FOREGROUND CONTAMINATION IN INTERFEROMETRIC MEASUREMENTS OF THE REDSHIFTED 21 cm POWER SPECTRUM

Judd D. Bowman[1,4], Miguel F. Morales[2], and Jacqueline N. Hewitt[3]
[1] California Institute of Technology, Pasadena, CA 91125, USA; jdbowman@caltech.edu
[2] University of Washington, Seattle, WA 98195, USA
[3] Massachusetts Institute of Technology, Kavli Institute for Astrophysics and Space Research, Cambridge, Massachusetts, USA

ABSTRACT

Subtraction of astrophysical foreground contamination from "dirty" sky maps produced by simulated measurements of the Murchison Widefield Array (MWA) has been performed by fitting a third-order polynomial along the spectral dimension of each pixel in the data cubes. The simulations are the first to include the unavoidable instrumental effects of the frequency-dependent primary antenna beams and synthesized array beams. They recover the one-dimensional spherically binned input redshifted 21 cm power spectrum within $\sim 1\%$ over the scales probed most sensitively by the MWA ($0.01 \lesssim k \lesssim 1$ Mpc^{-1}) and demonstrate that realistic instrumental effects will not mask the epoch of reionization signal. We find that the weighting function used to produce the dirty sky maps from the gridded visibility measurements is important to the success of the technique. Uniform weighting of the visibility measurements produces the best results, whereas natural weighting significantly worsens the foreground subtraction by coupling structure in the density of the visibility measurements to spectral structure in the dirty sky map data cube. The extremely dense uv-coverage of the MWA was found to be advantageous for this technique and produced very good results on scales corresponding to $|u| \lesssim 500\lambda$ in the uv-plane without any selective editing of the uv-coverage.

Key words: early universe – intergalactic medium – methods: data analysis – radio lines: general – techniques: interferometric

Online-only material: color figures

1. INTRODUCTION

Astrophysical foreground contaminants are 5 orders of magnitude brighter than the ~ 10 mK redshifted 21 cm emission expected from neutral hydrogen in the intergalactic medium (IGM) during the epoch of reionization (EoR). These foregrounds will severely complicate planned experiments and the interpretation of their results. Galactic synchrotron radiation dominates the sky at radio frequencies near 150 MHz ($z \approx 8$), accounting for $\gtrsim 70\%$ of the 200–10,000 K total brightness temperature (Shaver et al. 1999). Extragalactic continuum point sources are also especially strong and numerous, presenting a sea of confused point sources and comprising the bulk of the remaining $\sim 30\%$ of the sky brightness temperature. Galactic radio recombination lines (RRLs) and free-free emission from electrons in both the Galaxy and the IGM will additionally complicate the planned measurements.

Initial analyses indicated that these foregrounds were an insurmountable obstacle (Di Matteo et al. 2002; Oh & Mack 2003) because their angular variance dominates the expected fluctuations in the redshifted 21 cm background, but subsequent studies have suggested that multifrequency observations and the application of appropriate statistical techniques should provide methods to separate the foregrounds from the redshifted 21 cm signal by exploiting the large coherence of the foregrounds with frequency (Di Matteo et al. 2004; Zaldarriaga et al. 2004; Morales & Hewitt 2004; Furlanetto & Briggs 2004; Gnedin & Shaver 2004; Santos et al. 2005; Wang et al. 2006; McQuinn et al. 2006; Gleser et al. 2008). However, these studies have relied on a number of simplifying assumptions and, in particular, have neglected complications due to the frequency-dependent

instrumental response. The point-spread function (PSF) and instrumental field of view (FOV) are both frequency dependent, and can mix the angular structure that concerned Di Matteo et al. (2002) and Oh & Mack (2003) into the frequency direction, masking the EoR signal. This transfer of power from the angular to frequency dimensions has been dubbed "mode-mixing," and is a significant concern for redshifted 21 cm measurements.

In this paper, we explicitly model the mode-mixing effect for confusion-level contaminants and explore ways for minimizing this power transfer into the frequency domain. With these new techniques we show that, for the Murchison Widefield Array (MWA), contamination due to mode-mixing can be reduced well below that of the expected 21 cm signal strength during the reionization epoch for a reasonable foreground model. We begin in Section 2 by defining our instrument and sky models and establishing the mathematical foundation of the subtraction technique. In Section 3, we analyze the ability of the method to produce a model of the confusion-level foreground contamination and, in Section 4, we show that the input 21 cm power spectrum in our simulations can be recovered within $\sim 1\%$ (excluding thermal noise uncertainty) over most scales in the planned MWA measurements. We conclude in Section 5 with a brief discussion and an analytic approximation to provide context to the results of the simulation.

A detailed study of the dependence of this foreground-subtraction technique on the properties of the instrument and astrophysical sky has been conducted in parallel by Liu et al. (2009) for cases building on our fiducial MWA measurement framework. In addition, Jelić et al. (2008) have recently developed a sophisticated astrophysical foreground model and applied it to analysis of subtraction techniques in the context of the expected antenna configuration of the Low Frequency Array (LOFAR). Their results indicate that LOFAR should be

---

[4] Hubble Fellow.

able to constrain the reionization history over a wide range of redshifts by measuring the variance in image maps from individual spectral channels between 100 and 200 MHz. Our present focus is on foreground subtraction in the context of three-dimensional power spectrum measurements including the full frequency-dependent instrumental properties of the MWA, but many of the results of all three investigations are relevant to both measurements and instruments, as well as other active redshifted 21 cm experiments including the Precision Array to Probe the Epoch of Reionization (PAPER), the reionization project with the Giant Metrewave Radio Telescope (GMRT), and the future Square Kilometer Array (SKA).

2. STATEMENT OF THE PROBLEM

Galactic synchrotron radiation, extragalactic continuum sources, and free-free emission dominate brightness temperature maps of the low-frequency radio sky and yield diffuse angular structure that is several orders of magnitude more intense than the expected redshifted 21 cm background from the reionization epoch and Dark Ages. The fundamental problem faced by all redshifted 21 cm experiments, therefore, is how to effectively isolate the desired signal at high precision from these nuisance foregrounds. The strategies discussed in the literature for removing the foreground contamination all exploit a single, common spectral property that distinguishes the foregrounds from the expected signal. Along any one line of sight, each of the foreground components has a slowly varying power-law like spectrum. This results in a large spectral coherence scale and is in contrast to the redshifted 21 cm signal from the reionization epoch, which fluctuates relatively rapidly in all three spatial dimensions, and thus has a short coherence scale, both in frequency and angle (Morales & Hewitt 2004). In general, the spatial coherence length of the reionization signal is of order 10 Mpc, which translates to subdegree-scale fluctuations on the plane of the sky and sub-MHz fluctuations in frequency. Thus, although the angular fluctuations in the foreground intensity span the scales of interest for the 21 cm signal, the frequency fluctuations are on much larger scales than the desired signal.

In the image domain, the "per pixel" technique of subtracting a smooth spectral component from each pixel in a data cube should be able to effectively remove the foreground contribution (Furlanetto & Briggs 2004; Santos et al. 2005; Wang et al. 2006; McQuinn et al. 2006) and only sacrifice a few of the largest modes of the redshifted 21 cm fluctuations. A closely related approach is to produce internal linear combinations (ILCs) of maps from different spectral channels (e.g., difference maps) in analogy with the removal of Galactic emission from CMB measurements (Gnedin & Shaver 2004; Di Matteo et al. 2004). Unlike in CMB studies, however, template fitting of specific foreground components is not a viable approach for redshifted 21 cm foreground subtraction due to the extremely high level of accuracy needed and the scarcity of suitable templates at the target frequencies. In the Fourier domain, one can also fit low-order polynomials or derive ILCs to aid in removing foreground emission, and this has certain advantages for instruments with sparse u, v coverage (Zaldarriaga et al. 2004; U. Pen 2008, private communication). In all of these approaches and the additional statistical techniques discussed in Morales et al. (2006), the angular 21 cm fluctuations are effectively ignored in favor of the line-of-sight fluctuations.

There is a significant obstacle in implementing the techniques listed above in real-world applications: inverting the instrumental response matrix for the MWA and other low-frequency arrays

will not be possible. Thus, the effects of their frequency-dependent PSFs and FOVs will remain in the final data products and mix the strong angular foreground fluctuations into the frequency domain. This instrumental effect can produce sufficient spectral structure in the derived brightness temperature maps to destroy the large spectral coherence of the foregrounds and permanently mask the 21 cm signal. Even a modest observational bandwidth of 30 MHz at 150 MHz ($z \approx 8$) represents a fractional bandwidth of 20%. Thus, the array appears 20% larger in the uv-plane at the highest frequencies than the lowest, causing the PSF in uncorrected maps to change in angular size by 20% from the one end of the frequency band to the other. The FOV is similarly 20% larger at the bottom of the band than the top.[5] Consequently, the angular scale of ripples and peaks introduced into the image by the instrumental PSF changes with frequency, causing the features to appear in different locations at different frequencies. This is the mechanism that couples angular fluctuations into the frequency domain and is shown graphically in Figure 1. Similarly, the varying FOV significantly changes the input sky signal received at different frequencies. This is particularly important for Fourier domain-subtraction techniques, where a 20% change in FOV implies that $\sim 40\%$ of the signal is not in common for a visibility measured at the same point in the uv-plane, but at opposite ends of the band.

In order to study the effect of mode-mixing, we need to carefully simulate the instrument, foreground contamination, and analysis pipeline. In the following sections, we describe the simulation parameters used for this study.

2.1. Epoch of Reionization Observations with the MWA

The fiducial MWA design is described in Bowman et al. (2006, their Section 2). The array design consists of $N = 500$ antennas distributed within a $D = 1500$ m diameter circle. The density of antennas as a function of radius is taken to go as $\sim r^{-2}$, but capped at a maximum density of one antenna per 36 m^2. The antenna response is approximated by

$$W(\theta) = \cos^2\left(\frac{\pi}{2}\theta/\Theta_P\right), \qquad \theta < \Theta_P, \qquad (1)$$

where Θ_P is proportional to wavelength and is 31° at $\nu = 158$ MHz. The angular resolution of the array is given by λ/D and the total collecting area by $N dA$, where dA is the collecting area of each antenna and scales like $dA = 16(\lambda^2/4)$ for $\lambda < 2.1$ m and is capped for longer wavelengths. Finally, the full instantaneous bandwidth of the instrument is $B = 32$ MHz and the spectral resolution is 10 kHz, although the EoR observations will be binned to $\geqslant 40$ kHz resolution to reduce the data volume. All of the fiducial properties are summarized in Table 1. For the analysis in this paper, we define the observation to be of a single field with 360 hr integration during the most favorable circumstances. Additionally, we set the frequency coverage to $142 < \nu < 174$ MHz, which spans $9 > z > 7.1$.

The foreground-subtraction strategy planned for redshifted 21 cm measurements with the MWA is a multistage process consisting of three primary components:

1. *Bright source subtraction.* The first step is to subtract individual bright sources using a full deconvolution approach that incorporates the position, frequency, and

---

[5] One can design antennas with a frequency-dependent collecting area to counteract the effect of the aperture becoming smaller in wavelengths and produce a near constant FOV. However, for a number of technical reasons, none of the upcoming arrays use this kind of detector element.
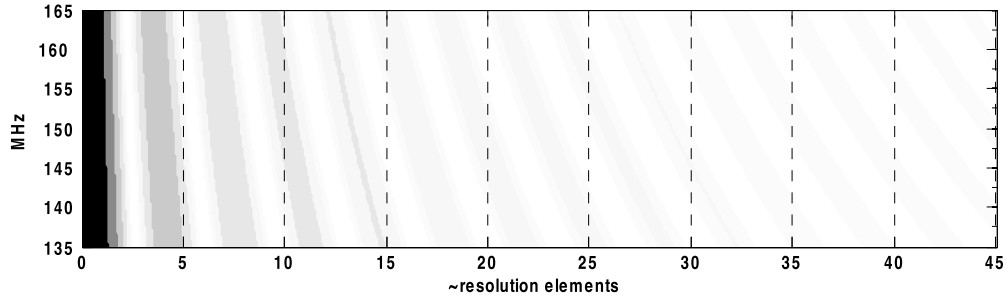
Figure 1. Cartoon of a frequency vs. position plot to illustrate the origin of mode-mixing from a frequency dependent PSF. At the origin of the horizontal axis a source has been mis-subtracted, and the residual flux ripples across the image due to the array's PSF. Since the PSF scales with frequency, the positions of the radial intensity peaks change with frequency resulting in diagonal frequency bands. At lines of sight away from the mis-subtracted source, this leads to frequency-dependent contamination, as indicated by the contribution along the vertical dashed lines. The frequency-dependent PSF is inherent in the measurement, and mixes a pure spatial component (mis-subtracted sources) into the frequency direction. In addition, as the frequency slope of the contamination depends on the distance to the residual sources, a wide range of line-of-sight k-modes are contaminated.
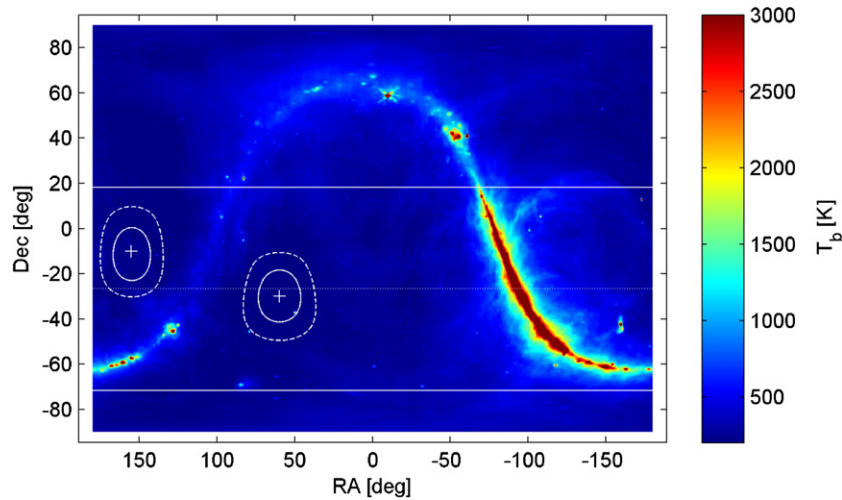

Figure 2. Brightness temperature of the Galactic synchrotron foreground at $\nu = 150$ MHz. The all-sky measurements of Haslam et al. (1982) were scaled from 408 MHz using a constant spectral index of $\beta = 2.6$ to produce this map. The declination of the zenith at the MWA latitude is shown (the horizontal dotted line), as is the declination range within zenith angle of $45°$ (horizontal solid lines). The MWA will have access to cold regions (~ 250 K) of the sky both north and south of the Galactic plane. The primary beam of the MWA antenna tiles is shown centered at the two planned target fields, marked by plus symbols at 60 R.A. by -30 decl. and 115 R.A. by 10 decl. The concentric circles indicate the 50% (solid) and 10% (dashed) power response relative to the peak response.

