DELAY SPECTRUM WITH PHASE-TRACKING ARRAYS: EXTRACTING THE H I POWER SPECTRUM FROM THE EPOCH OF REIONIZATION

SOURABH PAUL[1], SHIV K. SETHI[1], MIGUEL F. MORALES[2], K. S. DWARKANATH[1], N. UDAYA SHANKAR[1], RAVI SUBRAHMANYAN[1],
N. BARRY[2], A. P. BEARDSLEY[3], JUDD D. BOWMAN[3], F. BRIGGS[4,5], P. CARROLL[2], A. DE OLIVEIRA-COSTA[6], JOSHUA S. DILLON[7],
A. EWALL-WICE[6], L. FENG[6], L. J. GREENHILL[8], B. M. GAENSLER[5,9,10], B. J. HAZELTON[2], J. N. HEWITT[6], N. HURLEY-WALKER[11],
D. J. JACOBS[3], HAN-SEEK KIM[5,12], P. KITTIWISIT[3], E. LENC[5,9], J. LINE[5,12], A. LOEB[8], B. MCKINLEY[5,12], D. A. MITCHELL[5,13],
A. R. NEBEN[6], A. R. OFFRINGA[4,5], B. PINDOR[5,12], J. C. POBER[14], P. PROCOPIO[5,12], J. RIDING[5,12], I. S. SULLIVAN[2], M. TEGMARK[6],
NITHYANANDAN THYAGARAJAN[3], S. J. TINGAY[5,11,15], C. M. TROTT[5,11], R. B. WAYTH[5,11], R. L. WEBSTER[5,12], J. S. B. WYITHE[5,12],
ROGER CAPPALLO[16], M. JOHNSTON-HOLLITT[17], D. L. KAPLAN[18], C. J. LONSDALE[16], S. R. MCWHIRTER[16], E. MORGAN[6],
D. OBEROI[19], S. M. ORD[5,11], T. PRABU[1], K. S. SRIVANI[1], A. WILLIAMS[11], AND C. L. WILLIAMS[6]
[1] Raman Research Institute, Bangalore 560080, India
[2] Department of Physics, University of Washington, Seattle, WA 98195, USA
[3] School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA
[4] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[5] ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), Building A28, School of Physics, The University of Sydney, NSW 2006, Australia
[6] Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[7] Departments of Physics and Astronomy, University of California, Berkeley, CA, 94720, USA
[8] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
[9] Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia
[10] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, ON M5S 3H4, Canada
[11] International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia
[12] School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia
[13] CSIRO Astronomy and Space Science (CASS), P.O. Box 76, Epping, NSW 1710, Australia
[14] Department of Physics, Brown University, Providence, RI 02912, USA
[15] Istituto Nazionale di Astrofisica (INAF)—Istituto di Radio Astronomia, Via Piero Gobetti, Bologna, I-40129, Italy
[16] MIT Haystack Observatory, Westford, MA 01886, USA
[17] School of Chemical & Physical Sciences, Victoria University of Wellington, Wellington 6140, New Zealand
[18] Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA
[19] National Centre for Radio Astrophysics, Tata Institute for Fundamental Research, Pune 411007, India

ABSTRACT

The detection of redshifted 21 cm emission from the epoch of reionization (EoR) is a challenging task owing to strong foregrounds that dominate the signal. In this paper, we propose a general method, based on the delay spectrum approach, to extract H I power spectra that are applicable to tracking observations using an imaging radio interferometer ("Delay Spectrum with Imaging Arrays"). Our method is based on modeling the H I signal taking into account the impact of wide field effects such as the w-term, which are then used as appropriate weights in cross-correlating the measured visibilities. Our method is applicable to any radio interferometer that tracks a phase center and could be utilized for arrays such as the Murchison Widefield Array (MWA), Low Frequency Array (LOFAR), Giant Meterwave Radio Telescope (GMRT), Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER), and HERA. In the literature the delay spectrum approach has been implemented for near-redundant baselines using drift scan observations. In this paper we explore the scheme for non-redundant tracking arrays. This is the first application of delay spectrum methodology to such data to extract the H I signal. We analyze 3 hr of MWA tracking data on the EoR1 field. We present both two-dimensional ($k_\parallel$, $k_\perp$) and one-dimensional (k) power spectra from the analysis. Our results are in agreement with the findings of other pipelines developed to analyze the MWA EoR data.

Key words: cosmology: observations – cosmology: theory – dark ages, reionization, first stars – techniques: interferometric

1. INTRODUCTION

Probing the epoch of reionization (EoR) remains one of the outstanding aims of modern cosmology. In the past decade, intriguing details have emerged about this epoch from a host of cosmological observables. Gunn–Peterson (GP) tests on spectra of quasi-stellar objects (QSOs; Fan et al. 2006) in the redshift range $5.7 < z < 6.3$ suggest that the universe made a transition to full reionization during this period. On the other hand, cosmic microwave background radiation (CMBR) temperature and polarization anisotropy measurements (Komatsu et al. 2011; Planck Collaboration et al. 2015) suggest that the universe might have been fully ionized in a redshift range $8 < z < 10$ (Planck Collaboration et al. 2016). Both these observables have their strengths and weaknesses. The GP test, based on Lyα absorption, is not able to distinguish between a fully neutral medium from the one ionized to one part in a thousand. CMBR anisotropies, based on photon scattering off free electrons, are sensitive to the integrated Thompson scattering optical depth and therefore cannot reliably construct the tomography of the reionization epoch.

In the recent past, major experimental efforts have been undertaken to study the EoR in redshifted 21 cm line emission. In particular, many radio interferometers in the

frequency range $80 < \nu < 300$ MHz are currently operational that specifically aim to detect the EoR; for example, the Low Frequency Array (LOFAR, Van Haarlem et al. 2013), 21 Centimeter Array (21CMA, Zheng et al. 2016), Giant Meterwave Radio Telescope (GMRT, Paciga et al. 2013), Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER, Parsons et al. 2014), and the Murchison Widefield Array (MWA, Bowman et al. 2009; Lonsdale et al. 2009; Tingay et al. 2013). Even though the detection of the redshifted H I line from the EoR remains the most direct and, possibly the most promising, way to delineate the details of the epoch, this method is beset with its own set of issues. First, unlike CMBR anisotropies, the theoretical modeling of the H I signal from the EoR is considerably harder principally owing to uncertainty in the nature of ionizing sources and the details of their formation and evolution. Second, the signal is expected to be weak with brightness temperature $T_B \simeq 10$ mK. Many hundreds of hours of observation are needed to detect such a signal with current interferometers, with the attendant complication of maintaining instrumental stability for such long periods. Third, the foreground is expected to be many orders of magnitude larger than the signal (for details on the three issues see e.g., Barkana & Loeb 2001; Morales & Hewitt 2004; Furlanetto et al. 2006; Morales & Wyithe 2010; Zaroubi 2013 and references therein).

Radio interferometers to estimate the underlying power spectra have been successfully employed for CMB data analysis (Hobson et al. 1995). This method has also been suggested as a possible probe of the intensity correlations of the redshifted H I line, including from the EoR (Bharadwaj & Sethi 2001; Bharadwaj & Ali 2005; Datta et al. 2007).

Many different approaches have been discussed to detect the H I signal in the presence of dominant foregrounds, such as by Hazelton et al. (2013), Jelic et al. (2008), Harker et al. (2009), Liu & Tegmark (2011), Morales et al. (2012), Trott et al. (2012), and Dillon et al. (2013, 2014). They are all based on the expectation that foregrounds are smooth in frequency space as they arise from continuum emission, e.g., synchrotron radiation, in both our Galaxy and extragalactic sources. On the other hand the H I signal has significant structure in frequency space. It is conceivable that all these sources, both point and diffuse, can be subtracted from the images, leaving behind the H I signal and Gaussian noise. LOFAR partly relies upon this technique (Chapman et al. 2012, 2013). Another possible method is based on the isolation of foregrounds from the H I signal using the power spectrum of the observed signal in conjugate space to the observed frequency (Parsons et al. 2012a, 2012b; Pober et al. 2013; Thyagarajan et al. 2013). Variants of this "delay space" (Parsons et al. 2012a, 2012b) method are particularly relevant for interferometers such as MWA that have low angular resolution and have been used extensively for the analysis of PAPER data. In this approach the observed interferometric data—visibilities for each antenna pair as a function of frequency—are Fourier transformed along the frequency axis. The Fourier conjugate variable effectively captures signal delay between antenna pairs, which allows one to isolate foregrounds. In the context of the three-dimensional (3D) H I power spectrum this variable can be related to cosmological distance along the line of sight (for details of this mathematical correspondence see Parsons et al. 2012b; Liu

et al. 2014a). The "delay spectrum" constructed from the "delay space" approach can be used to recover the cosmological 3D H I power spectrum. In this approach, one deals with visibilities directly, which are primary data products of radio interferometers.

In this paper we propose a new method based on the delay space approach to extract the power spectrum of the H I signal in the presence of noise and foregrounds. Our method is based on modeling the H I signal taking into account the impact of the w-term (arising from non-coplanarity of the array, Cornwell et al. 2008) and the distortion of the intensity pattern during a tracking run. The information from the H I signal is used as weights to cross-correlate the measured visibilities. The proposed method ("Delay Spectrum with Imaging Arrays") is a general method applicable for tracking with radio interferometers with wide primary beams and arbitrary array configuration (e.g., MWA, LOFAR) and can also be applied to interferometers with redundant baselines. We apply the proposed method to analyze 3 hr of MWA data and compare our results with noise and foreground simulations.

The paper is organized as follows. In the next section, we describe the method of visibility correlation using the delay space approach in detail. In particular, this method is applied to the H I signal. In Section 3, we describe the MWA data and the initial analysis of these data based on the publicly available image processing software Common Astronomical Software Applications (CASA). In Section 4 the pipeline to extract the H I power spectrum is discussed. In Section 5, the results from MWA data are discussed and compared with simulations of foregrounds and noise. In the final section, we summarize our results and indicate possible future directions.

Throughout this paper, we have used the Planck+WP best fit values of cosmological parameters: $\Omega_{m0} = 0.3183$, $\Omega_{\Lambda 0} = 0.6817$, $\Omega_{b0} h^2 = 0.02203$, $h = 0.6704$, $\sigma_8 = 0.8347$, and $n_s = 0.9619$ (Planck Collaboration et al. 2014, 2015).

2. H I SIGNAL AND ITS CORRELATIONS

In this section we study the H I signal using visibility correlations in delay space. Using our formulation we derive, in addition to new results, many results already known in the literature. The main new results are: the impact of the w-term (Section 2.1) and changing intensity pattern in a tracking run (Section 2.2) on the H I correlations. Our results are valid for any radio interferometer but our aim here is to underline their applicability to MWA.

The most important inputs from the MWA array configuration for our study are: (a) the MWA primary beam and (b) the bandwidth. Other properties of the MWA array that have a bearing on our analysis are the non-coplanarity of the array and its baseline distribution.

