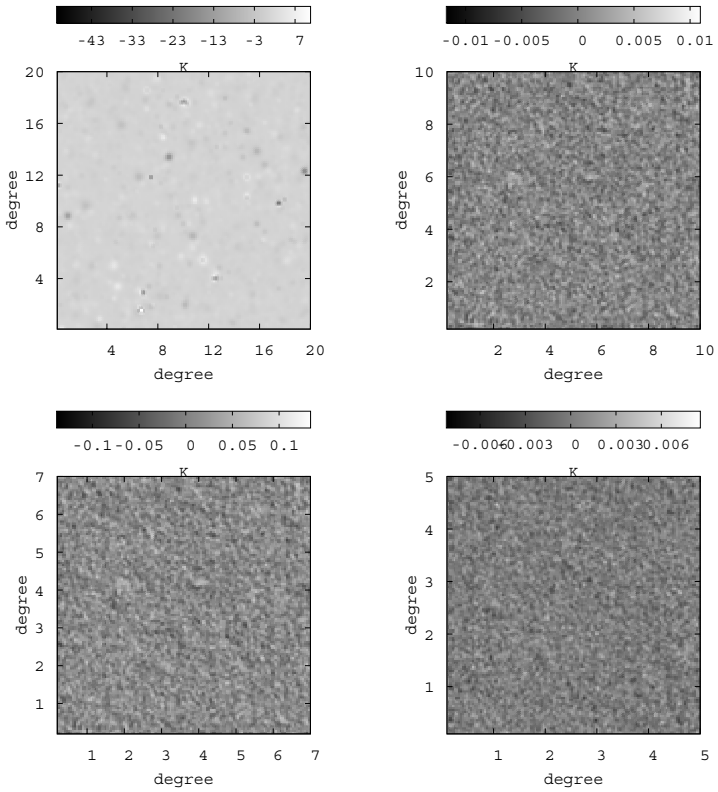

Fig. 5 Recovered maps of residual intensity including the 21 cm signal, detector noise and foreground residual at 50, 100, 150 and 200 MHz, respectively.

We have tested the analysis by changing parameters of the foreground models, such as the foreground amplitudes, parameters σ_α and α_0 in Equation (14). The results showed that the residuals are weakly independent of the selected parameters, which are consistent with what was found by Wang et al. (2006). They found that the effectiveness of this cleaning method is almost independent of the shape and amplitude of the relatively smooth foregrounds. However, we found that the pixel size has a subtle influence on the results. We repeated the analysis with 256^2 grids. The solid line in the top panel of Figure 6 represents the standard deviation σ of the simulated 21 cm signal, and the dashed line represents σ of the recovered 21 cm signal. We can find that the relative error in the bottom panel in maps with 256^2 grids are several times higher than those in maps with 128^2 grids.

We also test how the order of the polynomial fitting influences the 21 cm signal extraction. We change the order of the polynomial from $4^{\rm th}$ order to $6^{\rm th}$ order. The result shows that there is no essential difference between the $4^{\rm th}$ order fitting and $6^{\rm th}$ order fitting. We use a fourth order polynomial fitting in our analysis.

As mentioned in Section 2.2, the flux range of the extragalactic radio sources is $10^{-6} \sim 10^{-4}\,{\rm Jy}$. Now, we increase the maximum flux and test how the cut-off flux of the radio sources influences the result of the 21 cm signal extraction. The same analysis was done in the flux range $10^{-6} \sim 10^{2}\,{\rm Jy}$. The results of mapping on 128^2 grids and 256^2 grids show that an increase in the maximum of the flux has little influence on the result of foreground removal. This test indicates that our method can effectively remove the foregrounds as long as the total foreground signal can be well approximated by a log-log polynomial.