(A color version of this figure is available in the online journal.)

Table 1
MWA Design Specifications

Parameter	Value
Array layout, $\rho(r)$ (m^{-2})	$\sim r^{-2}$
Array diameter, D (m)	1500
Array latitude ($°$)	-27
Bandwidth, B (MHz)	32
Spectral resolution (kHz)	40
Number of antennas, N	500
Angular resolution, Θ_B ($°$)	0.073
Antenna collecting area, dA (m^2)	14.4
Antenna response scale, Θ_P ($°$)	31
Central frequency, ν_0 (MHz)	158
System temperature, $T_{\rm sys}$ (K)	440

Note. Parameters are listed with their fiducial values at $\nu = 158$ MHz, corresponding to $z = 8$ for redshifted 21 cm measurements.

time-dependent antenna calibrations to remove the sources to high precision. The MWA has been designed to aid in the mitigation of extremely intense extragalactic and Galactic continuum sources. It is common practice in radio astronomy to "peel" away bright sources in the FOV to improve the overall dynamic range in derived maps of the sky brightness and, thus, to increase the sensitivity to fainter signals. This process must be applied to the bright point sources in measurements by the MWA to a high degree of precision (better than 1 part in 10^5) in order to reveal the fluctuations in the redshifted 21 cm background. For typical interferometers, the sparse coverage of visibility measurements in the uv-plane means that a substantial amount of power from these point sources is spread across the derived map of the sky due to the sidelobes of the PSF of the synthesized beam. This poses a considerable deconvolution problem and makes it difficult to subtract them from the measurements since the locations and intensities of the point sources must be determined very accurately before their contributions to the individual visibility measurements can be removed to high precision. The large number of antennas used for the MWA reduces the sidelobes of the synthesized beam and provides high angular dynamic range in uncorrected maps of the sky. This is advantageous for isolating and subtracting the power from extremely intense extragalactic continuum sources and alleviates the severity of the bright source contamination.

2. *Diffuse spectral subtraction.* The next step is to remove diffuse emission from the Galaxy and faint confusion-level sources by fitting a polynomial to the spectrum along each line of sight in the observed sky maps in order to subtract the spectrally smooth foreground component. To aid this process, the angular size of the primary antenna beam (FOV) for the MWA has been matched to fit within the relatively cold regions that exist in the Galactic synchrotron emission at high Galactic latitudes. Figure 2 illustrates the typical brightness temperature of the synchrotron foreground at 150 MHz, as well as the primary field of view for an MWA antenna tile that is targeting a cold part of the sky.

3. *Statistical template subtraction.* The final step in the foreground-subtraction process is envisioned to take place as part of the 21 cm power spectrum analysis by fitting templates of the residual statistical structures of foregrounds that may remain even after the first-order subtraction of their emission from sky maps.

In the analysis that follows, we focus only on the second step: subtracting faint sources and diffuse emission that cannot be identified as individual contributions in the maps and deconvolved efficiently. We do not model the first step of bright source deconvolution, but rather assume it has been performed perfectly. At the end of this paper, we briefly look at the implications from the present analysis for the final planned foreground-subtraction step of statistical template fitting.

2.2. The Sky Model

In this section, we specify the details of the sky model for the astrophysical foregrounds and the redshifted 21 cm signal. We build our foreground sky model adhering closely to existing empirical findings. The underlying physical mechanisms behind the observed properties are not addressed in the development of this model since we are interested primarily in the functional consequences of their manifestations on the sky and in the instrument (for an alternative treatment, see Jelić et al. 2008). We begin by constructing an empirically motivated model of the astrophysical foregrounds that consists of two components: (1) discrete, faint extragalactic continuum sources, and (2) diffuse Galactic synchrotron emission. Together, these are expected to account for approximately 98% of the total intensity in the radio spectrum below 200 MHz (Shaver et al. 1999; Bridle 1967). We exclude free-free emissions as a separate component in our analysis since they have power-law spectra similar to the other components (Kogut et al. 1996), and they are easily subsumed by the uncertainty in the discrete continuum source contribution, although the free-free emissions have been shown to correlate with dust clouds at high Galactic latitudes, and thus have different angular structure than the discrete source population. We have also chosen to ignore supernova relics, radio clusters, and RRLs in this analysis. The first two categories of sources do not differ substantially from the power-law like components we include in the model, and the frequencies at which Galactic RRLs occur are known and can be excised from observations if they are determined to be a significant contribution. Diffuse RRLs have never been observed at high Galactic latitudes, but they are expected to be narrower than the 40 kHz frequency channel size of the MWA and occur approximately every few MHz, and therefore in only a few percent of the spectral channels (Erickson et al. 1995; Shaver 1975; Walmsley & Watson 1982, are good starting points for additional details about the observed properties and theoretical

treatment of RRLs). Thus, the cost of excising the RRLs is a minor complication to the window function, but one that is expected to be no more severe than that caused by excising radio-frequency interference (RFI).

2.2.1. Diffuse Galactic Synchrotron Radiation

For the diffuse Galactic synchrotron foreground, there are a number of observations on which the model is based (Bridle 1967; Landecker & Wielebinski 1970; Haslam et al. 1982; Reich & Reich 1988; Alvarez et al. 1997; Roger et al. 1999). The spectrum of the Galactic synchrotron emission has been found to be a nearly featureless power law with modest variations in the spectral index as a function of direction and frequency. At the frequencies of interest, the spectral index is given by $T \sim \nu^{-\beta}$ and has a typical value of $\beta \approx 2.5$ (Rogers & Bowman 2008). In general, it is steeper at high Galactic latitudes than toward the Galactic plane. The spectral index also steepens as a function of frequency to a maximum of $\beta \approx 2.8$ by $\nu \approx 1$ GHz, but is generally constant below about 200 MHz. The variation of the spectral index across the sky has a standard deviation of order $\sigma_\beta \approx 0.1$ on degree scales (Bridle 1967; Reich & Reich 1988; Platania et al. 1998; Roger et al. 1999) and appears to be weakly correlated to the angular structure (Reich & Reich 1986; Lawson et al. 1987; Platania et al. 2003). Angular structure in the diffuse Galactic emission has been shown to be well described by a power-law spectrum in Fourier space over a large range of scales. The angular power-law index is specified according to $C_\ell \sim \ell^{-\alpha}$, where typically $2.4 \lesssim \alpha \lesssim 2.9$ (Platania et al. 1998; Giardino et al. 2001, 2002; La Porta et al. 2008), and the spherical harmonic multipole is related to the uv-plane baseline length by $\ell \approx 2\pi u$.

We construct our model of the diffuse Galactic synchrotron radiation in two steps. First, we produce a Gaussian random field with angular power-law index of $\alpha = 2.55$. The amplitude of this field is normalized to both the mean and angular fluctuation power in the region of the all-sky map of Haslam et al. (1982) centered on R.A. $= 60°$ and decl. $= -30°$. This region has been chosen as the primary target field for the MWA due to its relatively low brightness temperature and its location roughly opposite the Galactic center. In order to produce appropriate amplitude normalizations, the temperatures in the Haslam et al. (1982) map are scaled from 408 MHz to 150 MHz using a constant spectral index of $\beta = 2.6$. The results of this scaling are shown in Figure 2. Since the Haslam et al. (1982) map is lacking angular information on scales smaller than about 1°, only multipoles below $\ell < 200$ were used to normalize the angular fluctuation power in our model. Figure 3 illustrates the angular power spectrum of our model at 150 MHz, as well as the power spectrum derived from the Haslam et al. (1982) map. To complete our synchrotron model, the second step is to assign a spectrum for each line of sight in the simulated map. We specified each spectrum as a power law with spectral index drawn from a second Gaussian random field. The spectral index field has $\bar{\beta} = 2.6$ and angular structure given by the sum of two terms: (1) a white noise contribution with $\sigma_\beta = 0.1$ and (2) the simulated brightness temperature map at 150 MHz renormalized such that it yields an additional variance in the spectral index map of $\sigma_\beta = 0.05$. Thus, the spectral index map is weakly correlated with the brightness temperature map. The level of correlation was set by eye from inspection of derived brightness temperature and spectral index maps including, in particular, the maps of Platania et al. (2003), although we note that those results were found at higher frequencies.
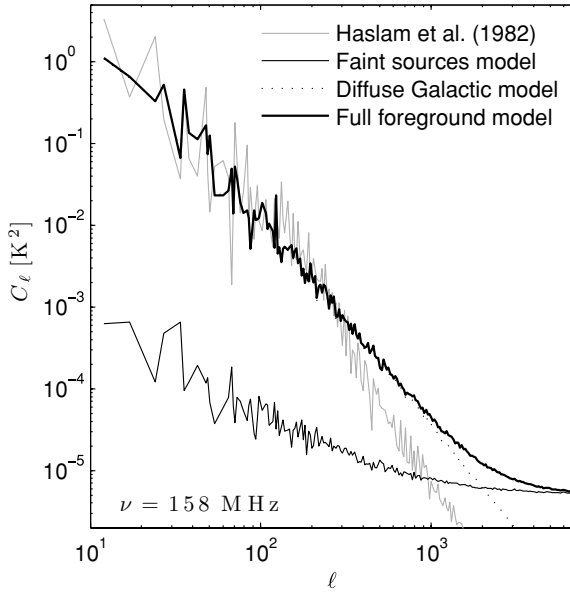
Figure 3. Angular power spectra of the two model foreground components along with the power spectrum derived from the Haslam et al. (1982) map (scaled to 150 MHz) over the region spanned by the MWA primary target field. The diffuse Galactic synchrotron emission dominates the foreground model at large scales (low ℓ), where its angular fluctuation amplitude and power-law index have been matched to the Haslam et al. (1982) map. For small, subdegree scales ($\ell \gtrsim 200$), the power in the Haslam et al. (1982) map drops off rapidly due to its limited angular resolution, whereas the Galactic component of the foreground model continues its pure $\alpha = 2.55$ power-law profile. At very small scales ($\ell \gtrsim 1000$), the angular power of the extragalactic continuum source model becomes comparable to the diffuse Galactic power since the continuum source fluctuations have a Poisson noise floor.

2.2.2. Extragalactic Continuum Sources

Normal galaxies, radio galaxies, and active galactic nuclei form the majority of the extragalactic continuum sources (Santos et al. 2005). Their contribution to the brightness temperature has been found to be of order 30–70 K at 178 MHz by Bridle (1967) by fitting an isotropic background component to measurements of spatial variations in the spectral index of the diffuse emission between 18 and 404 MHz. In the same analysis, it was found that the integrated brightness temperature due to extragalactic sources has a characteristic spectral index of $\beta = 2.7$, although with $\sim 10\%$ uncertainty, and therefore is somewhat steeper than the Galactic synchrotron spectral index. Individual sources have been observed to have significantly more variation in their spectral indices, however, spanning $-2 \lesssim \beta \lesssim 3$. The inverted cases ($\beta \approx -2$) are due to sources with strong synchrotron self-absorption, such that the spectrum has already flattened at the frequencies of interest (as opposed to flattening below ~ 10 MHz, as would be more typical). In general, the spectral index of discrete extragalactic sources is related to the intensity of the source. It is steeper for brighter sources (with maximum $\beta \approx 3$) than for fainter sources, where $\beta \approx 2.5$ is typical for sources with fluxes less than $S < 0.1$ Jy (Kellermann et al. 1969; Zhang et al. 2003; Cohen et al. 2004).