In delay space, spectrally smooth foregrounds lend themselves to ready interpretation. It can be shown that visibilities computed in delay space allow isolation of such foregrounds from the regions dominated by the EoR signal and noise (e.g., see Parsons & Backer 2009; Datta et al. 2010; Parsons et al. 2012a, 2012b; Vedantham et al. 2012; Thyagarajan et al. 2013, 2015a; Dillon et al. 2014; Liu et al. 2014a, 2014b). This can be achieved by Fourier transforming the raw visibilities in frequency space. However,

being 3D and statistical in nature, the properties of the H I signal can only be inferred by correlating the observed visibilities. Our approach, which is based on visibility correlations in delay space, allows us to develop a unified method to deal with both the H I signal and the foregrounds, which are discussed in Appendix A.

Radio interferometers measure the spatial correlation of the electric fields from the sky, with visibility $V_\nu(\boldsymbol{u}_\nu)$

$$V_\nu(\boldsymbol{u}_\nu) = \int A_\nu(\boldsymbol{\theta}) I_\nu(\boldsymbol{\theta}) e^{-i2\pi\boldsymbol{u}_\nu\cdot\boldsymbol{\theta}} d^2\theta \qquad (1)$$

Here $\boldsymbol{u}_\nu = (u_\nu, v_\nu)$ is the distance vector between the antennas of the interferometer (also called baseline vector) in units of wavelength, $\boldsymbol{\theta}$ denotes the position vector on the sky plane (which can be be expressed as a pair of direction cosines $\{l, m\}$), and ν is the frequency of observation.

We neglect the impact of the w-term in this section in writing the relation between the visibility and specific intensity in Equation (1). The w-term arises from non-coplanarity of the interferometric array and changes as the interferometer tracks a region. In Section 2.1 we show that the inclusion of the w-term causes an effective shrinking of the primary beam. For our study we calculate how the H I signal is affected owing to non-zero w as compared to the $w = 0$ case.

We explicitly express the frequency dependence of all the quantities. These quantities are: the background specific intensity $\bar{I}_\nu$, the primary beam $A_\nu(\boldsymbol{\theta})$, and the baseline $\boldsymbol{u}_\nu$. In terms of brightness temperature T_B, $I_\nu(\boldsymbol{\theta}) = 2kT_B(\boldsymbol{\theta})/\lambda^2$.

The sky intensity can be decomposed as

$$I_\nu(\boldsymbol{\theta}) = \bar{I}_\nu + \Delta I_\nu(\boldsymbol{\theta}) \qquad (2)$$

where $\bar{I}_\nu$ and $\Delta I_\nu(\boldsymbol{\theta})$ are the isotropic and fluctuating parts of the intensity distribution. Since the isotropic component does not contribute to interferometric measurement[20], the visibility recorded at frequency ν can be written as

$$V_\nu(\boldsymbol{u}_\nu) = \int A_\nu(\boldsymbol{\theta}) \Delta I_\nu(\boldsymbol{\theta}) e^{-i2\pi\boldsymbol{\theta}\cdot\boldsymbol{u}_\nu} d^2\theta. \qquad (3)$$

The measured visibility receives contributions from the redshifted H I line, the foregrounds, and the noise.

For the H I signal the observed intensity fluctuations can be related to the H I perturbations in Fourier space, $\delta_{H I}(\boldsymbol{k})$, as

$$\Delta I_\nu(\boldsymbol{\theta}) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{H I}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \qquad (4)$$

Here $\boldsymbol{r} = \{\boldsymbol{\theta}, r_\nu\}$ specifies the 3D position of the the H I emission; r_ν is the coordinate distance to the point of observation: $r_\nu = \int dz/H(z)$ with the limit of this integral extending from zero to redshift $z = \nu_e/\nu_0 - 1$. $\delta_{H I}(\boldsymbol{k})$ comprises many physical effects: density fluctuations, ionization inhomogeneity, density-ionization fraction cross-correlation, etc. (Zaldarriaga et al. 2004; Furlanetto et al. 2006). Together this can be expressed as

$$\delta_{H I}(\boldsymbol{k}) = \beta_b \delta_b + \beta_x \delta_x + \beta_\alpha \delta_\alpha + \beta_T \delta_T - \delta_{\partial v} \qquad (5)$$

Each quantity in the above equation corresponds to the fractional variation of a particular physical quantity: δ_b refers to fluctuation

---

[20] However, some methods have been discussed in the recent literature to extract the monopole signal from interferometric measurements (Presley et al. 2015; Singh et al. 2015).

in baryonic density, δ_α for the Lyα coupling coefficient x_α, δ_x for the neutral fraction, δ_T for T_K, and $\delta_{\partial v}$ for the line of sight peculiar velocity gradient. β factors denote the expansion coefficients of the corresponding quantity (Furlanetto et al. 2006).

Current experiments such as MWA, LOFAR, and PAPER aim at statistical detection of the EoR signal. The quantities of interest here are the correlation functions of the H I fluctuations. The most important correlation function one seeks to detect in an EoR experiment is the power spectrum, $P_{H I}$, defined as

$$\langle \delta^*_{H I}(\boldsymbol{k}) \delta_{H I}(\boldsymbol{k}') \rangle = (2\pi)^3 \delta^3(\boldsymbol{k} - \boldsymbol{k}') P_{H I}(k). \qquad (6)$$

The H I power spectrum can be constructed from the correlation of the observed visibilities. Substituting the form of fluctuation ΔI_ν from Equation (4) in the visibility expression (Equation (3)), we get

$$V_\nu(\boldsymbol{u}_\nu) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{H I}(\boldsymbol{k}) e^{ir_\nu k_\parallel} \int A_\nu(\boldsymbol{\theta})$$
$$\times \exp\left[-2\pi i\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\cdot\boldsymbol{\theta}\right] d^2\theta. \qquad (7)$$

Here we have decomposed the wave vector $\boldsymbol{k}$ as components on the plane of the sky $\boldsymbol{k}_\perp$ and along the line of sight $k_\parallel$. The integral over angles is the Fourier transform of the primary beam $A_\nu(\boldsymbol{\theta})$, which allows us to rewrite this equation as

$$V_\nu(\boldsymbol{u}_\nu) = \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{H I}(\boldsymbol{k}) e^{ir_\nu k_\parallel} a\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right) \qquad (8)$$

where

$$a\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right) \equiv \int A_\nu(\boldsymbol{\theta})$$
$$\times \exp\left[-2\pi i\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\cdot\boldsymbol{\theta}\right] d^2\theta. \qquad (9)$$

Using Equation (6), the visibility correlation function can be computed to be

$$\langle V_\nu(\boldsymbol{u}_\nu) V^*_{\nu'}(\boldsymbol{u}'_{\nu'}) \rangle = \bar{I}_\nu^2 \int \frac{d^3k}{(2\pi)^3} P_{H I}(k) e^{i\Delta r_\nu k_\parallel}$$
$$\times a\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right) a\left(\boldsymbol{u}'_\nu - \frac{\boldsymbol{k}_\perp r_{\nu'}}{2\pi}\right) \qquad (10)$$

Here $\Delta r_\nu = |r_{\nu'} - r_\nu|$ which for $|\nu' - \nu| \ll \nu$ can be simplified to $\Delta r_\nu = r'_\nu|\nu' - \nu|$, where $r'_\nu = |dr_\nu/d\nu|$. Equation (10) gives the correlation of the H I signal in three dimensions in which the two coordinates $\boldsymbol{u}$ correspond to Fourier components of the H I signal while the third ν refers to the coordinate of the fluctuation in position, r_ν, space (Bharadwaj & Sethi 2001).

To isolate the impact of foregrounds and obtain regions dominated by the H I signal and the noise ("EoR window"), we compute the the visibilities in delay space (Parsons & Backer 2009; Parsons et al. 2012a, 2012b, 2014; Liu et al. 2014a):

$$V_\tau(\boldsymbol{u}) = \int \exp(i2\pi\tau\nu) V_\nu(\boldsymbol{u}_\nu) d\nu \qquad (11)$$

Here τ, the conjugate variable of ν, defines the relevant variable in delay space. The delay space approach can be applied to data to isolate spectrally smooth foregrounds; we discuss the delay space approach as applied to such foregrounds in Appendix A. In Equation (11), we have suppressed the frequency dependence of the baseline on the lhs as the frequency dependence of all the quantities has been integrated. The baseline vector can be expressed as $\boldsymbol{u}_\nu = \boldsymbol{u}_{\nu_0}\nu/\nu_0$, where ν_0 is some fixed frequency that lies within the bandwidth. On the lhs of Equation (11), the frequency independent baselines $\boldsymbol{u} \equiv \boldsymbol{u}_{\nu_0}$. Throughout this paper, we assume $\nu_0 = 154$ MHz, the central frequency of the bandwidth we use for MWA data analysis.

The autocorrelation of $V_\tau(\boldsymbol{u})$ can be written as

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle = \int\int d\nu d\nu' \bar{I}_\nu \bar{I}_{\nu'} \int \frac{d^3k}{(2\pi)^3} P_{\mathrm{H\,I}}(k)$$
$$\times \exp(i2\pi\tau(\nu'-\nu) + i(k_\parallel(r_{\nu'}-r_\nu)))$$
$$\times \int d^2\theta \exp\left(i2\pi\boldsymbol{\theta}\cdot\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta})$$
$$\times \int d^2\theta' \exp(-i2\pi\boldsymbol{\theta}'$$
$$\cdot\left(\boldsymbol{u}'_{\nu'} - \frac{\boldsymbol{k}_\perp r_{\nu'}}{2\pi}\right)\right)A_{\nu'}(\boldsymbol{\theta}').$$
$$(12)$$

To make further progress, frequency dependent quantities are Taylor expanded. For baselines, this is a straightforward re-expression of the baseline vector as the vector is linear in frequency: $\boldsymbol{u}'_{\nu'} = \boldsymbol{u}'_\nu + \Delta\nu d\boldsymbol{u}'_\nu/d\nu$, where $\Delta\nu = \nu' - \nu$. It should be noted that $d\boldsymbol{u}'_\nu/d\nu$ is the physical baseline length measured in the unit of time.