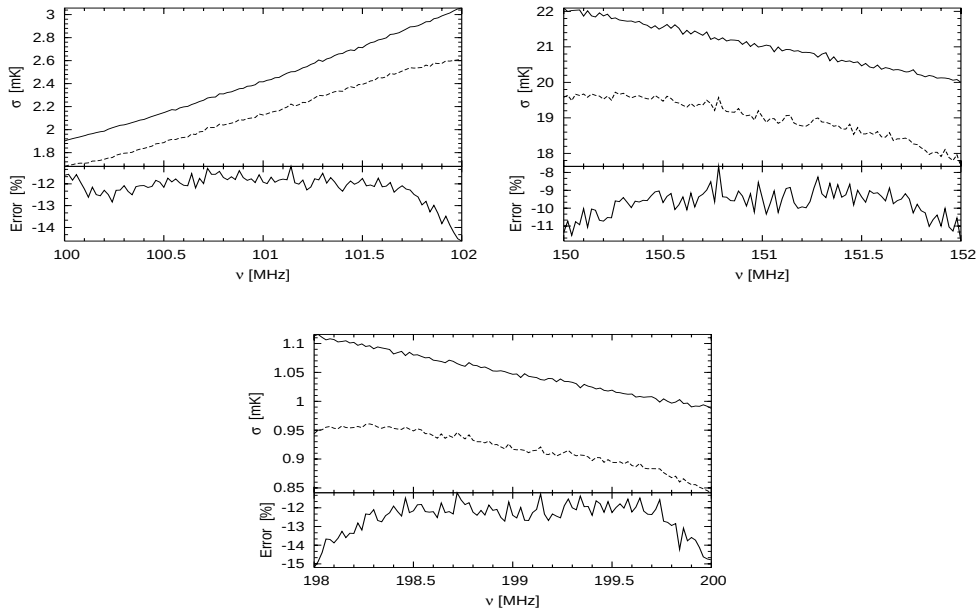

Fig. 6 Detection of 21 cm signal from maps with 256^2 grids around 100, 150 and 200 MHz. The solid line in the top panel represents the σ of simulated 21 cm signal, and the dashed line is the σ of recovered 21 cm signal. The bottom panel shows the relative error.

5 CONCLUSIONS

The measurement of highly redshifted 21 cm neutral hydrogen emission will provide unprecedented information about the epoch of reionization (Santos et al. 2005; Morales et al. 2006). Unfortunately, this experiment will suffer from a high degree of foreground contamination which is more than 5 orders of magnitude above the 21 cm fluctuation (Wang et al. 2006; Santos et al. 2005; Jelić et al. 2008). Several foreground-subtraction techniques have been suggested in the literature (Morales et al. 2006). Zaldarriaga et al. (2004) proposed that the smooth spectra models be fit to individual spatial-frequency pixels in the visibility space, which should be better at removing emission on larger angular scales. Wang et al. (2006) introduced a subtraction technique which removes the foreground contamination by fitting the many faint continuum sources in each line of sight. In this paper, we expand the pixel-based fitting algorithm to a 2-D image over the full relevant redshift range $5 \leq z \leq 20$.

We have examined how well the log-log polynomial fitting can influence removal of foreground contamination at the frequencies relevant for observations of the redshifted 21 cm signal. We have carried out a set of foreground simulations to model the observation of the 21 cm signal. The foreground simulations cover Galactic synchrotron and free-free emission, and extragalactic emission from radio sources. Maps with 128^2 grids are generated at frequencies $50\,\mathrm{MHz}$ ($20^\circ \times 20^\circ$), $100\,\mathrm{MHz}$ ($10^\circ \times 10^\circ$), $150\,\mathrm{MHz}$ ($7^\circ \times 7^\circ$) and $200\,\mathrm{MHz}$ ($5^\circ \times 5^\circ$).

Foregrounds are subtracted by fitting their intensity dependence on frequency with a quadratic polynomial in log-log space for each pixel in the sky, with σ scalings to determine the effects of foreground subtraction. Based on that, the foregrounds have much smoother frequency spectra than the 21 cm signal, so a high level of foreground removal is achieved which is really encouraging.