A number of surveys of radio sources have been performed at frequencies relevant to the MWA (see Cohen et al. 2004, their Figure 1). They include a survey at 74 MHz by Cohen et al. (2004), the 7C (McGilchrist et al. 1990) and 6C surveys (Hales et al. 1988) at 151 MHz, the 3CR survey (Laing et al. 1983) of bright (> 10 Jy) sources at 178 MHz, WENSS (Rengelink et al. 1997) at 327 MHz, and the 5C5 survey (Pearson 1975) at 408 MHz, as well as the FIRST survey (White

et al. 1997) and the NRAO VLA Sky Survey (NVSS) (Condon et al. 1998) at 1.4 GHz. An analysis of the catalogs of radio sources derived from these and other surveys has yielded a solid understanding of the statistical properties of radio sources (Perley & Erickson 1984; Windhorst et al. 1985; Wieringa 1991; Fomalont et al. 1991; Windhorst et al. 1993). The brightest sources in the catalogs are relatively nearby objects and, thus, have differential source counts given by $N(S) \sim S^{-2.5}$, as expected for a population of objects distributed evenly in a Euclidean space. Weaker sources are generally more distant and the observed differential counts fall off more gradually, as $N(S) \sim S^{-1.8}$, due to evolution of the population and redshift effects. At extremely faint flux levels (in the μJy regime), a new population of local starburst galaxies becomes visible and the differential counts again become steeper, with $N(S) \sim S^{-2.2}$. Differential source counts cannot asymptote to this power law as $S \to 0$, however, since it is steeper than $N(S) \sim S^{-2}$ and an infinite radio flux would result. The differential counts must eventually flatten again with decreasing flux.

For our extragalactic continuum source model, we first produce a simulated catalog of radio sources and then draw from that catalog to sum over individual sources within each pixel of a simulated sky map. We use the expression for differential source counts given by Subrahmanyan & Ekers (2002),

$$N(S) = 100 S_{\mu\text{Jy}}^{-2.2} \nu_{\text{GHz}}^{-0.8} \, \text{arcmin}^{-2} \mu\text{Jy}^{-1}, \qquad (2)$$

where $S_{\mu\text{Jy}}$ is the flux measured in μJy and ν_{GHz} is the frequency measured in GHz. Based on this equation, we estimate that it should be possible to identify individual sources in sky maps produced by the MWA down to a flux limit of approximately 10 mJy before the sources become confused (in other words, there will be more than approximately one source per pixel below 10 mJy and, hence, individual sources will not be identifiable). Since the integrated flux due to faint radio sources in Equation (2) converges around 10 μJy, and since we want to limit the number of sources in our simulated catalog, we only model sources between 10 μJy $< S <$ 10 mJy. For each source in our simulated catalog, we randomly assign a flux such that Equation (2) is satisfied. We also randomly assign a spectral index for each source from a distribution of spectral indices that is the sum of two Gaussian distributions. One distribution has $\bar{\beta} = 2.7$ and $\sigma_\beta = 0.1$, and the second distribution has $\bar{\beta} = 0$ and $\sigma_\beta = 0.6$ in order to account the large variation of spectral indices in extragalactic continuum sources. Only 10% of the sources have their spectral indices drawn from the second distribution.

In order to produce a realistic sky map from our simulated source catalog, the angular correlation of radio sources must be taken into account. Blake & Wall (2002) analyzed the 1.4 GHz NVSS and found evidence for weak clustering of the continuum point sources. The angular correlation function that they derived is

$$w(\theta) = 10^{-3} \theta^{-0.8}, \qquad (3)$$

for θ measured in degrees. The magnitude of this clustering is much lower than for the galaxy correlation functions determined optically, and is the result of projection effects. Deep radio catalogs contain sources spanning a large range of redshifts since their continuum spectra do not allow for grouping by photometric redshift measurements. This causes the angular correlation function of radio continuum sources to dilute since unrelated volumes of the universe are superimposed in the

observations. We follow the method of González-Nuevo et al. (2005) to make a realization of the source counts in pixels that includes both the Poisson variance and the proper angular clustering. The large-scale structure introduced by the angular clustering adds an additional variance to the distribution of point sources and the resulting brightness temperature in the sky map. Using this source count map, we populate each line of sight in the extragalactic continuum source foreground model by drawing the appropriate number of sources from our simulated catalog and summing their contributions. Figure 3 shows the angular power spectrum resulting from this process at 150 MHz.

2.2.3. Redshifted 21 cm Signal

Since the purpose of this work is to study the effects of a realistic interferometer on foreground subtraction, we employ a simple model for the redshifted 21 cm signal. Averaging over velocity field distortions, the brightness temperature of diffuse redshifted 21 cm emission can be described by

$$\delta T_{21}(\vec{\theta}, z) \approx 23(1 + \delta)x_{\mathrm{HI}}\left(1 - \frac{T_{\mathrm{CMB}}}{T_S}\right)$$
$$\times \left(\frac{\Omega_b h^2}{0.02}\right)\left[\left(\frac{0.15}{\Omega_m h^2}\right)\left(\frac{1+z}{10}\right)\right]^{1/2} \mathrm{mK}, \quad (4)$$

where δ is the matter density perturbation field, x_{HI} is the neutral fraction of hydrogen, $T_{\mathrm{CMB}} = 2.725$ K is the CMB temperature, and T_S is the spin temperature describing the relative population of the 21 cm hyperfine states. For our simple model, we assume that the IGM remains fully neutral at our target redshift of $z = 8$ and that the spin temperature of neutral hydrogen is $T_S \gg T_{\mathrm{CMB}}$. The only fluctuations in the 21 cm brightness temperature, therefore, are due to matter density perturbations, which we model as a Gaussian random field with a power spectrum given by the output of CMBFAST (Seljak & Zaldarriaga 1996) and scaled to the target redshift according to the linear growth factor.

2.3. Instrumental Simulation

We start with our model of the frequency-dependent sky brightness $I(\{\vec{\theta}, \nu\})$, as detailed in Section 2.2. For the remainder of this section, we will drop the explicit frequency dependence of all the terms to make the notation clearer. The result of observing the true sky with the MWA is a vector **m** of complex visibility measurements. This process can be expressed as

$$\mathbf{m}(v) = \mathbf{M}(\vec{\theta}, \vec{\theta})I(\vec{\theta}) + n(v), \quad (5)$$

where v indexes the individual measurements, **M** is the instrumental response matrix that maps the true sky to the resulting measurements, and n is a vector containing the system noise added to each measurement.

In observations with the MWA, the visibility data must be calibrated and compressed so that long integrations can be archived for later analysis, and this process is detailed in Morales & Matejek (2008) and Mitchell et al. (2008). Assuming perfect calibration, the resulting images and Fourier representations can be simulated as

$$I'(\vec{\theta}) = \mathbf{B}(\vec{\theta}, \vec{\theta})\,I(\vec{\theta}) + n(\vec{\theta}), \quad (6)$$

$$I'(\mathbf{u}) = \mathbf{B}(\mathbf{u}, \mathbf{u})\,I(\mathbf{u}) + n(\mathbf{u}). \quad (7)$$

Equations (6) and (7) are Fourier transforms of one another and the operator **B** represents the frequency-dependent array

beam and antenna field of view, and I is the true sky in either angular or transformed coordinates. The I' in the first line is the standard interferometric dirty map, and the gridded uv-plane in the second line. Depending on the application we will alternate between these two forms of the equation. The first line is useful when fitting spectra to individual image pixels, but has the disadvantage of a highly covariant noise matrix $N = n^T(\vec{\theta})n(\vec{\theta})$. The second line has (to good approximation) a diagonal noise matrix that is simply related to the density of visibility measurements in the uv-plane.

Relating these expressions to the polynomial-based foreground-subtraction investigations of Furlanetto & Briggs (2004), Wang et al. (2006), and Gleser et al. (2008), we find that these works treated simplified cases of this general approach by assuming complete coverage of the visibility measurements in the uv-plane, equivalent to using $\mathbf{B} = \mathbf{I}$, where **I** is the identity matrix, and by setting $n(\mathbf{u}) = 0$ or assuming $n(\mathbf{u})$ to be constant so that the noise covariance matrix, N, is diagonal in the image domain, even though this is not attainable for a realistic distribution of antenna elements. Jelić et al. (2008) have improved on these initial approximations by including the planned coverage of visibility measurements in the uv-plane for LOFAR, although the frequency dependence of their uv-coverage is specifically removed by excising baselines lacking corresponding measurements at all frequencies and they neglect the frequency dependence of the primary antenna beam. Here, we extend the growing foundation of these efforts by investigating the full frequency-dependent effects of including a realistic antenna distribution and primary antenna FOV for the MWA to calculate **B** and n. We expand **B** into two primary constituents

$$\mathbf{B} = \mathbf{B}_{\mathrm{array}}\mathbf{B}_{\mathrm{ant}}, \quad (8)$$

where $\mathbf{B}_{\mathrm{array}}$ accounts for the synthesized array beam due to the coverage of the visibility measurements in the uv-plane, and $\mathbf{B}_{\mathrm{ant}}$ accounts for the response of the primary beam of the antenna tiles. For $\mathbf{B}_{\mathrm{ant}}$ we use the window function in Equation (1), and for $\mathbf{B}_{\mathrm{array}}$ we compute the typical uv-coverage of an MWA observation including Earth-rotation synthesis for a 90° track through zenith as shown in Figure 4.

For the noise contribution to the simulated observations, we take $n(v)$ to be a vector of Gaussian random variables with standard deviation (Morales 2005)

$$\sigma_n = \frac{2k_B T_{\mathrm{sys}}}{dA\sqrt{d\nu\tau}}, \quad (9)$$

where k_B is the Boltzmann constant, $T_{\mathrm{sys}} = 440$ K is a conservative estimate of the system temperature (including sky noise) of the MWA at $\nu = 158$ MHz, $dA = 14.4$ m² is the effective collecting area of an MWA antenna tile at the same frequency, $d\nu = 40$ kHz is the spectral resolution of a single channel, and $\tau = 8$ s is the accumulation duration for each visibility measurement. We approximate $n(\mathbf{u})$ by pixelizing the uv-plane with cells equal in area to the inverse of the antenna FOV. The contribution of any single visibility measurement is then applied to only one grid cell. In this limit, the thermal noise is uncorrelated in the Fourier domain maps and given by, $n(\mathbf{u}) = \sigma_n/\sqrt{N(\mathbf{u})}$, where $N(\mathbf{u})$ is the number of visibility measurements contributing to each grid cell. Using the radiometer equation with the mean filling fraction for the MWA antenna configuration gives a fiducial thermal uncertainty for a derived image map of ~ 335 mK after 360 hr, although in practice the thermal noise is dependent on angular scale size
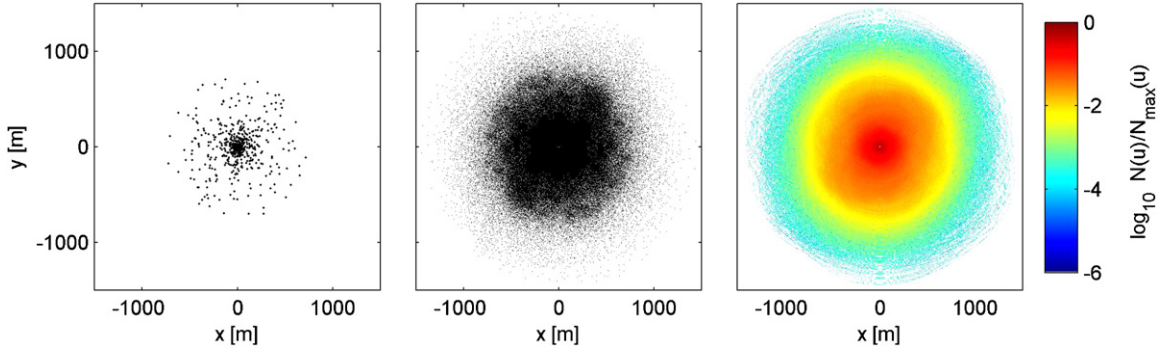
Figure 4. Realization of the MWA antenna layout (left) and corresponding distribution of baselines (middle). The 500 antennas are nearly coplanar and distributed over a 1500 m diameter region with a density that falls off as $\sim r^{-2}$. The baselines in the uv-plane are formed by the pairwise combination of all the antennas for an instantaneous observation of a target at the zenith. The right panel illustrates the relative density of visibility measurements after Earth-rotation synthesis has been performed by tracking a target field at R.A. $= -60°$ and decl. $= -30°$ for 6 hr as it transits.
(A color version of this figure is available in the online journal.)

and improves rapidly as scale size increases due to the highly condensed antenna distribution such that at degree scales the thermal uncertainty is only ~ 5 mK.