After the Taylor expansion of relevant quantities r_ν and $\boldsymbol{u}_\nu$, we obtain

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle$$
$$= \int\int d\nu d\nu' \bar{I}_\nu \bar{I}_{\nu'} \int \frac{d^3k}{(2\pi)^3} P_{\mathrm{H\,I}}(k)$$
$$\times \int d^2\theta \exp\left(i2\pi\boldsymbol{\theta}\cdot\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta})$$
$$\times \int d^2\theta' \exp(-i2\pi\boldsymbol{\theta}'$$
$$\cdot\left(\boldsymbol{u}'_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_{\nu'}(\boldsymbol{\theta}')$$
$$\times \exp[i\Delta\nu(2\pi\tau + k_\parallel dr_\nu/d\nu + d\boldsymbol{u}'_\nu/d\nu\cdot\boldsymbol{\theta}'$$
$$+ dr_\nu/d\nu\boldsymbol{k}_\perp\cdot\boldsymbol{\theta}'/(2\pi))].$$
$$(13)$$

Here $\Delta\nu = \nu' - \nu$; and all the quantities in Equation (13) have been written as explicit functions of ν and $\Delta\nu$. This allows us to simplify the integral further by making the coordinate transform $y = (\nu + \nu')/2$ and $x = (\nu' - \nu)/2$; the Jacobian of this transformation is unity. We can make a further simplification by using $y \simeq \nu$. This is justified for our case as we assume the bandwidth to be $\simeq$10 MHz around a central frequency of $\simeq$150 MHz. All the frequency dependent variables change by less than 10% for this case. This reduces

Equation (13) to

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle$$
$$= \int\int d\nu dx \bar{I}_\nu{}^2 \int \frac{d^3k}{(2\pi)^3} P_{\mathrm{H\,I}}(k)$$
$$\times \int d^2\theta \exp\left(i2\pi\boldsymbol{\theta}\cdot\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta})$$
$$\times \int d^2\theta' \exp\left(-i2\pi\boldsymbol{\theta}'\cdot\left(\boldsymbol{u}'_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta}')$$
$$\times \exp[i2x(2\pi\tau + k_\parallel dr_\nu/d\nu + d\boldsymbol{u}'_\nu/d\nu\cdot\boldsymbol{\theta}'$$
$$+ dr_\nu/d\nu\boldsymbol{k}_\perp\cdot\boldsymbol{\theta}'/(2\pi))].$$
$$(14)$$

Given the H I power spectrum $P_{\mathrm{H\,I}}(k)$ this integral could be computed numerically. However, it is possible to determine the correlation scales in both the transverse and line of sight directions by carefully examining Equation (14). The integral over angles shows that the dominant contribution comes from wavenumbers such that $\boldsymbol{k}_\perp \simeq 2\pi\boldsymbol{u}_\nu/r_\nu$. This relation allows us to simplify the integrals over θ, θ', and x. In particular, different terms in the exponent containing τ can be estimated. Using $d\boldsymbol{u}_\nu/d\nu = \boldsymbol{u}_\nu/\nu$, the last two terms in the exponents are on the order of $\boldsymbol{k}_\perp.\boldsymbol{\theta}'r_\nu/(2\pi\nu)$ (the term containing $dr_\nu/d\nu$ is slightly smaller because $\nu/r_\nu dr_\nu/d\nu \simeq 0.3$). For the MWA primary beam $\theta_0 \simeq 0.3$, and for the MWA baseline distribution, the term $k_\parallel dr_\nu/\nu$ generally dominates over these terms, especially in the regions dominated by EoR. For all our calculations we use parameters specific to MWA, in particular, the primary beam of MWA. However, the formulation presented here is general enough to be applicable to other arrays.

By dropping the last two terms, which are subdominant, in the exponent containing τ, we can separate the integrals over x and angles. This gives us

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle$$
$$= \int d\nu \bar{I}_\nu{}^2 \int \frac{d^3k}{(2\pi)^3} P_{\mathrm{H\,I}}(k)$$
$$\times \int d^2\theta \exp\left(i2\pi\boldsymbol{\theta}\cdot\left(\boldsymbol{u}_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta})$$
$$\times \int d^2\theta' \exp\left(-i2\pi\boldsymbol{\theta}'\cdot\left(\boldsymbol{u}'_\nu - \frac{\boldsymbol{k}_\perp r_\nu}{2\pi}\right)\right)A_\nu(\boldsymbol{\theta}')$$
$$\times \int dx \exp[i2x(2\pi\tau + k_\parallel dr_\nu/d\nu)].$$
$$(15)$$

The integral over x can readily be carried out now. The dominant contribution to the integral comes from $\tau \simeq k_\parallel/(2\pi)dr_\nu/d\nu$, which establishes the correlation scale in the direction along the line of sight. The variation of frequency dependence of integrals over θ and θ' is expected to be small for the bandwidth of MWA and therefore these integrals can be computed at some frequency that lies within the bandwidth. In this paper we assume these values to be fixed at the central frequency $\nu = 154$ MHz and use them to compute the relations in Equation (16). If the frequency dependence of the primary beam and the background intensity are neglected, the integral over ν is trivial. As noted above, this is a good approximation for MWA. The power spectrum of the H I signal, based on Equation (15), is shown in Figure 1. In this calculation, we use
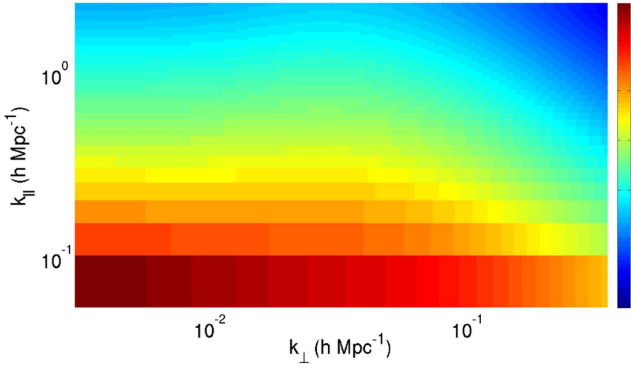
Figure 1. Power spectra of the expected cosmological H I signal. The power is plotted as $\log_{10}P$ where P is in units $mK^2 (Mpc/h)^3$.

the H I power spectrum $P_{H I}(k)$ given by the simulation of Furlanetto et al. (2006).

Equations (15) and (10) can be used to infer many important properties of the H I signal:

1. There exists a near one-to-one relation between $\boldsymbol{u}$, τ, and the Fourier components of the H I power spectrum (Morales & Hewitt 2004; Parsons et al. 2012a, 2012b; Paul et al. 2014):

$$u_\nu = \frac{k_{\perp 1} r_\nu}{2\pi}, \ v_\nu = \frac{k_{\perp 2} r_\nu}{2\pi}, \ \tau = \frac{k_\| c (1 + z)^2}{2\pi H_0 \nu_{21} E(z)} \quad (16)$$

where ν_{21} is the rest frame frequency of the 21 cm line, r_ν is the transverse comoving distance, $E(z) = \sqrt{\Omega_M (1 + z)^3 + \Omega_k (1 + z)^2 + \Omega_\Lambda}$, and z is the redshift corresponding to the observed frequency ν. The relation between $k_\|$ and τ follows from the relation: $2\pi\tau = k_\| dr_\nu/d\nu$ (Equation (15)). As noted above, all the frequency dependent quantities in Equation (16) are computed at a fixed frequency $\nu_0 = 154$ MHz.

2. The correlations in the sky plane and along the line of sight are nearly separable. This allows us to compute weights in the plane of the sky owing to the w-term and the distortion of intensity pattern in a tracking run (the next two subsections) without the additional complication owing to frequency dependence of these quantities.

3. Equations (15) and (16) allow us to simplify the relation between visibility correlation and the H I power spectrum. Equation (15) can be solved in the limit defined by Equation (16) to give (e.g., Morales 2005; McQuinn et al. 2006; Pen et al. 2009; Thyagarajan et al. 2015a)

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}) \rangle = \frac{\bar{I}_\nu^2 \Delta\nu \theta_0^2}{r_\nu^2 dr_\nu/d\nu} P_{H I}(k). \quad (17)$$

Here the MWA primary beam solid angle $\theta_0^2 = \lambda^2/A_{\rm eff}$. For MWA $A_{\rm eff} = 21.5$ m^2 at 150 MHz (Tingay et al. 2013). $\Delta\nu = 10.2$ MHz is the total bandwidth we use in this work. The mean specific intensity $\bar{I}_\nu = 2k_B T_B/\lambda^2$. This allows us to express the H I signal as the square of the product of mean brightness temperature and the H I power spectrum $T_B^2 P_{H I}(k)$ in the units $(mK)^2 (Mpc/h)^3$. It should be emphasized that Equation (17) provides the suitable normalization only when $\boldsymbol{u}' = \boldsymbol{u}$, $w = 0$, as has been assumed throughout this section, and the impact of sky intensity distortion while tracking a region is not considered. All these effects act to lower the rhs of

Equation (17), the measured visibility correlation, for a fixed signal $T_B^2 P_{H I}(k)$. This is accounted for by appropriate weights, which we discuss in the next two sections.

2.1. H I Signal and w-term

From Equations (14) and (15), we can gauge the impact of the w-term. These equations and the discussion following them show that the angular integrals depend only weakly on the line of sight variables. The main effect of the w-term is to alter the integrals over angles which we study here.

For a given baseline $\boldsymbol{b}$, $w_\nu = \boldsymbol{b} \cdot \boldsymbol{s}_0 \nu/c$, where $\boldsymbol{s}_0$ is the phase center at any time. As a region is tracked, the w-term changes owing to the drift of the phase center. For a tracking run, $u_\nu^2 + w_\nu^2$ is left invariant at any frequency; this result simply follows from the fact the the baseline length is fixed.

After the inclusion of the w-term, the measured visibility for a given intensity distribution is given as

$$V_\nu(\boldsymbol{u}_\nu, w_\nu) = \int A_\nu(l, m) \Delta I_\nu(l, m)$$
$$\times \exp[-i2\pi\{u_\nu l + v_\nu m$$
$$- w_\nu/2(l^2 + m^2)\}]dldm. \quad (18)$$

Here we have replaced $\boldsymbol{\theta}$ with its components (l, m) and also made the approximation $\sqrt{1 - l^2 - m^2} - 1 \simeq -(l^2 + m^2)/2$, which is valid for the MWA primary beam. Substituting Equation (4) into Equation (18) gives us

$$V_\nu(\boldsymbol{u}_\nu, w_\nu) = \bar{I}_\nu \int \frac{d^3 k}{(2\pi)^3} \delta_{H I}(\boldsymbol{k}) e^{ir_\nu k_\|} \int A_\nu(l, m)$$
$$\times \exp\left[-2\pi i \left\{ \left(u_\nu - \frac{k_{\perp 1} r_\nu}{2\pi} \right) l \right. \right.$$
$$+ \left(v_\nu - \frac{k_{\perp 2} r_\nu}{2\pi} \right) m$$
$$\left. \left. - \frac{w_\nu}{2} (l^2 + m^2) \right\} \right] dldm. \quad (19)$$

Each MWA tile being approximately a square aperture, the primary beam $A_\nu(l, m)$ can be written as

$$A_\nu(l, m) = \frac{\sin(\pi L_x l)}{\pi L_x l} \frac{\sin(\pi L_y m)}{\pi L_y m}. \quad (20)$$

Here L_x and L_y are dimensionless. They correspond to the ratio of the length of the tile along the x- and y-axis to the wavelength. For the central wavelength of the observation $L_x = L_y \simeq 2$. Equations (18) and (20) show that integrals over l and m are separable and identical. These integrals cannot be computed analytically but under certain approximations meaningful analytic expressions can be found. Let us define

$$Q_\nu(k_{\perp 1}; u_\nu, w_\nu) = \int dl \exp\left[-2\pi i \left\{ \left(u_\nu - \frac{k_{\perp 1} r_\nu}{2\pi} \right) l \right. \right.$$
$$\left. \left. - \frac{w_\nu}{2} l^2 \right\} \right] \frac{\sin(\pi L_x l)}{\pi L_x l} \quad (21)$$

where $Q_\nu(k_{\perp 1}; u_\nu, w_\nu)$ is a function of $k_{\perp 1}$ and is parametrized by u_ν and w_ν. First we consider $w = 0$. In this case, it can be shown that if the limits of the integral are allowed to go from

minus infinity to plus infinity, we obtain

$$Q_\nu(k_{\perp 1}; u_\nu, 0) = \frac{1}{L_x} \text{ if } |(u_\nu - r_\nu k_{\perp 1}/(2\pi))|/L_x < 1/2$$
$$= 0 \text{ otherwise.}$$
(22)

We notice that the approximation used is good because the function has a compact support provided by the primary beam. As $L_x \simeq 1/\theta_0$, where θ_0 is the extent of the primary beam, this result means that, for a given u_ν, the wavenumbers that contribute to the integral are the ones that are bounded by the extent of the primary beam. This result is already implied by Equation (14).