We did the same analysis on maps with 256^2 pixels to test the influence of the pixel size. The errors increase with the pixel number. We found that improving the number of fitting parameters from 4 to 6 gave no further improvement and increasing the maximum flux had little influence on the result. All our

tests show that the foreground removal method can work well with the angular resolution we assumed with the $4^{\rm th}$ order polynomial. The effectiveness of this method is weakly dependent on the flux of the extragalactic sources.

In conclusion, polynomial fitting can effectively remove the foreground in the frequency range $100 \sim 200\,{\rm MHz}$ where the foreground can be well approximated by the fitting function. However, in real observation, the foreground is more complex, thus more details need to be considered, such as the instrument response, polarization response, and other non-astrophysical factors. Furthermore, the process of spectral fitting will also introduce errors due to uncertainty in model parameter estimates.

Acknowledgements I would like to thank Xiang-Ping Wu, Quan Guo, Huan-Yuan Shan and Xiao-Chun Mao for helpful comments and discussions. I also wish to thank Hui-Hai He for a careful reading of the draft manuscript.

References

Ali, S. S., Bharadwaj, S., & Chengalur, J. N. 2008, MNRAS, 385, 2166
Cen, R. 2003, ApJ, 591, L5
Chen, X., & Miralda-Escudé, J. 2004, ApJ, 602, 1
Ciardi, B., & Madau, P. 2003, ApJ, 596, 1
Di Matteo, T., Ciardi, B., & Miniati, F. 2004, MNRAS, 355, 1053
Di Matteo, T., Perna, R., Abel, T., et al. 2002, ApJ, 564, 576
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Furlanetto, S. R., Sokasian, A., & Hernquist, L. 2004a, MNRAS, 347, 187
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004b, ApJ, 613, 16
Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581
Holder, G. P., Haiman, Z., Kaplinghat, M., et al. 2003, ApJ, 595, 13
Iliev, I. T., Shapiro, P. R., Ferrara, A., et al. 2002, ApJ, 572, L123
Iliev, I. T., Scannapieco, E., Martel, H., et al. 2003, MNRAS, 341, 81
Jelić, V., Zaroubi, S., Labropoulos, P., et al. 2008, astro-ph/0804.1130
Lawson, K. D., Mayer, C. J., Osborne, J. L., et al. 1987, MNRAS, 225, 307
Kumar, A., Padmanabhan, T., & Surbramanian, K. 1995, MNRAS, 272, 544
Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
Morales, M. F., & Hewitt, J. N. 2004, ApJ, 615, 7
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Peacock, J. A. 1999, Cosmological Physics (1st, Cambridge: Cambridge Univ. Press)
Pen, U.-L., Wu, X.-P., & Peterson, J. 2004, astro-ph/0404083
Pierpaoli, E. 2003, ApJ, 589, 58
Platania, P., Bensadoun, M., Bersanelli, M., et al. 1998, ApJ, 505, 473
Rohlfs, K. 1986, Tools of Radio Astronomy (1st, Berlin: Springer)
Santos, M. G., Cooray, A., Haiman, Z., et al. 2003, ApJ, 598, 756
Santos, M. G., Cooray, A., & Knox, L. 2005, ApJ, 625, 575
Shaver, P. A., Windhorst, R. A., Madau, P., et al. 1999, A&A, 345, 380
Tegmark, M., Eisenstein, D. J., Hu, W., et al. 2000, ApJ, 530, 133
Tegmark, M. 1997, ApJ, 480, L87
Tozzi, P., Madau, P., Meiksin, A., et al. 2000, ApJ, 528, 597
Venkatesan, A., Giroux, M. L., & Shull, J. M. 2001, ApJ, 563, 1
Wang, X.-M., Tegmark, M., Santos, M. G., et al. 2006, ApJ, 650, 529
Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, ApJ, 608, 622