The final component of modeling the foreground subtraction is to actually fit and subtract the foreground model from the sky brightness map derived from the observations. There are several reasonable methods that can be used to derive the sky brightness from the archived measurements. The most ideal map to use would be a minimally biased estimate of the true sky. This would require inverting the instrumental response matrix **B** and the noise covariance matrix (Tegmark et al. 1997). In practice, it will not be feasible to invert **B**, and our goal is to set a worst-case limit due to the effects of the instrumental response of the MWA on subtracting foreground contamination, and a minimally biased estimate of the sky should remove many of these effects. Thus, we turn, instead, to what is commonly called visibility weighting of the dirty sky map in radio astronomy. This is the map that results by simply multiplying the uv representation by a position-dependent weight, or equivalently Wiener filtering. The weighted dirty sky map is given by

$$I'_U(\theta) = \mathbf{F}(\theta, \mathbf{u})\mathbf{U}(\mathbf{u}, \mathbf{u})I'(\mathbf{u}), \tag{10}$$

where **F** is the Fourier transform operator, and **U** is a weighting function for each pixel in the Fourier map (Sault 1984; Briggs et al. 1999; Thompson et al. 2001). The smallest uncertainty in the derived sky map due to thermal noise is achieved by weighting by the inverse of the variance due to the (Gaussian) thermal noise in each pixel in the Fourier domain. This is commonly called natural weighting, and in this formalism corresponds to $\mathbf{U} = \mathbf{B}$ if the noise in each visibility is the same. Although it produces the least uncertainty, this method typically emphasizes the information contained in short baselines because short baselines are more numerous in radio interferometers than long baselines. Thus, the effective resolution of the derived sky brightness map is lower than the λ/D expectation. Another common method, called uniform weighting, alleviates this under-resolution at the expense of introducing more uncertainty into the derived sky map. Uniform weighting, like its name suggests, is constant for each pixel in the Fourier domain. This allows the information in the long baselines to be emphasized and restores the effective resolution of the sky map to the expected level. Figure 5 illustrates the difference in the weighting functions. We will test both of these weighting methods, as well as an approximation to the natural weighting function derived by fitting a double exponential to the
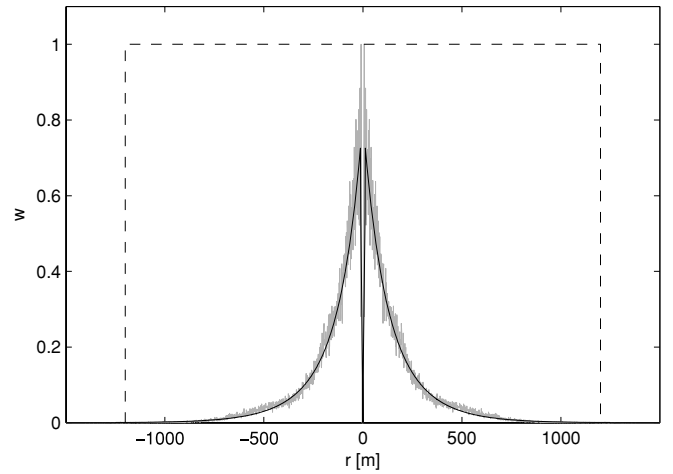

Figure 5. Illustration of three different weighting functions used for forming the dirty sky maps in Figures 7–9. Each profile is for a north–south cross section through the center of the baseline distribution after Earth-rotation synthesis on the target field. The profiles are for uniform weighting (the dashed line), natural weighting (the gray line), and the smoothed double exponential approximation to natural weighting (the black line). The profiles are normalized in this plot such that the maximum weight is one. The natural weighting profiles favor the short baselines, while the uniform weight accentuates the long baselines.

distribution of the number of baselines per uv-cell. We will see if the foreground-subtraction technique is robust to any effects the weighting scheme might introduce.

Once a dirty sky map has been generated, a low-order polynomial is fit to each pixel in the map and subtracted. The residual contamination, $r(\theta)$, after this subtraction is given by

$$r_U(\theta) = I_U(\theta) - d(\theta), \tag{11}$$

where $d(\theta)$ is the "dirty" foreground model generated from the polynomial fits. However, regardless of the weighting used to determine the model $d(\theta)$, we want to subtract this model from the unweighted data to preserve the signal-to-noise ratio (S/N) of the original signal. Leaving the weighting function in the Fourier domain, the residual in the unweighted data is given by

$$\begin{aligned} r'(\theta) &= \mathbf{F}^{-1}(\theta, \mathbf{u})\mathbf{U}^{-1}(\mathbf{u}, \mathbf{u})\mathbf{F}(\mathbf{u}, \theta)r_U(\theta) \\ &= \mathbf{F}^{-1}(\theta, \mathbf{u})\mathbf{U}^{-1}(\mathbf{u}, \mathbf{u})\mathbf{F}(\mathbf{u}, \theta)\left[I_U(\theta) - d(\theta)\right] \\ &= I'(\theta) - \mathbf{F}^{-1}(\theta, \mathbf{u})\mathbf{U}^{-1}(\mathbf{u}, \mathbf{u})\mathbf{F}(\mathbf{u}, \theta)d(\theta) \\ r'(\theta) &= I'(\theta) - d'(\theta), \tag{12} \end{aligned}$$

where $d'(\theta) = \mathbf{F}^{-1}(\theta, \mathbf{u})\mathbf{U}^{-1}(\mathbf{u}, \mathbf{u})\mathbf{F}(\mathbf{u}, \theta)d(\theta)$. The residual, r', in the Fourier domain map is the quantity we are interested in determining, and that, we hope, will yield the expected redshifted 21 cm fluctuations.

3. MEASURING THE CONFUSION-LEVEL FOREGROUND

For all the cases discussed in this section, a third-order polynomial is used to fit the full 32 MHz of observed spectrum along each pixel of the sky map according to

$$d(\vec{\theta}, \nu) = a_3(\vec{\theta})\nu^3 + a_2(\vec{\theta})\nu^2 + a_1(\vec{\theta})\nu + a_0(\vec{\theta}), \qquad (13)$$

where d is the model fit introduced in Equation (11), and the coefficients a_n are solved on a per pixel basis to minimize the least-squared difference between the weighted dirty sky map, I_U, and the polynomial fit. The spectral fits are performed in linear coordinates. Fitting polynomials in logarithmic space has been considered by Wang et al. (2006) and others, but we find that, since the response of the interferometer removes the mean value of each spectral channel and causes the value of some pixels to be less than zero, logarithmic fits pose extra complications.

Before we consider the effects of the instrumental response, we first review the ability of the polynomial fit technique to remove the foreground contributions in the image domain representation of our foreground sky model alone (with no 21 cm signal). This corresponds to the principal scenarios addressed by previous efforts (Di Matteo et al. 2004; Zaldarriaga et al. 2004; Furlanetto & Briggs 2004; Gnedin & Shaver 2004; Santos et al. 2005; Wang et al. 2006; McQuinn et al. 2006; Gleser et al. 2008). Figure 6 illustrates the results of performing the subtraction on the foreground-only sky model with no instrumental response. The top row of the figure shows the input sky data cube. The left panel is a map from a single frequency channel, whereas the right panel is a ν-θ_y plane that slices along the frequency axis. The second row of the figure shows the residuals in the data cube following the polynomial subtraction (again the left panel is a map from a single frequency channel, whereas the right panel is a slice along the frequency axis). Finally, the bottom row of the figure shows the Fourier domain representation of the residuals in the data cube. The left panel in this case is the uv-plane for a single frequency channel and the right panel is a $\nu-\nu$ plane slicing along the frequency direction. The residuals can be seen in the second row to be of order ~ 1 mK. This represents the best-case scenario that can be achieved by the third-order polynomial fit method to our foreground model. In general, the amplitude of the residuals could be made arbitrarily low by increasing the order of the polynomial used in the fits, but only at the expense of extracting additional power from the more rapidly varying spectral fluctuations. Since, in actual observations, the polynomial fit will remove power from the redshifted 21 cm fluctuations as well as the foreground contributions, it is desirable to use the lowest order polynomial feasible because this will limit the effects on the redshifted 21 cm power spectrum to only the longest length scales in the spectral domain. The bottom panel shows that the residual angular power is peaked at large scales (low u). The dark vertical bands in the bottom-right panel are at the zero-crossings of the residual power, which tend to occur at roughly the same frequencies for all sight lines. As concluded by the previous efforts in the literature, it appears that, in an ideal case at least, the polynomial-subtraction technique is sufficient to remove the

foreground contribution in the sky to the level required to detect the expected 21 cm signal during reionization.

Next, we include the instrumental response in the analysis. Now there is an additional degree of freedom. The weighting function, $\mathbf{U}$, that is applied to the simulated gridded measurements before transforming from the Fourier to the image domain (see Equation (10)) must be specified before the calculation can be performed. As discussed in the last section, we model uniform weighting, natural weighting, and a smoothed approximation to the natural weighting found by fitting a double exponential profile to the distribution of baselines (see Figure 5). The results of the foreground-subtraction technique performed on the simulated dirty sky maps for each of these weightings are presented in Figures 7–9, respectively. Figure 10 shows a direct comparison of the dirty map-subtraction residuals for the same line of sight using the three different weighting schemes.

3.1. Uniform Weighting

The first case is that of uniform weighting and is shown in Figure 7. Here, the top row of the figure shows the dirty sky map, $I_D(\theta)$, produced by this weighting, the second row shows the results of the foreground subtraction in the image domain, $r(\theta)$, and the bottom row shows the results of the foreground subtraction after transforming back to the Fourier domain and removing the weighting, $r'(u)$. In all cases, the thermal noise contribution has been artificially removed following the analysis to illustrate the foreground residuals more clearly. As discussed in Section 2.3, this weighting gives the best effective angular resolution, but increases the uncertainty due to thermal noise in the dirty sky map.

At first glance, the results of the foreground subtraction in this case do not look especially promising since there are of order ~ 1 K fluctuations in the residual map, $r(\theta)$, in the second row of Figure 7. However, after the residual map is transformed back to the Fourier domain, it becomes evident that the polynomial fit has actually done an excellent job of subtracting the foreground contamination from baselines within a radius of $u \lesssim 500\lambda$, and only a poor job for baselines beyond this radius. The transition from the region where the foreground subtraction was successful to the region where it failed is rapid and coincides with the radius where visibility measurements with the MWA become sufficiently sparse that there is no longer complete coverage. In this outer region, a given uv-pixel is typically not sampled at all frequencies. This creates variations at the corresponding angular scales between image maps at different frequencies and results in additional spectral structure in the image domain data cube. Thus, on large angular scales, uniform weighting produces a dirty map with an effective PSF, resembling a clean δ-function yielding no corresponding sidelobe confusion since there is complete uv-coverage over all frequencies, whereas for small angular scales, the effective PSF is poorly behaved with large sidelobes. This comes at the expense of elevated thermal noise in the dirty map due to the suboptimal weighting of the uv-plane measurements when producing the map.

The bottom row of Figure 7 illustrates that it may not be necessary to selectively tailor the uv-plane coverage to be frequency independent for the MWA. As long as the scales of interest are sampled completely for all frequencies, the sparsely sampled region does not interfere with the scales of interest following foreground subtraction. Nevertheless, if the extra sidelobe noise in the dirty map due to the sparse coverage above $u \gtrsim 500\lambda$ is not desired, then the long baselines
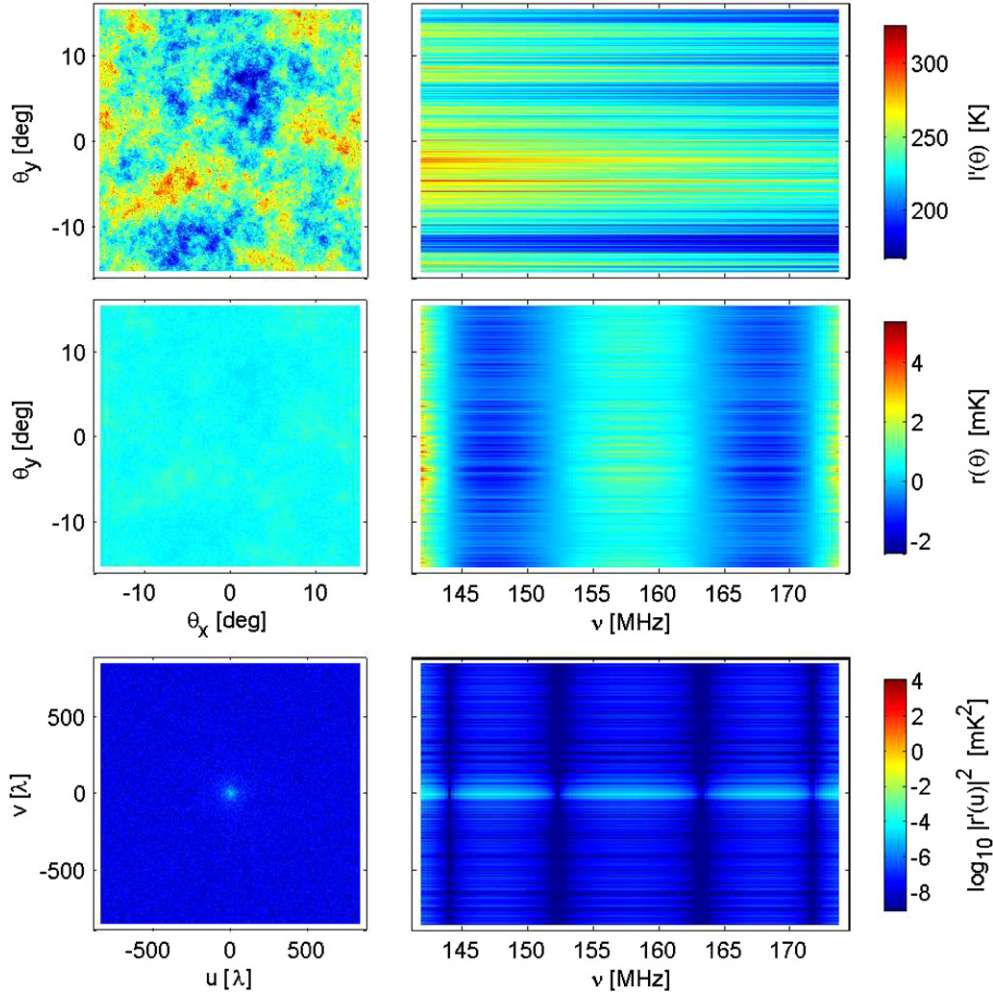
Figure 6. Polynomial foreground-subtraction technique applied to the input foreground sky model *excluding* the simulated instrumental response. Thermal noise and the 21 cm signal have been excluded from this simulation, as well, in order to provide a best-case reference for the foreground-subtraction technique. The first row shows two cuts through the foreground sky model. The left panel is an image plane for a spectral channel at $\nu = 157$ MHz, thus the pixels in the plot are indexed by $I'(\theta_x, \theta_y)$. The right panel is a slice through frequency and one angular dimension, thus the pixels in the plot are indexed through $I'(\nu, \theta_y)$. The smooth spectral properties of the foreground model are easily visible in the right-hand plot. The middle row displays the residuals, r, in the image domain after the third-order polynomial has been fit and subtracted from the spectra along each line of sight in the data cube. Again, the left panel is an image plane and the right panel is a slice through frequency and one angular dimension. The bottom row displays the residuals in the Fourier domain. Although it appears in the middle row that the residuals after foreground subtraction will be comparable to the ~ 25 mK redshifted 21 cm signal, after transforming the residuals back to the Fourier domain (the bottom row), it is evident that the polynomial fit and subtraction removed the foreground contribution well for all but the largest angular scales.