Equation (21) cannot be analytically approximated so readily for non-zero w. We use the stationary phase approximation to analytically evaluate the integral. For this assumption to hold, the phase of the exponent should be much larger than the slow variation of the primary beam. This would be the case if wl^2 is large. In this approximation, we obtain

$$Q_\nu(k_{\perp 1}; u_\nu, w_\nu) = \sqrt{\frac{2i}{w}} \frac{\sin(\pi L_x(u_\nu - k_{\perp 1}r_\nu/(2\pi))/w_\nu)}{\pi L_x(u - k_{\perp 1}r_\nu/(2\pi))/w_\nu}$$
$$\times \exp\left[-i(u_\nu - k_{\perp 1}r_\nu/(2\pi))^2/(4\pi w_\nu)\right].$$
(23)

The main impact of the inclusion of the w-term can be discerned from this expression. In the limit of large w, the impact of the w-term is to shrink the MWA beam and the primary beam tends to $1/\sqrt{w}$ (Cornwell et al. 2008).[21] However, this also means that the spread of $k_{\perp 1}$ for which the integral is non-zero also increases, as seen in the terms involving the sine function. If the decrease of the primary beam results in a loss of signal-to-noise, an increase in the correlation length $|\boldsymbol{u} - \boldsymbol{u}'|$ gains signal-to-noise. We can write visibility correlation for pairs of $\boldsymbol{u}$, w and $\boldsymbol{u}'_\nu$, w'_ν as

$$\langle V_\nu(\boldsymbol{u}_\nu, w_\nu) V_\nu^*(\boldsymbol{u}'_\nu, w'_\nu)\rangle$$
$$= \int \frac{d^3k}{2\pi^3} P_{H1}(k) Q(k_{\perp 1}; u_\nu, w_\nu)$$
$$\times Q(k_{\perp 2}; v_\nu, w_\nu) Q(k_{\perp 1}; u'_\nu, w'_\nu)$$
$$\times Q(k_{\perp 2}; u'_\nu, w'_\nu).$$
(24)

Equation (24) can be computed numerically. In Figure 2, we show how the H I correlation function is affected in the presence of the w-term. These expressions are also valid for diffuse foregrounds which have a different 2D power spectrum and frequency dependence, e.g., optically thin synchrotron radiation for which the angular and frequency dependence is separable. As Equation (24) can be used to compute the impact of the w-term at any frequency, it can readily be generalized to study diffuse foregrounds.

2.2. Time-dependent Coordinate System and w-term

In a tracking interferometric observation, a phase center is tracked and snapshots are taken at regular intervals with short
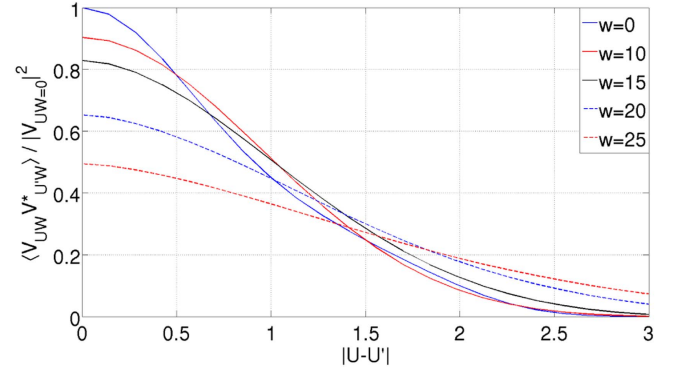
---

[21] The impact of the w-term can be more readily computed if the beam is Gaussian (e.g., Appendix B of Paul et al. 2014). In this case, the primary beam approaches $(1/\theta_0^2 + \pi w_\nu)^{-1/2}$ for non-zero w.

Figure 2. H I signal is plotted as a function $|\boldsymbol{u} - \boldsymbol{u}'|$ for different values of w. The impact of the w-term is to decrease the overall signal owing to shrinking of the primary beam and increase the correlation length $|\boldsymbol{u} - \boldsymbol{u}'|$.

duration. Each of these snapshots can be imaged and the images added if the successive fields of view can be assumed to be coplanar. This approximation breaks down for wide field-of-view instruments such as MWA. One manifestation of the wide field of view is the w-term whose impact was studied in the previous subsection. In this section we generalize the discussion of the last subsection to take into account the time dependence of the non-coplanarity of the tracked region (Taylor et al. 1999).

As the region is tracked, the relation between the image and astronomical coordinates changes, which distorts the intensity pattern with respect to the phase center being tracked. More specifically, this effect arises from the projection of a non-coplanar array on a plane, which is necessary to perform the Fourier transform for imaging. It is best illustrated with a set of point sources. These sources appear to move with respect to the phase center (e.g., Figure 19-9 in Taylor et al. 1999, p. 180). The distortion of the intensity pattern corresponds to non-uniform stretching and it increases for sources further away from the phase center. Thus, this effect cannot be corrected by a standard shift of coordinate. The non-uniform stretching makes the situation complex, and the standard grid approach is difficult to implement in this case. For a set of point sources, the correction for this effect could be applied iteratively in the image plane (Chapter 19, Taylor et al. 1999, p. 180).

For a small field of view, this effect can be neglected and a unique coordinate system (e.g., time-independent direction cosines $\{l, m\}$) can be used to relate the image coordinates with the sky intensity pattern for a long tracking run. However, it is not possible to define such a coordinate system when either the field of view is large or the tracking period is long.

Our aim here is not to correct for this effect but rather to estimate its impact on the correlation of visibilities at two different times during a tracking run. Suppose we measure visibilities within a small cell in the u–v plane (the size of the cell will be discussed in a later section) centered around a baseline $\{u, v\}$ at $t = 0$. At a later time $t = t'$ another baseline might enter this cell. From the discussion in the previous subsections (e.g., Figure 2) the two visibilities are expected to correlate strongly with each other (even if the values of w differ significantly for these two sets, this statement is generally true). However, visibilities measured at two different times do not correspond to the same intensity pattern. Our aim here is to estimate the level of de-correlation caused by the distortion of intensity pattern during a tracking run. In this paper, we construct a time-dependent coordinate system which allows us

to analyze this distortion of intensity pattern. We assess the impact of this effect when a region is tracked using the MWA primary beam. In particular, we consider this effect on the visibilities produced by the EoR H I signal.

We start by recalling the definition of direction cosines for a point on the sky whose coordinates, declination and hour angle $\{\delta, h\}$, are written as (Christiansen & Hogbom 1969)

$$l = \sin(h)\cos(\delta), \tag{25}$$

$$m = \cos(\delta)\cos(h)\sin(\phi) - \sin(\delta)\cos(\phi), \tag{26}$$

$$n = \cos(\delta)\cos(h)\cos(\phi) + \sin(\delta)\sin(\phi). \tag{27}$$

It can be shown that $n = \sqrt{1 - l^2 - m^2}$. The phase center is always defined as $l = 0$, $m = 0$, $n = 1$; for the coordinates defined above it is $h = 0$ and $\delta = \phi$.

As a phase center is tracked owing to the rotation of the Earth, δ remains fixed but the hour angle changes. For a wide field of view, this can result in distortion of the intensity pattern of the sky. To take this effect into account, we can define a time-dependent coordinate system:

$$l(\delta h) = \sin(h + \delta h)\cos(\delta) - \sin(h_0 + \delta h)\cos(\delta_0). \tag{28}$$

Here h_0 and δ_0 define the phase center for $\delta h = 0$; δh defines the flow of time. m can be similarly defined and n can be computed from l and m. This definition gives a time-dependent coordinate system where the coordinates are always defined with respect to the phase center. It is easy to verify that for small field of view and for small tracking times, which corresponds to cases when higher order terms in h, δh, and δ can be dropped, l is independent of time which means that the distance of a point from the phase center is left invariant under tracking. In such cases, the intensity pattern on the sky corresponds to the unique intensity pattern defined by sky coordinates δ and h and remains unchanged as the phase center is tracked.

However, when this approximation breaks down, l becomes a function of time and it is impossible to define a unique relation between direction cosines and sky coordinates. This means that any quantities defined with respect to sky coordinates (e.g., intensity pattern) become time dependent. The visibility for the H I signal is given by

$$V_\nu(\boldsymbol{u}_\nu, w_\nu, \delta h)$$
$$= \bar{I}_\nu \int \frac{d^3k}{(2\pi)^3} \delta_{\mathrm{H\,I}}(\boldsymbol{k}) e^{ir_\nu k_\parallel} \int A_\nu(l, m)$$
$$\times \exp\left[-2\pi i \left\{\left(u_\nu - \frac{k_{\perp 1} r_\nu}{2\pi}\right)l \right.\right.$$
$$+ \left(v_\nu - \frac{k_{\perp 2} r_\nu}{2\pi}\right)m$$
$$\left.\left. - \frac{w_\nu}{2}(l^2 + m^2)\right\}\right]\sin\delta\, d\delta\, dh. \tag{29}$$

The direction cosines l and m are now functions of time. The angular integral is carried out over δ and h. Unlike the earlier case (fixed grid) this is not a product of two 1D integrals. The correlation of the visibilities $V_\nu(\boldsymbol{u}_\nu, w_\nu, \delta h)$ can be computed using the same methods as outlined in the previous sections.