(A color version of this figure is available in the online journal.)

could be completely discarded or selectively cut before producing the dirty map in order to yield a frequency-independent uv-coverage. A variation of this approach was applied in Jelić et al. (2008) for the planned LOFAR configuration.

3.2. Natural Weighting

The foreground-subtraction results for the natural weighting case are shown in Figure 8. It is clear even in the second row of Figure 8 that the residual fluctuations in the dirty map are on much larger angular and spectral scales and have larger amplitudes than for the uniform weighting case, with ~ 10 K fluctuations that are an order of magnitude greater than the expected 21 cm signal. And, in this case, the residuals persist at all scales after returning the Fourier domain and plotting $r'(u)$ in the third row of the figure. The best explanation for the poor subtraction obtained with the natural weighting can be found in the profiles of the weighting functions shown in Figure 5. The fine, hashlike fluctuations that are visible in the profile for the natu-

ral weighting due to the discrete antenna positions in the array are responsible for introducing ripples and fluctuations into the dirty sky cube generated with this weighting. So, although natural weighting would result in the least thermal uncertainty in the dirty sky map, it does so at the expense of coupling the variations in the density of visibility measurements in the uv-plane to the dirty sky map. This is exactly what we would like to avoid in the sky map used for the foreground subtraction.

We can estimate the importance of structure in the uv coverage at a given length scale. In general, the synthesized array beams at each frequency measured by an interferometer will be identical, but proportional in overall size to the inverse of the frequency, since the distribution of visibility measurements is identical, but scales with the wavelength in the uv-plane. The coherence length of fluctuations in visibility coverage in the uv-plane is mapped to the spectral domain by this frequency-dependent scaling. The relationship between the coherence length in the uv-plane, l_u, and the coherence length in the
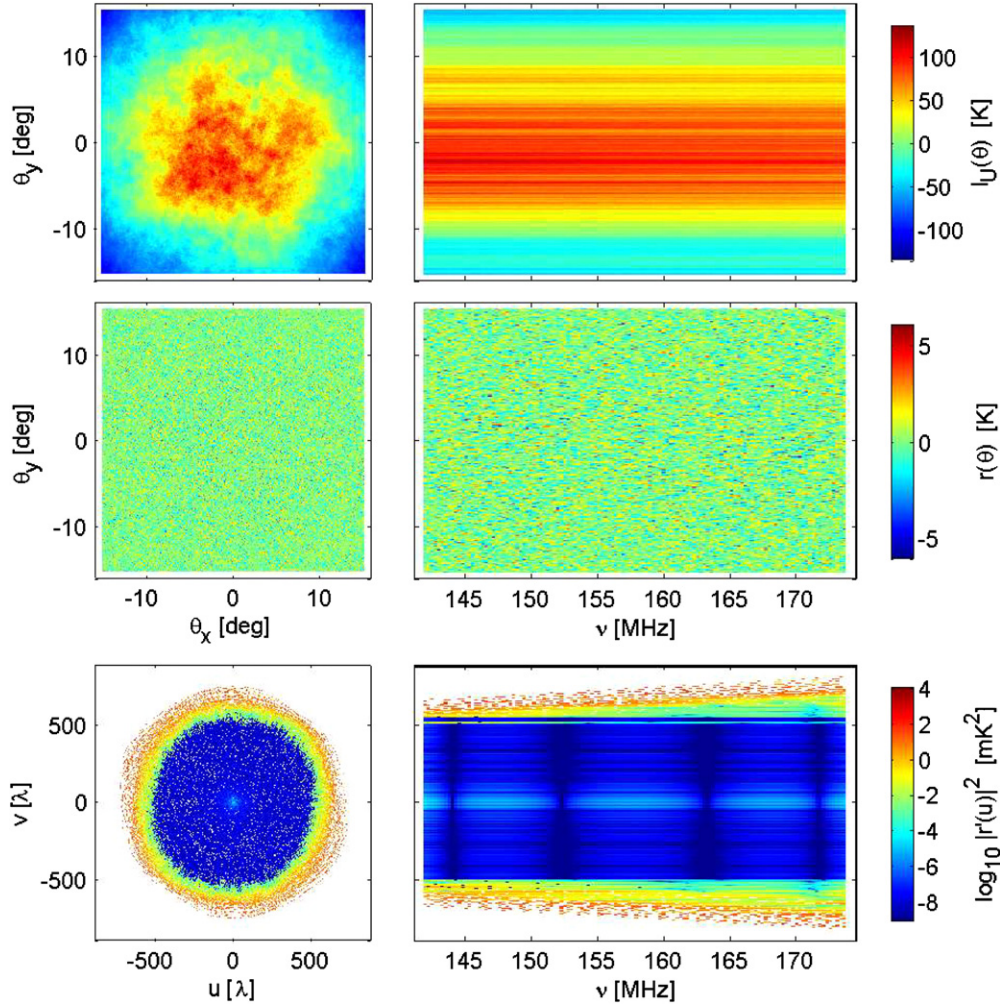
Figure 7. Polynomial foreground-subtraction technique applied to the simulated dirty sky map generated with a uniform weighting of the visibility map. The sky model used for the subtraction included only the foreground and thermal noise terms and *excluded* the 21 cm signal. The thermal noise contribution, however, has been artificially removed from these plots in order to better illustrate the foreground-subtraction residuals. In this case, the bottom row displays the unweighted residuals, r', in the Fourier domain. Similar to the reference case in Figure 6, it appears in the middle row that the residuals after foreground subtraction will be much larger this time than the ~ 25 mK redshifted 21 cm signal. However, after transforming the residuals back to the Fourier domain and unweighting, we find that the polynomial fit and subtraction removed the foreground contribution well again for most baselines. Aside from the short baselines that were poorly cleaned in even the instrument-free case, only baselines longer than $\sim 500\lambda$ are significantly contaminated when performing the subtraction in the dirty map generated with uniform weighting.

(A color version of this figure is available in the online journal.)

frequency domain, l_ν, is given by

$$\frac{l_u}{|u|} = \frac{l_\nu}{\nu_0},\tag{14}$$

where $|u|$ is the radius of interest in the uv-plane at frequency ν_0. From this expression, it is easy to see that modest features in the coverage of the visibility measurements near the origin of the uv-plane do not introduce significant fluctuations into the spectral domain because $l_\nu \rightarrow \infty$ as $|u| \rightarrow 0$, whereas variations in the coverage or weighting at large radii in the uv-plane will have much shorter coherence lengths in the spectral domain because $l_\nu \rightarrow 0$ as $|u| \rightarrow \infty$. In order to produce fluctuations in the spectral domain that have sufficiently long periods that they can be fit by a low-order polynomial across the observed bandwidth requires $l_\nu \gtrsim B$, where B is the bandwidth. For the MWA, with $\nu_0 \approx 150$ MHz and $B = 8$ to 32 MHz, this condition requires that

$$l_u \gtrsim |u|\frac{B}{\nu_0} \approx \frac{|u|}{10}\tag{15}$$

in order to prevent fluctuations in the spatial coverage from impacting the polynomial fits used in the foreground subtraction. Referring to Figure 5, this condition is clearly not met for the natural weighting function with its fine-scale features.

3.3. Smoothed Natural Weighting

We have so far seen that the uniform weighting allows foreground subtraction to succeed well below the required level, whereas the natural weighting does not. However, the uniform weighting succeeds at the expense of increasing the thermal noise in the derived dirty map by nearly an order of magnitude compared to the natural weighting. Since the natural weighting fails due to the fine-scale structure in the baseline distribution of the array, we explored smoothing the natural weighting function by approximating the distribution of baselines with a double exponential profile given by (see Figure 5)

$$N(u) \sim e^{-|u|/l_1} + e^{-|u|/l_2}\tag{16}$$

where $l_1 = 90\lambda$ and $l_2 = 190\lambda$, and the amplitude is set to match the true baseline distribution. The function is cored for $|u| \leqslant 6\lambda$,
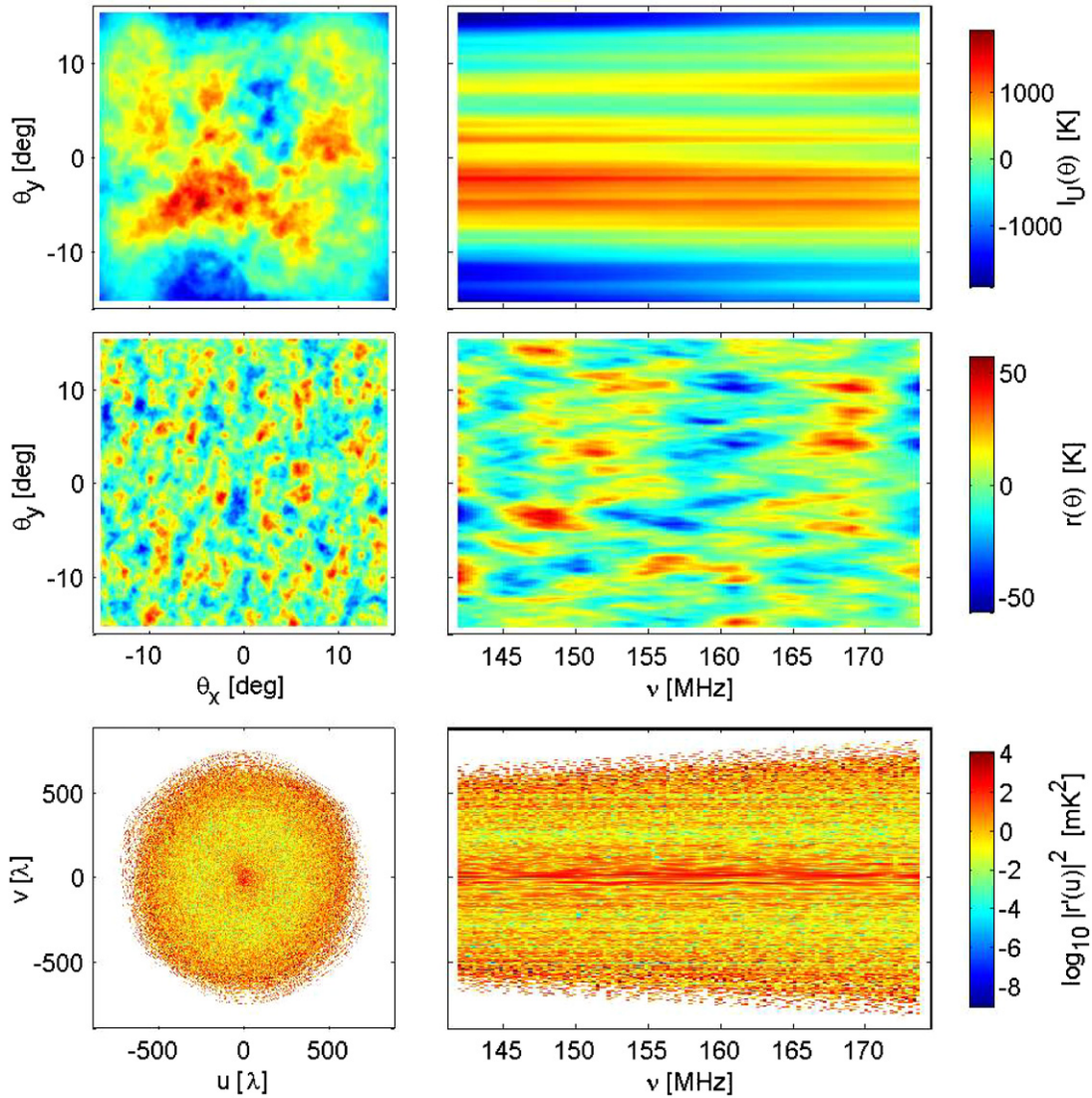
Figure 8. Same as Figure 7, but for the polynomial-subtraction technique applied to the dirty sky map generated with a natural weighting of the visibility map. No 21 cm signal was included in the analysis, and the thermal noise has been artificially removed from the plots. As with the uniform weighting, the residuals, r, in the image domain appear much larger than for the best-case fits to the sky model alone. In this case, however, even after transforming back to the Fourier domain and unweighting, the residuals remain much larger than for the other weighting methods.

(A color version of this figure is available in the online journal.)

approaching 0 linearly, in order to prevent overweighting these scales compared to the true natural weighting. This weighting function retains the emphasis on the short baselines inherent in the natural weighting case, but without introducing the fine-scale variations in the density of the visibility measurements into the weighting function. In our simulation, it resulted in only a 7% increase in thermal noise in the dirty map compared to the true natural weighting. The results of the foreground subtraction for this case are shown in Figure 9. This weighting function produces the best results in $r(\theta)$ and nearly matches the uniform weighting case in $r'(u)$.

3.4. Review

We pause for a moment to review the results of the analysis so far. At this stage, we can conclude that subtraction of confusion-level foreground contamination in dirty sky maps produced by the MWA is possible using the simple "per pixel" polynomial-subtraction technique. Uniform weighting of the gridded visibility measurements during the dirty map

generation process produces the best results of the three cases we tested. If maximizing the S/N in the derived dirty map, while preserving the purity of the spectral domain from mode-mixing contamination, is desired, then we have shown that a smooth approximation to the natural weighting function yields similar results to the uniform weighting, but without significantly increasing the thermal noise in the dirty map above the best-case natural weighting level. With additional studies, it seems likely that the smoothed variation of the natural weighting could be made to match the performance of the uniform weighting.

It may seem counterintuitive that the success of the polynomial-subtraction algorithm should be largely independent of the level of thermal noise in the input dirty map, so we note that this appearance is due, in part, to the fact that we have artificially removed thermal noise from the residuals plotted in Figures 7–9 in order to facilitate an unhindered comparison of the different weighting schemes. The derived foreground models do include the contributions of thermal noise, as do the resulting residuals. However, our interest is not in the uncertainty of the
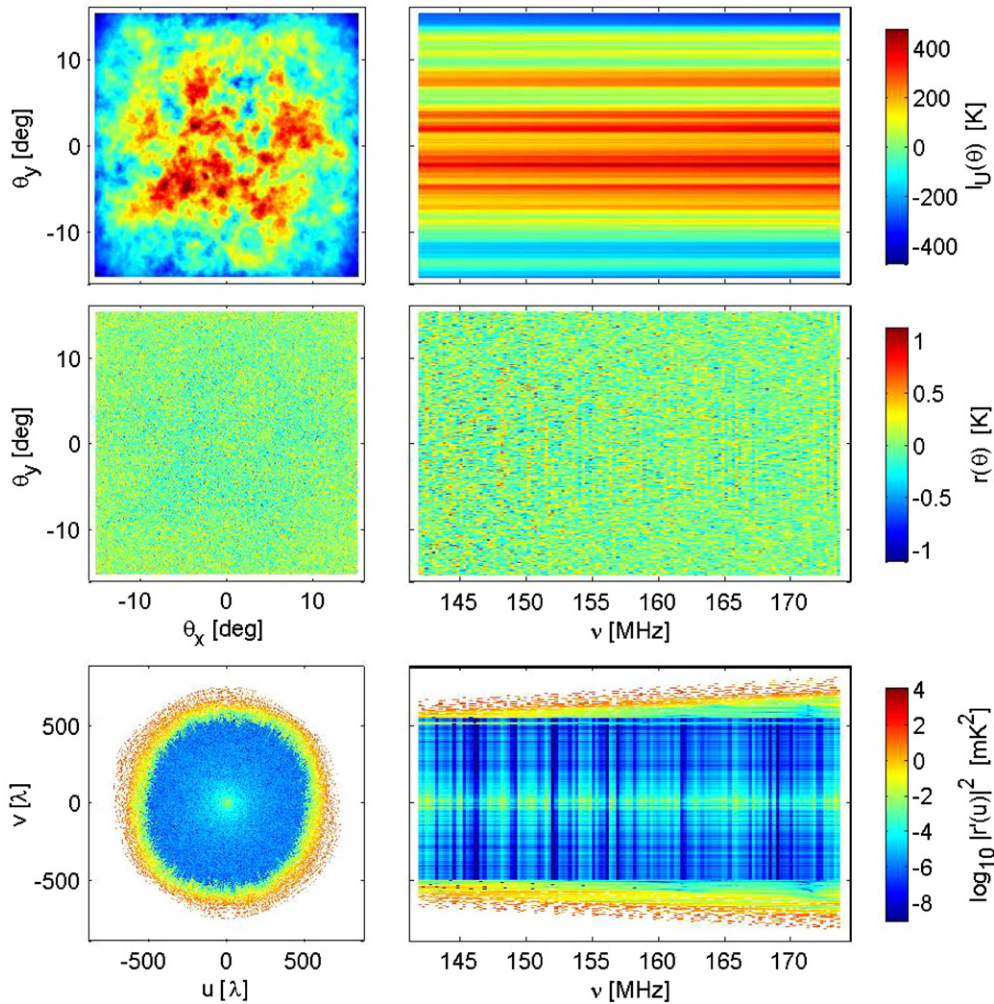
Figure 9. Same as Figure 7, but for the polynomial-subtraction technique applied to the dirty sky map generated with the smoothed approximation of the natural weighting of the visibility map. No 21 cm signal was included in the analysis, and the thermal noise has been artificially removed from the plots. For this case, the residuals, r, in the image domain are the lowest for the three dirty sky maps tested, but are still 2 orders of magnitude greater than the best-case subtraction for the sky model with no instrumental response. Transforming back to the Fourier domain and unweighting yield residuals that are nearly as low as the uniform weighting case.

(A color version of this figure is available in the online journal.)

foreground model, $d(\theta)$, but rather the uncertainty in the recovered 21 cm power spectrum, estimated from $r'(\mathbf{u})$. We have thus far postponed our book-keeping of the thermal uncertainty so as not to confuse to the two. In the remainder of this paper, we address this issue of thermal noise in the measurements and the ability of the MWA to detect the redshifted 21 cm signal after applying the foreground-subtraction steps discussed above. Since we have seen that the uniform weighting case is sufficient for foreground subtraction, we will use only uniform weighting in the analysis below for clarity of presentation.

4. RECOVERING THE EPOCH OF REIONIZATION SIGNAL

After foreground subtraction, the $\sim 10^8$ individual measurements comprising an MWA data cube contain largely thermal noise power and only a very small contribution from the redshifted 21 cm signal (and foreground residuals). To detect the EoR signal, we need to average over all of these measurements to obtain 10–20 estimators with sufficient S/N to detect and characterize the reionization signal. The foundation for a
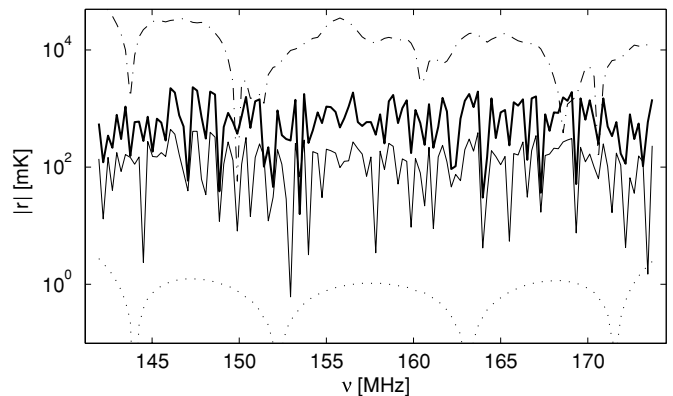

Figure 10. Polynomial-subtraction residuals along the same line of sight in the dirty sky data cubes for the three weighting schemes as well as for the pure foreground sky model with no instrumental response simulated. No 21 cm signal was included, and thermal noise has been artificially removed. The weighting schemes are uniform (thick solid line), natural (dashed-dotted line), exponential (thin solid line), and the foreground model alone (dotted line).

statistical EoR power spectrum measurement of the brightness temperature fluctuations in low-frequency, wide-field radio observations has been developed in the literature by Morales & Hewitt (2004) and Zaldarriaga et al. (2004). These efforts are built on the similar approach employed for interferometric measurements of CMB anisotropies (White et al. 1999; Hobson & Maisinger 2002; Myers et al. 2003). The primary approach is to convert the three-dimensional measurement cube to a one-dimensional power spectrum.

Following foreground subtraction, the full 32 MHz data cube will be divided into 8 MHz sub-bands, each of which will be reduced into a power spectrum independently. This approach is planned for the MWA analysis pipeline because the redshifted 21 cm signal is expected to evolve considerably over the $\Delta z \approx 2$ redshift range spanned by the full 32 MHz observation. By treating four sub-bands independently, any evolution of the signal can be largely neglected in the interpretation of the results since each sub-band will span only $\Delta z \approx 0.5$. The effect that this simplification has on the power spectra measurements is to eliminate measurements of the largest spatial scales, corresponding to the lowest $k_\parallel$ modes, along the line-of-sight direction. Since these modes contain most of the foreground power, these are not expected to have yielded usable information about the 21 cm signal anyway.

Each residual data sub-cube, $r'(\{\mathbf{u}, \nu\})$, is transformed into a three-dimensional wavenumber by applying a Fourier transform along the frequency axis, $\nu \rightarrow \eta$, and then applying a coordinate transformation into cosmological coordinates $\mathbf{k}$

$$r'(\mathbf{k}) = \mathbf{J}(\mathbf{k}, \{\mathbf{u}, \eta\})\mathbf{F}(\{\mathbf{u}, \eta\}, \{\mathbf{u}, \nu\})r'(\{\mathbf{u}, \nu\}), \qquad (17)$$

where $\mathbf{J}(\mathbf{k}, \{\mathbf{u}, \eta\})$ is given by the Jacobian of the coordinate transformation from $\mathbf{u}, \eta$ (in units of λ and Hz^{-1}) to $\mathbf{k}$ (in units of cMpc^{-1}, see Morales & Hewitt 2004, for details). The residual $r'(\mathbf{k})$ can be thought of as the three-dimensional Fourier transform of the foreground-subtracted image cube as expressed in comoving Mpc.

Due to the isotropy of space, the power spectrum is approximately spherically symmetric in $\mathbf{k}$ coordinates. We can thus square $r'(\mathbf{k})$ and average the result in spherical annuli

$$P(k) = \langle |r'(\mathbf{k})|^2 \rangle_{|\mathbf{k}|=k} \qquad (18)$$

to form the one-dimensional power spectrum (Morales & Hewitt 2004), or the more common dimensionless power spectrum given by $\Delta^2 = k^3 P(k)/(2\pi^2)$. However, it is useful to break the averaging from the three-dimensional $\mathbf{k}$-space to the one-dimensional k-space into two steps since both the foregrounds and a full treatment of the predicted redshifted 21 cm signal have *aspherical* structure in the Fourier domain. Following the approach of McQuinn et al. (2006), we first average over the angular sky direction to obtain $P(k_\perp, k_\parallel)$. This is conceptually similar to averaging over the m values and keeping the ℓ values in a CMB analysis, except we still have the line-of-sight dimension. In practice, the two-dimensional and one-dimensional power spectra are determined following the maximum-likelihood formalism that reduces to weighting the individual measurements by the inverse of the per-cell noise variance in the limit of no covariance terms between the cells. For the two-dimensional power spectrum this yields

$$P(k_\perp, k_\parallel) = \frac{\sum_{|\mathbf{k}_\perp|=k_\perp} \mathbf{U}_\mathbf{n}(\mathbf{k}, \mathbf{k})|r'(\mathbf{k})|^2}{\sum_{|\mathbf{k}_\perp|=k_\perp} U_n(\mathbf{k})} \qquad (19)$$

and for the one-dimensional power spectrum

$$P(k) = \frac{\sum_{|\mathbf{k}|=k} \mathbf{U}_\mathbf{n}(\mathbf{k}, \mathbf{k})|r'(\mathbf{k})|^2}{\sum_{|\mathbf{k}|=k} U_n(\mathbf{k})}, \qquad (20)$$

where $\mathbf{U}_\mathbf{n}$ is related to the natural weighting function used to generate dirty maps from the gridded visibility measurements, but with spectral structure converted into spectral covariance terms by the Fourier transform along the frequency axis. Neglecting the effects of RFI and RRL excision, the $k_\parallel$ covariance is highly peaked for the MWA due to the dense uv-coverage, and we make the simplifying approximation in Equations (19) and (20) to treat it as a delta function in this analysis so that the covariance matrix for the $\mathbf{k}$-space Fourier data cube is taken to remain diagonal and $\mathbf{U}_\mathbf{n}$) can be treated as a multiplicate operator independent of $k_\parallel$.

Theoretical models indicate that the true redshifted 21 cm signal will be composed of contributions (due to velocity field and other astrophysical effects) from modes that are modulated according to powers of $\mu \equiv \hat{k} \cdot \hat{n}$, where μ is the cosine of the angle between the line of sight and the wavevector (Kaiser 1987; Barkana & Loeb 2005; McQuinn et al. 2006). Thus,

$$P_{21}(\vec{k}) = P_{\mu^0}(k) + \mu^2 P_{\mu^2}(k) + \mu^4 P_{\mu^4}(k) + \cdots, \qquad (21)$$

and can also be fully described by $P_{21}(k_\perp, k_\parallel)$. The simple model of the 21 cm signal we are using in this paper, however, contains only a single P_{μ^0} contribution, and the MWA will have insufficient sensitivity initially to detect the higher order aspherical effects independently.