In Figure 3 we show the results when the effect of the time-dependent coordinate system ("moving grid") is included. The initial phase center ($\delta h = 0$) is chosen to be $h = 0$ and $\delta = \phi$.
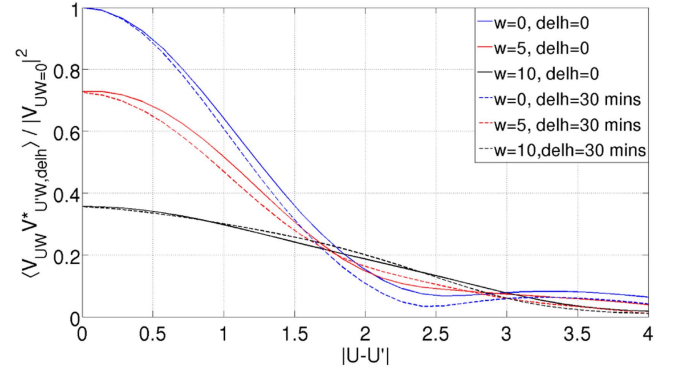

Figure 3. H I signal is plotted as a function $|\boldsymbol{u} - \boldsymbol{u}'|$ for different values of w and δh. The initial phase center is chosen to be $h = 0$ and $\delta = \phi$.

The results are shown for two different values of δh and a range of w values. We only show the autocorrelation function for a given value of δh. But the results shown in Figure 3 can be used to assess the cross-correlation of visibilities measured at two different times. For our case the value of this cross-correlation lies between the autocorrelations of visibilities measured at the same time. The moving grid does not introduce another scale in the problem and the results in this case are not very different from the case for a fixed grid. In both cases the dominant correlations occur for $\boldsymbol{u} - \boldsymbol{u}' < 1.5$.

Figures 2 and 3 are based on the MWA primary beam. However, it is possible to glean generic information applicable for other primary beams from them. First, the decorrelation length $|\boldsymbol{u} - \boldsymbol{u}'|$ scales as the inverse of the primary beam (e.g., Paul et al. 2014). So for a smaller beam, the decorrelation seen in the figures as a function of $|\boldsymbol{u} - \boldsymbol{u}'|$ would be shallower. The impact of the w-term for a different primary beam can be partially gauged from Equation (23), which is valid for large values of w. In this limit, the primary beam tends to $1/\sqrt{w}$, irrespective of the primary beam of the telescope. It is difficult to analytically estimate the impact of the w-term when this limit does not hold. But it can be shown that the impact of the w-term diminishes for a smaller primary beam, e.g., a Gaussian beam for which the primary beam tends to $(1/\theta_0^2 + \pi w_\nu)^{-1/2}$ for non-zero w (e.g., Paul et al. 2014). As noted above, the distortion of intensity pattern during a tracking run is a wide field effect. For a smaller primary beam, the level of decorrelation seen in Figure 3 would be smaller but it is difficult to analytically estimate it.

2.3. Weights for Cross-correlation

Equation (29) can be used to compute the counterpart of Equation (14) which takes into account the impact of a non-zero w-term and the distortion of intensity pattern. We compute this expression for the visibility correlation in delay space numerically. In this formulation, the measured visibility is a function of four parameters: τ, $\boldsymbol{u}$, w, t. Here, as noted above, $\boldsymbol{u}$ and w are the values of these variables at a fixed frequency which we choose to be $\nu_0 = 154$ MHz.

We define the weight on a given cross-correlation as

$$\mathcal{W}(\boldsymbol{u}, w, t; \boldsymbol{u}', w', t') = \frac{\langle V_\tau(\boldsymbol{u}, w, t) V_\tau^*(\boldsymbol{u}', w', t')\rangle}{\langle V_\tau(\boldsymbol{u}, 0, t) V_\tau^*(\boldsymbol{u}', 0, t')\rangle}. \tag{30}$$

The weights are defined with respect to the H I cross-correlation computed in Equation (14) for $\boldsymbol{u} = \boldsymbol{u}'$, $w = w' = 0$, $t = t'$, and
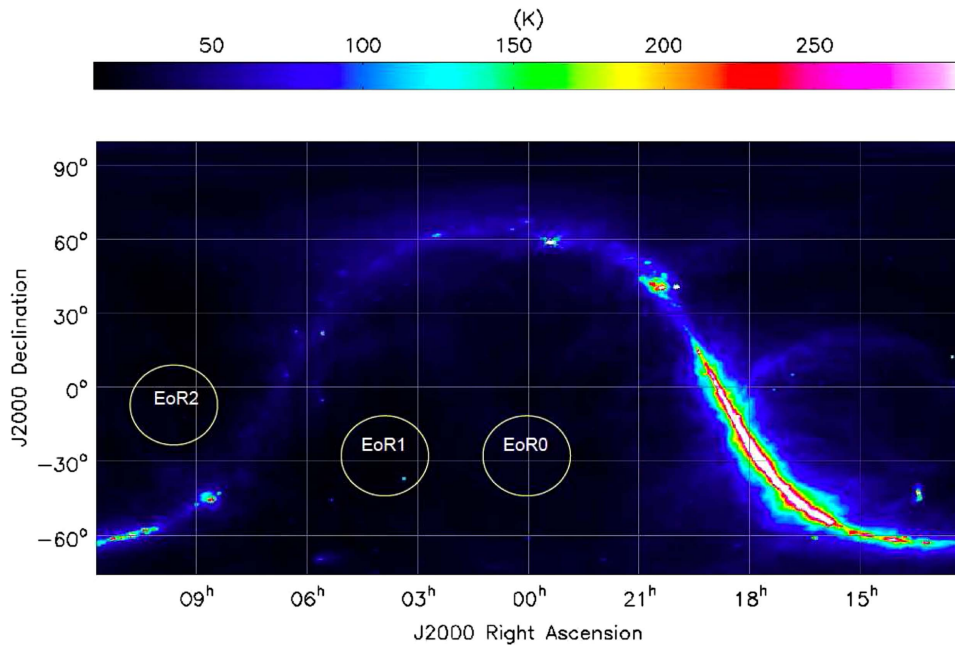
Figure 4. EoR fields are shown on the Galactic map at 408 MHz (Haslam et al. 1982). Three fields, away from the Galactic plane, have been identified for the MWA reionization study. These are EoR0, EoR1, and EoR2 respectively as shown in circles. In this paper we focus on the EoR1 field centered at R.A. = 4^h, decl. = $-27°$.

$\tau = \tau'$. We only consider the case $\tau' = \tau$ for the computation of weights.

Using Equation (30) allows one to recover the H I power spectrum for a fixed wavenumber k from visibility cross-correlations.

3. ANALYSIS OF MWA DATA

MWA is a low frequency radio interferometer array located in Western Australia. It consists of 128 antenna tiles with each tile comprising 16 crossed dipole antennas over a metal ground screen in a 4×4 grid. MWA bandwidth is 30.72 MHz, divided into 24 coarse channels of width 1.28 MHz each. The total bandwidth is divided into 768 fine channels. With the use of an analog beamformer appropriate phase delays are introduced in each individual dipole antenna to track the pointing center of the beam across the sky. For more information on MWA, please see Tingay et al. (2013) and Lonsdale et al. (2009).

To minimize the effect of Galactic synchrotron emission, the MWA EoR community has chosen three fields on the sky away from the Galactic plane. These fields have been named as EoR0, EoR1, and EoR2 and are shown in Figure 4. In this paper we present 3 hr of tracking analysis of the EoR1 field centered at R.A. = 4^h, decl. = $-27°$.

Many research groups are currently developing pipelines to extract statistical information from radio interferometric data, with an aim to detect the H I signal from EoR (Dillon et al. 2015; Jacobs et al. 2016; Trott et al. 2016; B. J. Hazelton et al. 2016, in prepration). These can be divided broadly into two categories: image based and visibility based pipelines. For foreground subtraction and imaging these pipelines use the following imaging algorithms: Real Time System (Mitchell et al. 2008; Ord et al. 2010) and Fast Holographic Deconvolution (Sullivan et al. 2012).

The image based pipelines (Dillon et al. 2015; B. J. Hazelton et al. 2016, in prepration) use a source catalog created through the deconvolution of the data which are subtracted to obtain a residual image cube. The Fourier transform of this image cube

with some further processing yields the power spectra. On the other hand, the visibility based pipelines (Trott et al. 2016) use the data in the visibility domain for power spectra estimation, after initial processing in the image domain to obtain the foreground model. A detailed comparison of the outputs from all the methods described above is provided in Jacobs et al. (2016). Thyagarajan et al. (2015a, 2015b) describe the impact of a wide field of view in power spectra estimation.

A special variant among the visibility based estimators is the "delay spectrum" (Parsons et al. 2012b, 2014; Pober et al. 2013), which directly Fourier transforms each calibrated visibility along its frequency channels. The radio interferometer PAPER uses this approach extensively; it relies upon redundant baselines to calibrate the interferometer and east–west and near east–west baselines for power spectrum estimation (Parsons et al. 2012b, 2014; Ali et al. 2015). This particular scheme has been discussed and implemented only for redundant drift scan observations. In this paper we explore the possibility of applying this approach for non-redundant imaging arrays and tracking measurements.

In this and the next section, we discuss in detail our method of MWA data analysis and power spectrum estimation from the data.

We summarize below the major ingredients of the method and then describe each of the steps in detail in subsequent sections:

1. CASA (McMullin et al. 2007) is used for initial processing of the data to calibrate raw visibility measurements. This is followed by the creation of a model sky image from clean components. This model is then subtracted in the visibility domain to obtain residual visibilities. We use both the calibrated and residual visibilities for computing the power spectrum.

2. Each visibility is then Fourier transformed in frequency space (Equation (31)). This process is needed to isolate foregrounds in the $k_\perp$–$k_\parallel$ plane. We note the our method utilizes both the subtraction of foregrounds and their

isolation, but it does not employ an external point source catalog.

3. The procedure outlined above yields complex visibilities as a function of five variables: $V_\tau(u, v, w, t)$. For computing the power spectrum we cross-correlate these visibilities for $t' \neq t$ to remove the noise bias. To weigh each cross-correlation we assume that there exist regions in the $k_\perp - k_\parallel$ plane which are dominated by only noise and the H I signal. This allows us to compute a weight for each cross-correlation based on the expected H I signal. For computing these weights we take into account the impact of the w-term and the distortion of intensity pattern in a tracking scan. The relevant method is elaborated in detail in Sections 2 and 2.1–2.3 and summarized in Section 4.

4. In Section 4.1, we describe the power spectrum estimator, taking into account weights given by the expected H I signal, in three, two, and one dimensions. We also discuss our method to compute the errors on the estimated power spectrum.

3.1. CASA Processing

MWA data were collected at 2 minute intervals with a time resolution of 0.5 s and frequency resolution of 40 kHz. The central frequency of these observations is 154.24 MHz. For preprocessing we have used the Cotter pipeline (Offringa et al. 2015) to average to 10 s of integration; we have not performed any averaging over the frequency channels. Cotter also uses the in-built AOFlagger to flag and remove radio frequency interference. The edge channels of each coarse band are flagged with Cotter due to aliasing effects. After this preprocessing the Cotter pipeline delivers the data in the CASA readable "Measurement set (ms)" format for further processing.

Once the "ms" files are produced for each 2 minute data set, we process each of these 2 minute data sets in CASA to produce an image. The Hydra A source is used to calculate the bandpass solutions which are applied to the uncalibrated data. We next construct a sky model from these data so that we can subtract it to obtain the residual visibility. After bandpass calibration the first round of "clean" is applied on each 2 minute data set. The multi-scale multi-frequency synthesis algorithm (Rau & Cornwell 2011) has been used for imaging. We have created images of size 3072×3072 pixels with 1 arcmin cell size using Cotton–Schwab CLEAN (Schwab 1984) with a uniform weighting scheme. After the first round of CLEAN we perform self-calibration (both phase and amplitude +phase) and apply the CLEAN loop until the rms value of the residual image converges. The threshold limits for the CLEAN steps are chosen to be 5 sigma. The W-projection algorithm (Cornwell et al. 2008) is also used to correct for the errors arising due to non-coplanarity of baselines. Once we obtain the best model of the sky for each 2 minute observation, the model visibilities are then subtracted from the calibrated data using the UVSUB algorithm in CASA to obtain the residual data. This process is followed for both XX and YY polarizations separately. A flow chart of the data pipeline is shown in Figure 5. In Figure 6, we present a sample image of 2 minute deconvolution.

As noted above we process the data for only 2 minutes to ensure the primary beam does not substantially change during
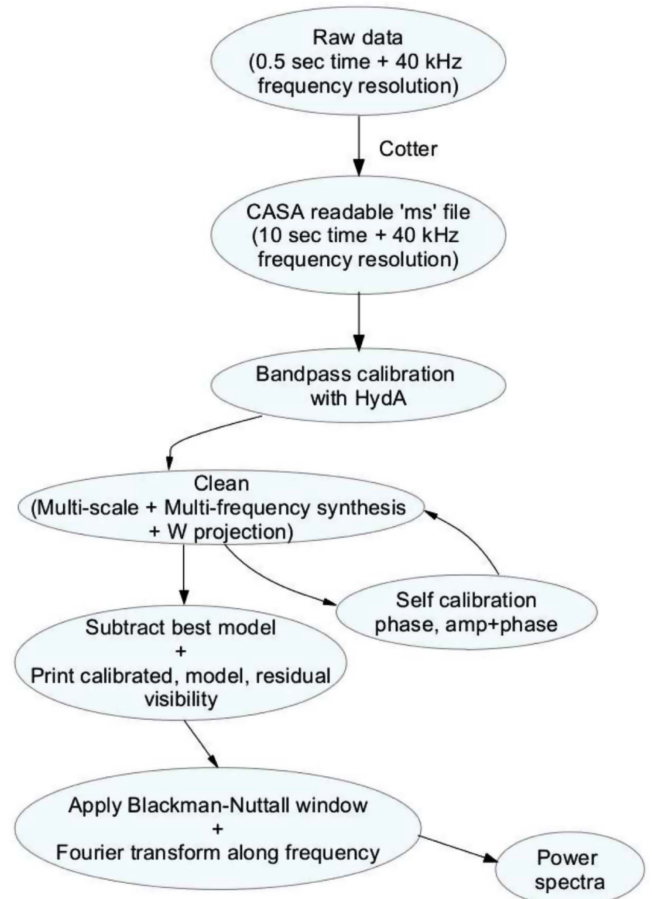

Figure 5. Schematic of the power spectra pipeline.

the run. For a 2 minute scan we obtain an rms of nearly 40 mJy beam^{-1}.

The residual visibility $V_\nu(u_\nu, v_\nu, w_\nu, t)$ is a function of five variables. We compute the discrete Fourier transform of the residual visibilities in the frequency space weighted by the Blackman–Nuttall (Nuttall 1981) window B_ν to suppress leakage into the EoR window (Thyagarajan et al. 2013, 2016):

$$V_\tau(u, v, w, t) = \sum \Delta \exp(i2\pi\nu\tau) V_\nu(u_\nu, v_\nu, w_\nu, t) B_\nu. \quad (31)$$

Notice that in Equation (31) the frequency dependence of the baseline vector $\boldsymbol{b}_\nu = \{u_\nu, v_\nu, w_\nu\}$ is integrated over. Therefore, the labels $\{u, v, w\}$ on the lhs of Equation (31) need further explanation. As noted above (in the discussion following Equation (11)) they can be chosen to denote a given baseline vector at a fixed frequency ν_0. We choose this frequency to be the central frequency of the band $\nu_0 = 154$ MHz. Parsons et al. (2012a, 2012b) provide detailed implications of the frequency dependence of the baseline vector. Here $\Delta = 40$ kHz and 256 channels are used for our study, which correspond to a total bandwidth 10.24 MHz in the frequency range 149.09–159.34 MHz.

4. POWER SPECTRUM

The visibilities (Equation (31)) are cross-correlated with weights determined from the H I signal (Section 2.3) to estimate the power spectrum. For each pair of parameters,
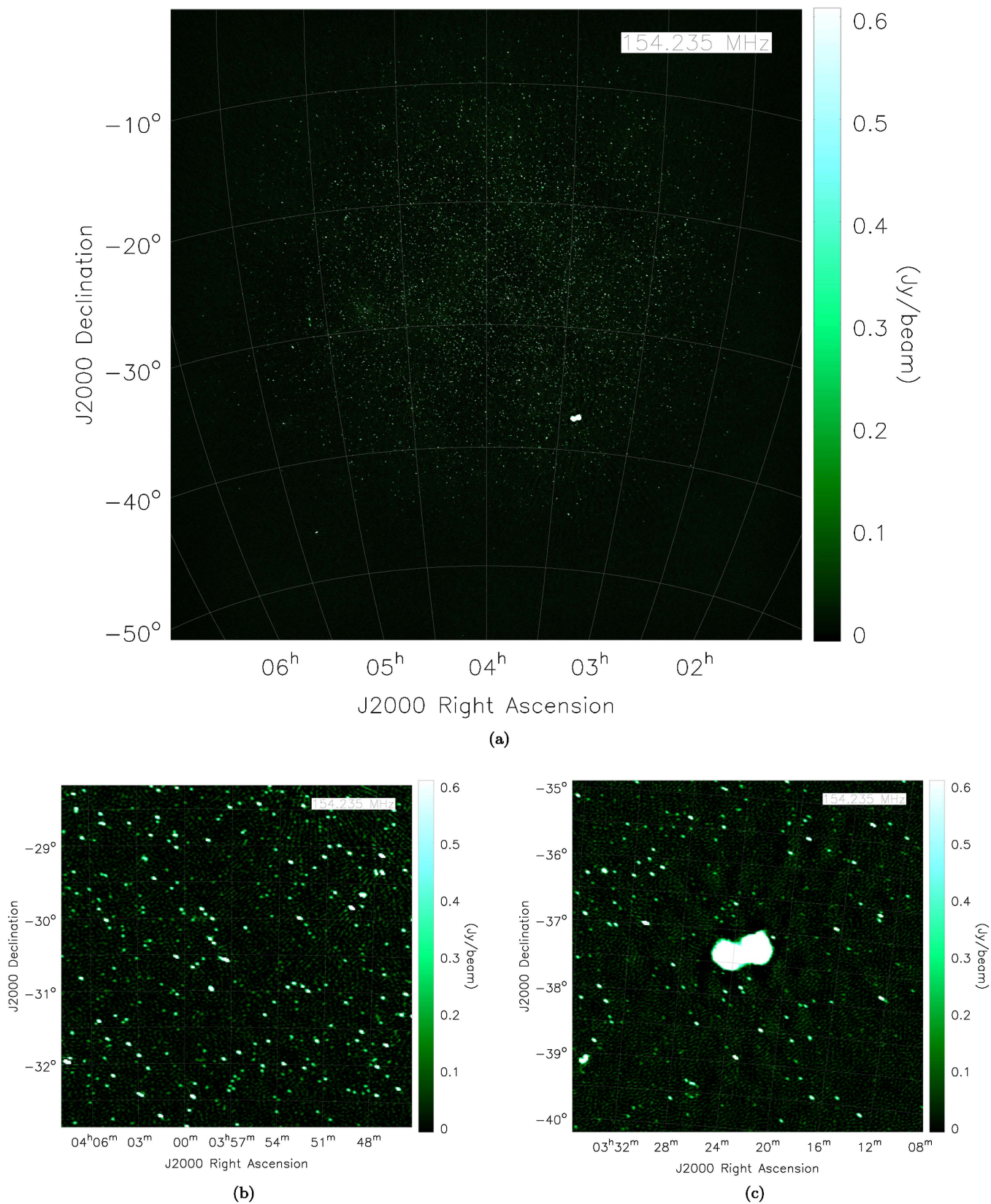
Figure 6. Image of the EoR1 field at 154.24 MHz in the top panel for two minutes of data. The bottom panels display two regions from the image. The right panel displays the region containing the Fornax A.

e.g., $\{u, u'\}$, the weights are generally different. It is computationally prohibitive to deal with weights for all cross-correlations. We make several simplifying assumptions to make the problem tractable based on the properties of the H I signal. In Sections 2, 2.1, and 2.2 we discuss in detail the H I signal and how it is affected by the inclusion of the w-term and

the additional complication arising from distortion of the field of view as a region is tracked for MWA.

We summarize the main results of these sections as applied to the data:

1. In Section 2 the H I signal and its correlations are discussed in detail. Equation (15) shows that correlations in the sky plane are nearly independent of correlations along the line of sight. This allows us to compute weights for correlations in the plane of the sky independent of the third axis. Equation (15) allows us to derive a relation between the measured correlation and the inferred H I power spectrum (Equation (17)). Equation (17) defines the scale of cross-correlation weights. The weight function $\mathcal{W}(u, v, w, t; u', v', w', t')$ is unity when $\boldsymbol{u} = \boldsymbol{u}'$, $w = w' = 0$, and $t \simeq t'$. Equation (17) refers to this case.

2. In Section 2.1, the impact of the w-term on the H I signal is computed. Equation (23) and Figure 2 capture the effect of non-zero w on H I correlations. The w-term diminishes the signal by shrinking the effective primary beam and increases the correlation length scale $|\boldsymbol{u} - \boldsymbol{u}'|$. We use the analytic expression based on Equation (23) for computing weights for $w > 30$.

3. In Section 2.2, we attempt to assess the impact of time-dependent distortion of the intensity pattern in a tracking run for MWA. Figure 3 shows the combined effect of the moving grid and w-term. The distortion of intensity pattern generally acts to enhance decorrelation but is found to be not significant and does not alter the main features of the signal. For our computation, we only update the weights after every 10 minutes to account for this effect.