Even in reduced power spectra, the 21 cm signal is weaker than the total thermal noise power, and additional steps must be taken to isolate it. There are two possible approaches for extracting the 21 cm contribution from the noise-dominated measurement: (1) calculate the thermal noise contribution to the autocorrelation generated power spectrum and subtract it from the final result, or (2) generate the power spectrum by dividing the observation into two epochs of equal duration and then cross-correlating the data cubes from the two epochs (M. Tegmark 2007, private communication). This second approach preserves the persistent 21 cm signal and eliminates the thermal noise power (which will be independent between the two observing epochs and, therefore, average to zero during the cross-correlation), leaving only the thermal uncertainty. It has the advantage that the system temperature does not need to be known independently to high precision in order to recover the 21 cm signal (as long as the thermal noise properties of the instrument do not vary between epochs). We have chosen to follow this second approach and calculate the one-dimensional and two-dimensional power spectra as cross-power spectra by replacing the $|r'(\mathbf{k})|^2$ terms in Equations (18)–(20) with $(r_1'^* r_2' + r_2'^* r_1')/2$ such that Equation (18) becomes

$$P(k) = \left\langle \frac{r_1'^* r_2' + r_2'^* r_1'}{2} \right\rangle_k, \qquad (22)$$

where r_1' and r_2' are the residuals for the two epochs of the simulated observation.

The panels in Figure 11 illustrate the postsubtraction residuals in the two-dimensional power spectrum of the full simulated instrumental response. The left column of panels shows the postsubtraction power spectra of the individual and combined sky model components with (1) the foreground-only contribution
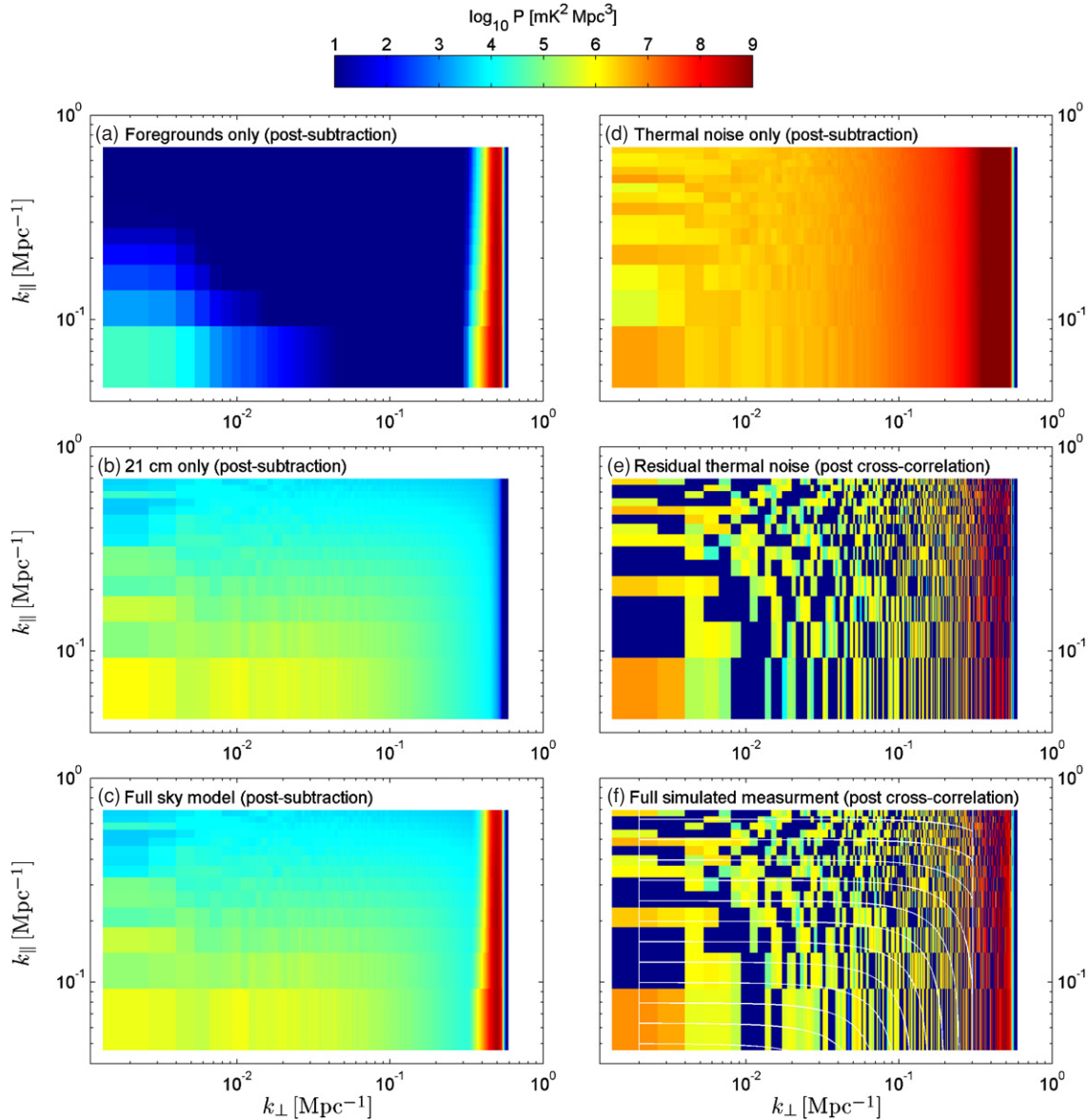
Figure 11. Two-dimensional power spectra of postsubtraction residuals from dirty maps generated with uniform weighting. For the left column of panels, the thermal noise has been artificially removed. It is evident that the foreground-only residuals following subtraction are much lower than the 21 cm signal in the range $k_\perp \lesssim 3 \times 10^{-1}$, beyond which corresponds to the outer annulus of poor foreground subtraction in the bottom panel of Figure 7. Panel (c) confirms that the 21 cm signal dominates the recovered power from the full-sky model since it appears nearly identical to panel (b). The white arcs in panel (f) illustrate the spherical shells of constant k that are used for the one-dimensional power spectrum in Figure 12.

(A color version of this figure is available in the online journal.)

to the sky model, (2) the redshifted 21 cm only contribution, and (3) the full combined sky model. Thermal noise has been artificially removed from these three panels to highlight the structure of the astrophysical sky. It is clear that the residual power in the postsubtraction, foreground-only contribution in Figure 11(a) is well below the 21 cm only power in Figure 11(b). Figure 11(c) confirms that the foreground subtraction recovers the 21 cm signal since the subtraction results on the full-sky model are nearly identical to the 21 cm only results. The residual foreground contamination in Figure 11(a) does not have the same symmetry as the redshifted 21 cm signal, and tends to have functional forms given by $P(\mathbf{k}) = P(k_\perp)P(k_\parallel)$, as opposed to the spherical and μ-modulated 21 cm symmetries. Morales et al. (2006) have discussed using these differences in shape to further separate the 21 cm power spectrum from the foregrounds, although we do not pursue this step in the present analysis.

The right column of Figure 11 illustrates the thermal noise properties and the removal of the mean thermal noise power from the data cube by the cross-correlation. Figure 11(d) shows the total thermal noise in the simulated power spectrum after 360 hr of observation with the MWA. Despite the long integration, the total thermal noise still dominates the desired 21 cm signal by over an order of magnitude. After dividing the simulated observation into two epochs and calculating the cross-correlation, however, the mean thermal noise power is removed from the final power spectrum, leaving only the residual thermal uncertainty power shown isolated in Figure 11(e) and with the full simulated measurement including the sky signal in Figure 11(f). The MWA will not have enough sensitivity within its first year of observations to produce a high S/N two-dimensional power spectrum, so the one-dimensional power spectrum will be the primary output product. The white curves
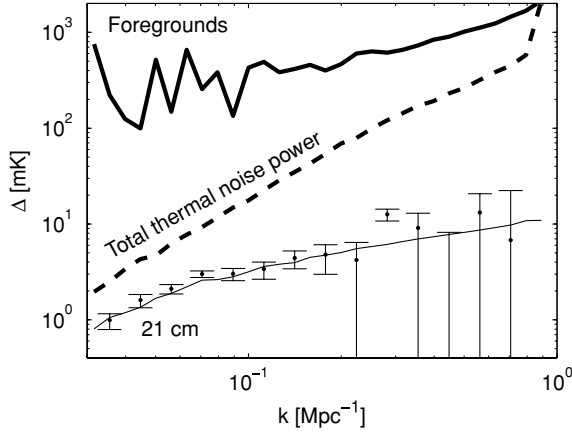
Figure 12. Simulated measurement of the one-dimensional spherically binned redshifted 21 cm power spectrum by the MWA in a single 8 MHz sub-band with 360 hr of integration. The thin solid line is the input 21 cm signal. The error bars give the thermal uncertainty on the recovered measurement and are 1σ. Measurements below $k \lesssim 10^{-1}$ Mpc^{-1} are highly coupled due to the limited depth of the observed comoving volume, as is evident from the constant-k contours in the bottom right panel of Figure 11. The thick solid line shows the total power (before subtraction) in the simulated sky data cube including thermal noise and foregrounds, and the thick dashed line shows the total power of the thermal noise alone.

in Figure 11(f) indicate the spherical shells of constant k that are used to calculate the one-dimensional power spectrum in Figure 12. In addition to the recovered 21 cm power spectrum, Figure 12 also illustrates the total sky power and total thermal noise power in the one-dimensional power spectrum before subtraction and cross-correlation.

We note that, during a year-long season of observing on the primary target window, the MWA should produce four power spectrum measurements, each comparable to the simulated results in Figures 11(f) and 12, spanning the full redshift range $7.1 < z < 9$ covered by the 32 MHz instrumental bandwidth in blocks of $\Delta z \approx 0.5$, corresponding to the individual 8 MHz sub-bands of the divided data cube. In addition, a secondary target field should yield an independent set of observations producing four additional similar power spectra, for a total of eight unique power spectra documenting the 21 cm emission during the reionization epoch.

4.1. Characterization of Subtraction Effects

The foreground-subtraction process introduces errors into the uncorrected derived estimate of the redshifted 21 cm power spectrum. The two principal sources of possible error are (1) the unavoidable removal of some of the large-scale power from the redshifted 21 cm signal by the polynomial fitting and subtraction algorithm, and (2) the contamination of the derived power spectrum by any residual foreground power resulting from an imperfect subtraction. These effects must be quantified through modeling to interpret real measurements accurately. As an initial step, we introduce a subtraction characterization factor, f_s, that captures the ratio of the recovered power spectrum to the true power spectrum such that

$$f_s(\mathbf{k}) = P_{21}(\mathbf{k})/P'_{21}(\mathbf{k}), \qquad (23)$$

where P'_{21} is the true input 21 cm power spectrum, and P_{21} is the recovered output power spectrum following the simulated measurement and foreground-subtraction process.

Figure 13 plots the net characterization factor calculated for our fiducial model. The bias due to performing the polynomial
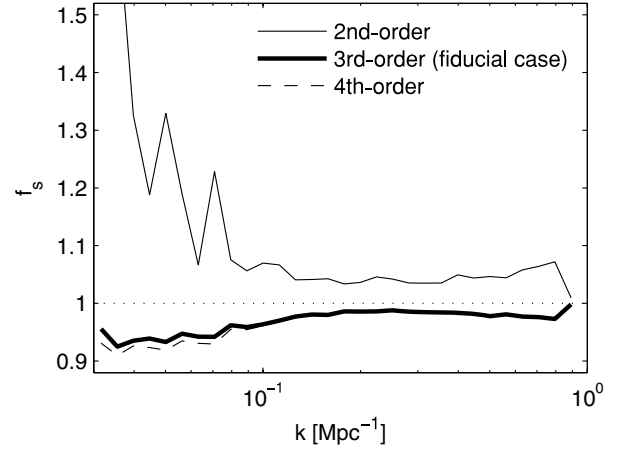

Figure 13. Subtraction characterization factor, f_s, for the one-dimensional binned power spectrum. The characterization factor is shown for three different levels of polynomial subtraction. As expected, the largest error is for large spatial modes (low k). Over most scales, however, the correction for our fiducial third-order polynomial subtraction is only $\sim 1\%$.

fit and subtraction in the presence of foregrounds reduces the overall power in the recovered power spectrum at all scales (such that $f_s(k) < 1$), although this effect is minor— less than 10% at any scale, and of order only $\sim 1\%$ for $k \gtrsim 10^{-1}$ Mpc^{-3}. This suggests that the polynomial-subtraction process is removing more power from the 21 cm signal than any residual foreground power following subtraction is adding to the recovered estimate. This conclusion is confirmed when we calculate the characterization factor for the recovered signal without any foregrounds in the sky model and find less than a 0.1% difference between this 21 cm only case and the result for the full-sky model shown in the figure.

As a counterpoint, we include in Figure 13 a separate curve (the thin solid line) for the characterization factor calculated for the subtraction process on the full-sky model using only a second-order polynomial. In this case, residual foreground power resulting from the poorer fit by the second-order polynomial contributes more to the recovered 21 cm power spectrum than the polynomial fit itself removes and yields $f_s > 1$ over the entire spectrum. This is particularly evident for $k < 10^{-1}$ Mpc^{-3}, where it is clear that the lower order polynomial failed to remove much of the foreground power. Finally, we note that increasing the polynomial to fourth-order has a minor effect on the recovered 21 cm signal (shown as the dashed line in Figure 13) that is evident only at large scales. This suggests that there is additional margin in the polynomial-subtraction approach to handle more complicated foregrounds than currently anticipated without substantially degrading the recovered 21 cm signal.

In principle, the characterization factor is dependent not only on the foreground-subtraction algorithm, but also on the specific shapes of the redshifted 21 cm signal and the foregrounds. Since we have assumed a simple 21 cm model with structure based only on a Gaussian random field of matter density perturbations as well as a fairly general model of the foreground contribution, the f_s derived from this analysis should be taken only as a guide.