4.1. Power Spectrum Estimator

As shown above, each correlation receives a different weight depending on the values of $\{u, v, w, t\}$ of the baselines being correlated. As noted, we define the weights $\mathcal{W}(u, v, w, t; u', v', w', t')$ such that they approach unity when $u = u'$, $v = v'$, $w = w' = 0$, and $t \simeq t'$ such that the effect of the moving grid is not important ($t \neq t'$ for all correlations to remove noise bias).

The H I power spectrum is a function of $k = \sqrt{k_\perp^2 + k_\parallel^2}$; Equation (16) gives the relation between the Fourier components of the H I signal and $\{\boldsymbol{u}, \tau\}$. All cross-correlations for which the wave vector lies in some range k and $k + dk$ can be used to construct the unbiased H I signal: $1/N \sum V_\tau(u, v, w, t) V_\tau(u', v', w', t')/\mathcal{W}$; where N is the number of all pairs for which k lies in the range specified above. However, this estimator, though unbiased for the H I signal, could be dominated by small values of weights $\mathcal{W}$, which does not make it the lowest noise (or optimal) estimator.

As the observed signal is dominated by noise, we consider an optimal estimator for our study:

$$\hat{P}_\tau(u, v) = \frac{1}{\sigma} \sum_{u', v', w, w', t, t'} (V_\tau V_\tau^* \times \mathcal{W}) \tag{32}$$

where $\sigma = \sum_{u', v', w, w', t, t'} \mathcal{W}^2(u, v, w, t, u', v', w', t')$. To avoid noise bias, $t \neq t'$ for all cross-correlations. For a given $\{u, v, \tau\}$, Equation (32) allows us to compute the power

spectrum by optimally weighing over all the cross-correlations. However, as Figures 2 and 3 show the correlations fall substantially for $\boldsymbol{u} - \boldsymbol{u}' \geqslant 1.5$ (see also Paul et al. 2014 and references therein). This motivates us to pixelize the $\boldsymbol{u}$-plane and consider only those visibility pairs for which the correlations are significant. We consider cells of different sizes and present results here for $\Delta u = \Delta v = 0.5$. The number of visibility measurements in a cell varies depending on the (u, v) values. The shortest baselines have higher population as expected for MWA. For 3 hr of analysis and $u, v < 50$, the number of visibilities in a cell lies in the range ~ 1000–3500 where each visibility has a time resolution of $\Delta t = 10$ s. All the cross-correlations within a cell are computed using Equation (32).

For averaging over different cells, each cell is assigned an average weight corresponding to the rms of the power spectrum for a cell, σ_p. These weights are then used for optimally averaging the power spectrum (Equation (32)) over other cells (for details see Appendix B). Note that this procedure allows us to separate large correlations of the H I signal, the ones for which $\mathcal{W}$ is close to unity, from the ones which are expected to be incoherent because $\mathcal{W}$ is small.

The schematic of the two processes—the computation of power spectrum in three and two dimensions—is displayed in Figure (7): the top panel delineates the process of computing cross-correlations within each cell and the bottom panel depicts how azimuthal average for a fixed baseline length $\sqrt{u^2 + v^2}$ is computed. For MWA data, $k_\parallel \gg \boldsymbol{k}_\perp$, which means the value of k is dominated by the value of $k_\parallel$. This suggests the following method for computing the 1D power spectrum, which we adopt: all the cells for a given τ are optimally averaged using the method described above. This procedure yields a complex number. In the figures that display 2D and 1D power spectra we plot the absolute value of the estimated power spectrum.

The power spectrum error in one dimension is computed by first estimating the rms for each cell σ_p. σ_p are then used as weights for optimal averaging over all the cells for a fixed τ. The resultant rms after averaging over the cells approaches $\sigma_{\text{fin}} \simeq (1/\sum \sigma_p^2)^{-1/2}$ if the power spectrum across cells is uncorrelated. This holds for noise but, as noted above, is an approximation for the H I signal. We expect this assumption to be valid in our case as the observed signal is dominated by noise (for a detailed explanation, see Appendix B).

5. RESULTS

In Figure 8, the power spectra computed from 3 hr of data are shown in the $k_\perp$–$k_\parallel$ plane. The power spectra for both XX and YY polarizations are shown for calibrated and residual visibilities.

We first discuss discernible features in the power spectra:

1. In Fourier space the foreground contributions occupy a wedge shaped region (also called the "foreground wedge") owing to the smooth spectral characteristics of foreground sources (Datta et al. 2010; Parsons et al. 2012b; Vedantham et al. 2012; Thyagarajan et al. 2013, 2015a; Dillon et al. 2014; Liu et al. 2014a). The region beyond the foreground wedge is expected to be free from foreground contamination and dominated by thermal noise and expected H I signal. This foreground isolation approach
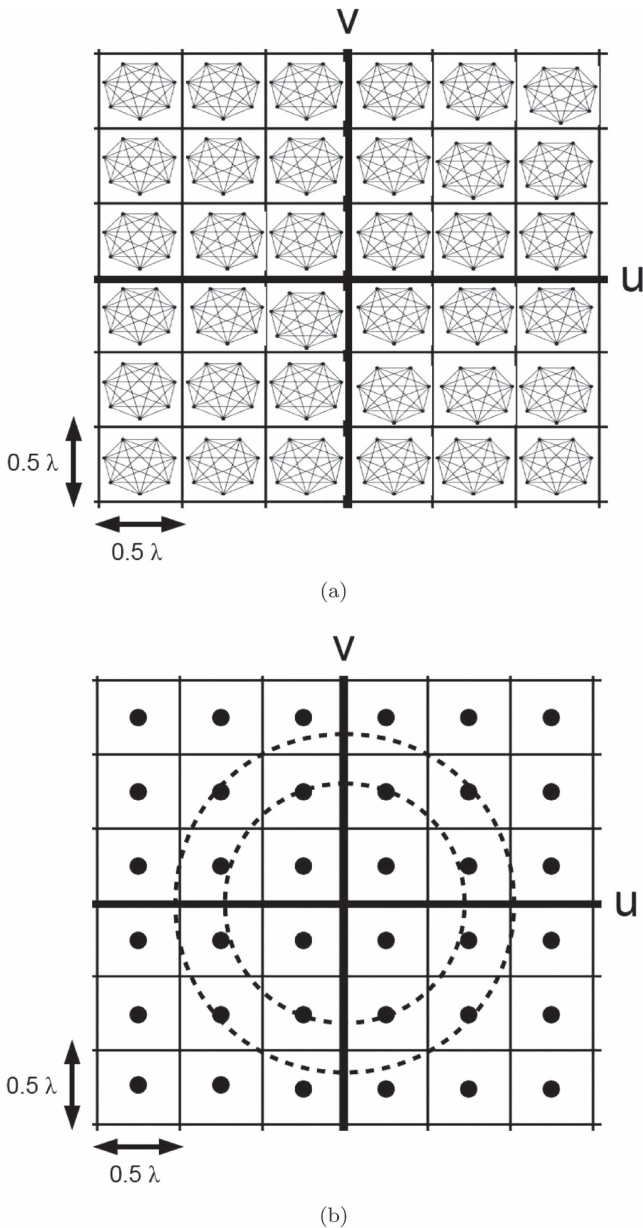
V

u

0.5 λ

0.5 λ

(a)

V

u

0.5 λ

0.5 λ

(b)

Figure 7. (a) The population of visibilities within different *uv* bins for each τ. The small dots denote visibilities due to individual baselines. These are cross-correlated with each other within a given *uv* bin. (b) The black dots within a bin are the power value at each bin. An optimal azimuthal average, based on weights of each cell, is done to collapse the *uv* axes into a single axis of baseline length.

is particularly useful for MWA as it has low angular resolution. The strongest H I signals lie in the shortest baselines (low $k_\perp$ values) and decrease rapidly with increasing $k_\perp$ values. The upper and lower panels of Figure 8 display the dirty (calibrated with no foreground subtraction) and the residual (clean components subtracted) power spectra, respectively. Figure 8 bears out the assumption that foregrounds have smooth spectral characteristics as they are seen to form the "foreground wedge," this separation is in good agreement with the expectation from foreground simulations (Figure 10). The first few $k_\parallel$ modes exhibit maximum foreground

contributions, the $k_\parallel = 0$ mode being the strongest. The amplitude at this mode is roughly 10^{14}–10^{15} mk^2 (Mpc/h)^3 which is in good agreement with the results of other MWA EoR pipelines (Jacobs et al. 2016). A clear decrement in power in the "foreground wedge" is visible in residual power spectra as compared to the dirty one.

2. As described in the previous section, MWA has missing channels on either side of coarse bands of width 1.28 MHz. This leads to a periodicity of missing data across the frequency axis in visibility, the effect of which is reflected in the Fourier-transformed power spectra as the horizontal bright lines at fixed $k_\parallel$.

Figure 9 shows the 1D power spectra; obtained from regions that exclude the foreground wedge and bright coarse bands in Figure 8. For computing the 1D power spectrum, the foreground wedge and the bright coarse horizontal bands of the 2D power spectrum (Figure 8) are rejected. More specifically, the channels corresponding to $k_\parallel \leqslant 0.14$ h Mpc^{-1} and $k_\perp > 0.1$ h Mpc^{-1} are not considered. For each coarse band, the central brightest channel along with two channels on either side are excluded. The remaining contiguous regions are used in estimation of the 1D power spectrum. For instance, for a given $k_\parallel$ that meets the criterion outlined above, all the cells that correspond to $k_\perp < 0.1$ h Mpc^{-1} are used for the computation of the 1D power spectrum. The errors in the binned power spectra are computed using a scheme outlined in Appendix B.

The 2D power spectra obtained from data (Figure 8) can be compared with Figure 10 that shows the expected power spectrum based on simulations of foregrounds and noise. In particular, this comparison allows us to assess the structure of coarse channels and the foreground wedge. It also indicates the range of scales of the 2D power spectrum. However, Figure 10 is based on a single realization of noise and a model of foregrounds based on a random distribution of point sources, and therefore a more detailed comparison between the data and simulations is not possible. We shall return to this detailed comparison in future work.

6. CONCLUSIONS

In this paper, we propose a new method to extract the H I power spectrum from MWA visibility data in delay space. The proposed method is applicable when a region is tracked using imaging radio interferometers.

One of the crucial factors in power spectrum estimation is how the *w*-term is dealt with within the pipeline. Our findings are that the *w*-term causes an effective shrinking of the primary beam which reduces the contribution of the H I signal. We carefully model the H I signal by taking the *w*-term into account, the weights calculated are then applied to cross-correlate the measured visibilities. Moreover, the cross-correlation approach is particularly useful to minimize various systematics in the system. We also model and account for the impact of changing intensity pattern in a tracking run. We find this effect to be sub-dominant to the *w*-term correction.