5. DISCUSSION

In the preceding sections, we have shown that the effects of the frequency-dependent instrumental response of the MWA do not hinder the simple polynomial-fit diffuse foreground-subtraction scheme to recover the redshifted 21 cm signal. This

is particularly encouraging because the fit has been performed in dirty sky maps, which are very sensitive to instrumental effects. Thus, we have established, in some sense, a worst-case scenario that appears manageable.

Dense uv-coverage is a new advancement for radio arrays. In the traditional paradigm, the uv-plane is almost always sparsely sampled (as with the VLA and other existing arrays), and there are essentially no regions where the coverage can be said to be "complete," as is the case for the MWA within $u \lesssim 500\lambda$. The reason that this is generally not a problem for telescopes like the VLA is that, in most cases, these instruments are not attempting to make observations at or below the confusion limit imposed by faint extragalactic continuum sources and diffuse Galactic structure. So, although the VLA has sparse coverage in the uv-plane, the sky that it is observing is also essentially sparse and deconvolution algorithms are able to reconstruct a reasonable model of the point sources in the field of view.

It is useful to construct an analytic approximation for the simulations discussed in the previous sections in order to foster a more intuitive understanding of the interaction between the array properties and the success of the foreground subtraction. Condon (1974), Perley & Erickson (1984), and Artyukh (2003) provide a foundation for such a model, and Sault (2006) and Wayth (2006), both MWA project documents, have adapted the results of these efforts for analyzing confusion noise in measurements of the MWA.

As outlined in Section 2.3, a dirty sky map produced by an interferometer is, neglecting calibration errors, the result of multiplying the true sky by the primary antenna tile beam and then convolving by the synthesized array beam. If the sky consists only of the faint extragalactic continuum sources with Poisson statistics, then the variance in the intensity of the dirty sky map, σ_D^2, will be related to the variance in the true sky by

$$\sigma_D^2(\theta_x, \theta_y) = \sigma_T^2 \iint B^2(\theta_x - \theta_x', \theta_y - \theta_y') P^2(\theta_x, \theta_y) d\theta_x d\theta_y, \tag{24}$$

where σ_T^2 is the variance in the intensity of the true sky, B describes the response of the synthesized array beam as a function of angle, and P describes the primary antenna tile beam. This equation can be simplified by assigning simple models for the response profiles of the synthesized array beam and the primary antenna tile beam. We take the beams to be described by top hat functions, such that the response is defined to be one within a region of diameter Θ_B (for the synthesized array beam; Θ_P, for the primary antenna tile beam). Outside this region, the response is taken to be $B_{rms} \ll 1$ for the synthesized beam, and zero for the primary antenna tile beam. Thus, we include an allowance for the sidelobes of the synthesized beam, but not for the primary beam. With these simplifications, the integrals in Equation (24) can be performed and the expression reduces to

$$\sigma_D^2 \approx \sigma_S^2 \left(1 + B_{rms}^2 \Omega_P / \Omega_B \right), \tag{25}$$

where σ_S^2 is the variance due to the faint sources in our model and accounts for the extra variance due to the weak angular clustering, $\Omega_B \approx \Theta_B^2$ is the solid angle of the synthesized array beam, and $\Omega_P \approx \Theta_P^2$ is the solid angle of the primary antenna tile beam.

Inspecting this simplified form of Equation (24), we see that when the synthesized array beam has no sidelobes, the variance

in the dirty sky map is equal to that of the true sky map (after gridding). This is equivalent to complete coverage of the uv-plane and corresponds generally to the cases considered by Di Matteo et al. (2004), Zaldarriaga et al. (2004), Furlanetto & Briggs (2004), Gnedin & Shaver (2004), Santos et al. (2005), Wang et al. (2006), McQuinn et al. (2006), Gleser et al. (2008). But when sidelobes are included in the model of the synthesized array beam, the variance in the dirty map is increased above that of the true sky map. For sparse coverage in the uv-plane, B_{rms}^2 will be given approximately by the inverse of the number of independent measurements in the uv-plane. In the limiting case of a single, very short integration (such as a "snapshot" observation), the number of independent measurements is equal to the number of baselines. For the MWA, with its $\sim 125,000$ instantaneous baselines, $B_{rms}^2 \approx 10^{-5}$ in the worst-case, and is even better (lower) for long integrations when Earth-rotation synthesis increases the number of independent measurements. Using $\Theta_B \approx 5'$ and $\Theta_P \approx 30°$, we can estimate for the MWA that

$$B_{rms}^2 \Omega_P / \Omega_B \approx 1. \tag{26}$$

Thus, the variance due to the sidelobes in the synthesized array beam is comparable in magnitude to the inherent variance in the intensity of the true sky. A large field of view acts to oppose the advantages of dense coverage of the uv-plane in this instance by introducing more variance into the dirty sky map. In this regime, to produce an estimate of the sky without the effects of the sidelobes of the synthesized array beam, the full inversion of $\mathbf{A}$ is required since it cannot be approximated as a sparse matrix.

Nevertheless, sidelobe-induced noise in the dirty sky maps does not necessarily mean that the foreground subtraction will be adversely affected. If the sidelobe patterns were the same for each image plane in the data cube, then the spectrum along each pixel would simply receive a faint mirror contribution from the smooth spectra along all the other pixels. It is the variation with frequency of the sidelobe-induced noise that can cause trouble in the planned foreground-subtraction technique. As found in Section 3, however, the density of the MWA uv-coverage effectively eliminates the frequency-dependence of the sidelobe structure of the array beam over the large range of scales where the uv-plane is completely sampled at all frequencies. Although it does not appear to be necessary for the MWA, less densely sampled arrays could control the spectral dependence of their array beam sidelobes by masking out regions of the uv-plane that are not sampled at all frequencies and only processing regions that do have coverage in each spectral channel (as considered in Jelić et al. 2008, for the more sparsely sampled LOFAR).

Despite the increased susceptibility of long baselines in introducing undesirable fluctuations into the spectral domain, it is important to point out that they serve an important purpose. Long baselines are still expected to be necessary in order to peel away the bright sources in the field of view since the peeling process requires a precise knowledge of the position of each bright source, which is greatly improved with long baselines.

It has been shown previously that the ideal baseline distribution for minimizing thermal uncertainty in measurements of the redshifted 21 cm one-dimensional power spectrum is extremely condensed (Bowman et al. 2006; Lidz et al. 2008). In this work, we have further demonstrated that a compact array like the MWA also produces the dense uv-coverage necessary to efficiently subtract faint and diffuse foregrounds from dirty maps.

REFERENCES

Alvarez, H., Aparici, J., May, J., & Olmos, F. 1997, A&AS, 124, 315
Artyukh, V. S. 2003, A&A, 409, 787
Barkana, R., & Loeb, A. 2005, ApJ, 624, L65
Blake, C., & Wall, J. 2002, MNRAS, 329, L37
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, ApJ, 638, 20
Bridle, A. H. 1967, MNRAS, 136, 219
Briggs, D. S., Schwab, F. R., & Sramek, R. A. 1999, in ASP Conf. Ser. 180, Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 127
Cohen, A. S., Röttgering, H. J. A., Jarvis, M. J., Kassim, N. E., & Lazio, T. J. W. 2004, ApJS, 150, 417
Condon, J. J. 1974, ApJ, 188, 279
Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. 1998, AJ, 115, 1693
Di Matteo, T., Ciardi, B., & Miniati, F. 2004, MNRAS, 355, 1053
Di Matteo, T., Perna, R., Abel, T., & Rees, M. J. 2002, ApJ, 564, 576
Erickson, W. C., McConnell, D., & Anantharamaiah, K. R. 1995, ApJ, 454, 125
Fomalont, E. B., Windhorst, R. A., Kristian, J. A., & Kellerman, K. I. 1991, AJ, 102, 1258
Furlanetto, S. R., & Briggs, F. H. 2004, New Astron. Rev., 48, 1039
Giardino, G., Banday, A. J., Fosalba, P., Górski, K. M., Jonas, J. L., O'Mullane, W., & Tauber, J. 2001, A&A, 371, 708
Giardino, G., Banday, A. J., Górski, K. M., Bennett, K., Jonas, J. L., & Tauber, J. 2002, A&A, 387, 82
Gleser, L., Nusser, A., & Benson, A. J. 2008, MNRAS, 391, 383
Gnedin, N. Y., & Shaver, P. A. 2004, ApJ, 608, 611
González-Nuevo, J., Toffolatti, L., & Argüeso, F. 2005, ApJ, 621, 1
Hales, S. E. G., Baldwin, J. E., & Warner, P. J. 1988, MNRAS, 234, 919
Haslam, C. G. T., Stoffel, H., Salter, C. J., & Wilson, W. E. 1982, A&AS, 47, 1
Hobson, M. P., & Maisinger, K. 2002, MNRAS, 334, 569
Jelić, V., et al. 2008, MNRAS, 389, 1319
Kaiser, N. 1987, MNRAS, 227, 1
Kellermann, K. I., Pauliny-Toth, I. I. K., & Williams, P. J. S. 1969, ApJ, 157, 1
Kogut, A., Banday, A. J., Bennett, C. L., Gorski, K. M., Hinshaw, G., & Reach, W. T. 1996, ApJ, 460, 1
La Porta, L., Burigana, C., Reich, W., & Reich, P. 2008, A&A, 479, 641
Laing, R. A., Riley, J. M., & Longair, M. S. 1983, MNRAS, 204, 151
Landecker, T. L., & Wielebinski, R. 1970, Aus. J. Phys. Astrophys. Suppl., 16, 1
Lawson, K. D., Mayer, C. J., Osborne, J. L., & Parkinson, M. L. 1987, MNRAS, 225, 307
Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Hernquist, L. 2008, ApJ, 680, 962
Liu, A., Tegmark, M., & Zaldarriaga, M. 2009, MNRAS (http://www3.interscience.wiley.com/journal/122232273/abstract)
McGilchrist, M. M., Baldwin, J. E., Riley, J. M., Titterington, D. J., Waldram, E. M., & Warner, P. J. 1990, MNRAS, 246, 110
McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, ApJ, 653, 815
Mitchell, D. A., Greenhill, L. J., Wayth, R. B., Sault, R. J., Lonsdale, C. J., Cappallo, R. J., Morales, M. F., & Ord, S. M. 2008, IEEE J. of Selected Topics in Signal Processing, 2 , 707
Morales, M. F. 2005, ApJ, 619, 678
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7
Morales, M. F., & Matejek, M. 2008, arXiv:0810.5107
Myers, S. T., et al. 2003, ApJ, 591, 575
Oh, S. P., & Mack, K. J. 2003, MNRAS, 346, 871
Pearson, T. J. 1975, MNRAS, 171, 475
Perley, R. A., & Erickson, W. C. 1984, A Proposal for a Large, Low Frequency Array Located at the VLA Site, Memorandum VLA–Scientific–146, National Radio Astronomy Observatory
Platania, P., Bensadoun, M., Bersanelli, M., de Amici, G., Kogut, A., Levin, S., Maino, D., & Smoot, G. F. 1998, ApJ, 505, 473
Platania, P., Burigana, C., Maino, D., Caserini, E., Bersanelli, M., Cappellini, B., & Mennella, A. 2003, A&A, 410, 847
Reich, P., & Reich, W. 1986, A&AS, 63, 205
Reich, P., & Reich, W. 1988, A&AS, 74, 7
Rengelink, R. B., Tang, Y., de Bruyn, A. G., Miley, G. K., Bremer, M. N., Roettgering, H. J. A., & Bremer, M. A. R. 1997, A&AS, 124, 259
Roger, R. S., Costain, C. H., Landecker, T. L., & Swerdlyk, C. M. 1999, A&AS, 137, 7
Rogers, A. E. E., & Bowman, J. D. 2008, AJ, 136, 641
Santos, M. G., Cooray, A., & Knox, L. 2005, ApJ, 625, 575
Sault, R. J. 1984, The Weighting of Visibility Data, Memorandum VLA, Scientific, 154, National Radio Astronomy Observatory
Sault R. J. 2006 (available on the Murchison Widefield Array internal web server)
Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437
Shaver, P. A. 1975, Pramana, 5, 1
Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, A&A, 345, 380
Subrahmanyan, R., & Ekers, R. D. 2002, arXiv:astro-ph/0209569v1
Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22
Thompson, R. A., Moran, J. M., & Swenson, G. W. 2001, Interferometry and Synthesis in Radio Astronomy (2nd ed.; New York: Wiley-Interscience)
Walmsley, C. M., & Watson, W. D. 1982, ApJ, 260, 317
Wang, X., Tegmark, M., Santos, M. G., & Knox, L. 2006, ApJ, 650, 529
Wayth R. 2006 (available on the Murchison Widefield Array internal web server)
White, M., Carlstrom, J. E., Dragovan, M., & Holzapfel, W. L. 1999, ApJ, 514, 12
White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479
Wieringa, M. H. 1991, PhD thesis, Rijksuniversiteit Leiden
Windhorst, R. A., Fomalont, E. B., Partridge, R. B., & Lowenthal, J. D. 1993, ApJ, 405, 498
Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., & Koo, D. C. 1985, ApJ, 289, 494
Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, ApJ, 608, 622
Zhang, X., Reich, W., Reich, P., & Wielebinski, R. 2003, A&A, 404, 57