We analyze 3 hr of MWA data from the EoR1 field, one of the fields identified by the MWA community for EoR science. CASA has been used for calibration and to create a foreground model using the clean components. Both the dirty (calibrated with no foreground subtraction) and residual (foreground
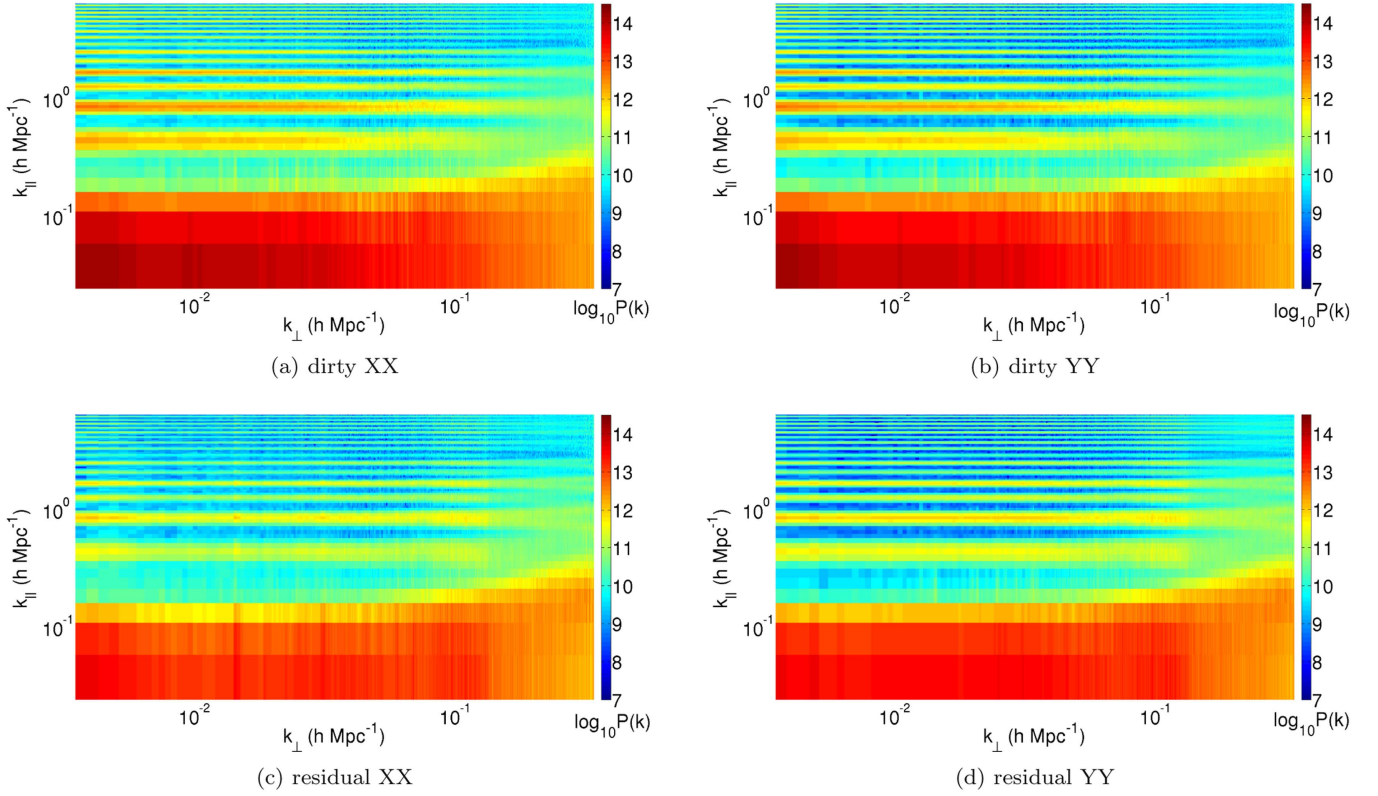
(a) dirty XX   (b) dirty YY

(c) residual XX   (d) residual YY

Figure 8. Two-dimensional power spectra for 3 hr of data on the EoR1 field for XX and YY polarization. The power is plotted as $\log_{10} P$ where P is in units $mK^2 (\mathrm{Mpc/h})^3$.
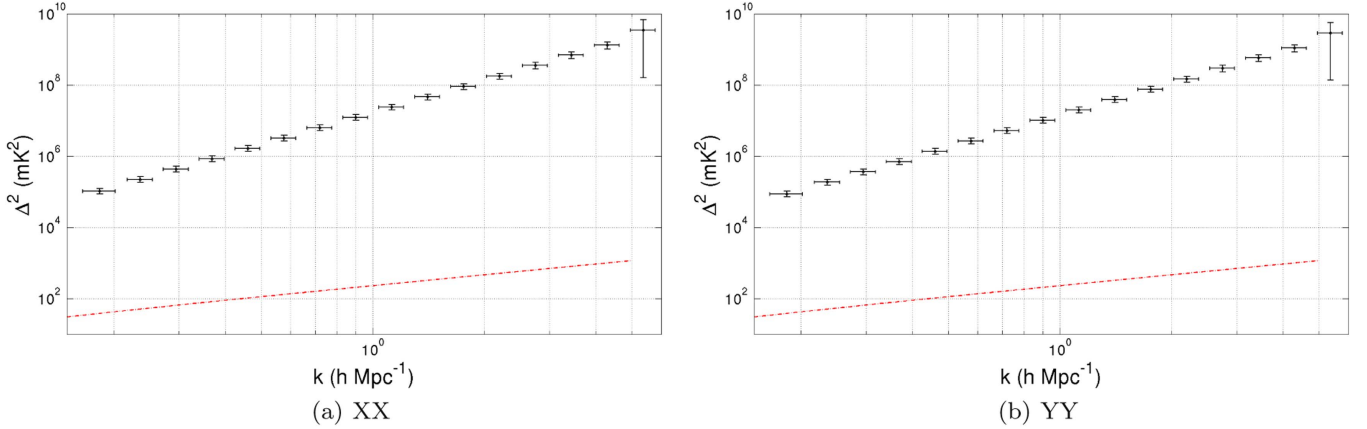

(a) XX   (b) YY

Figure 9. One-dimensional power spectra ($\Delta^2 = k^3 P(k)/(2\pi^2)$ in units $(mK)^2$) for the XX (left panel) and YY (right panel) polarization are shown along with the errors for each band. The dotted–dashed (red) curve show the expected H I signal (Furlanetto et al. 2006).

model subtracted) power spectrum in delay space are presented. Our results (Figures 8 and 9) are in good agreement with analyses of other MWA EoR pipelines (Jacobs et al. 2016).

In the future we plan to apply the method proposed here in cases of more integration time and, in particular, to a longer single tracking run. The decorrelation caused by the w-term and the changing intensity pattern should be more dominant in the latter case. This will allow us to test the efficacy of our method for more extreme cases and might indicate the best possible way of detecting the H I signal from the Eor.
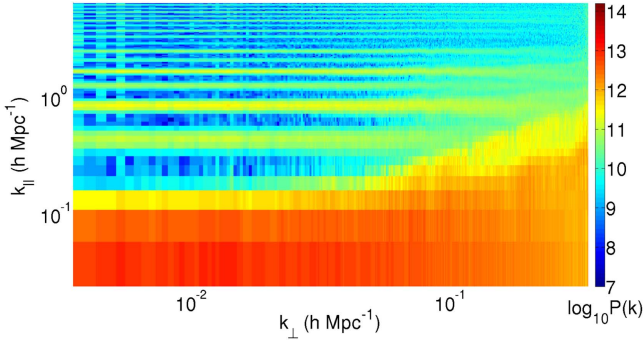
13

Figure 10. Results of the simulation of noise and foregrounds are shown. The power spectrum pipeline developed to analyze the data is also applied to simulated foregrounds and noise for a 3 hr observational run for a single polarization.

APPENDIX A
FOREGROUNDS AND NOISE SIMULATIONS

The primary contribution to foregrounds come from spectrally smooth point and diffuse sources. They differ from the H I signal in both spatial and spectral behavior. However, it is the latter difference that allows us to potentially isolate foregrounds from the H I signal in the power spectrum estimation.

To understand the impact of foregrounds in the data, we model them as a set of point sources. We note that if both the point and the diffuse sources have smooth spectra across the instrumental bandwidth, their impact on the power spectra are similar and therefore point sources allow us to capture adequately our ability to isolate foregrounds from the signal. In this section, for analytic work, we assume $w = 0$. We note without further proof that this assumption does not alter our main inferences.

For a set of point sources, the intensity distribution is given by

$$I_\nu(\boldsymbol{\theta}) = \sum_i F_{i\nu}\delta_D^2(\boldsymbol{\theta} - \boldsymbol{\theta}_i). \tag{33}$$

Here $F_{i\nu}$ and $\boldsymbol{\theta}_i$ are the source fluxes and positions, respectively. This allows us to compute the visibility for a given baseline $\boldsymbol{u}_\nu$ and frequency ν:

$$V_\nu(\boldsymbol{u}_\nu) = \sum_i F_{i\nu} \exp(i2\pi\boldsymbol{u}_\nu.\boldsymbol{\theta}_i)A_\nu(\boldsymbol{\theta}_i). \tag{34}$$

Here A_ν is the primary beam. As discussed earlier, we also define a visibility in the conjugate space by taking the Fourier transform with respect to ν (Equation (11)): $V_\tau(\boldsymbol{u})$. Our aim is to compute the correlation of this visibility such that

$$\langle V_\tau(\boldsymbol{u}) V_{\tau'}^*(\boldsymbol{u}')\rangle = \sum_i\sum_j \int d\nu \int d\nu' F_{i\nu} F_{j\nu'}$$
$$\times \exp(i2\pi[\boldsymbol{u}_\nu.\boldsymbol{\theta}_i - \boldsymbol{u}'_{\nu'}.\boldsymbol{\theta}_j + (\tau\nu - \tau'\nu')]$$
$$\times A_\nu(\boldsymbol{\theta}_i)A_{\nu'}(\boldsymbol{\theta}_j). \tag{35}$$

Both the source flux F_ν and primary beam $A_\nu(\boldsymbol{\theta})$ are functions of frequency so it is difficult to analytically compute this expression. However, assuming smooth and small variations of both of these quantities across the bandwidth, we can make meaningful analytic estimates; we verify this assumption from detailed simulations and the analysis of the data. The main frequency variation in this case comes from the phase of the integral (the terms in the exponent) and in particular from the change in the baseline length as frequency changes. We note here that multiple correlations are available to us for this analysis for different pairs of $\{\tau, \tau'\}$ and $\{\boldsymbol{u}, \boldsymbol{u}'\}$. Here we assume $\tau = \tau'$.

Delay space—foreground wedge. Here we expand the same baseline in frequency space: $\boldsymbol{u}'_{\nu'} = \boldsymbol{u}_\nu + d\boldsymbol{u}_\nu/d\nu(\nu' - \nu)$. In this case, $\boldsymbol{u}_\nu = \nu/\nu_0\boldsymbol{u}_{\nu_0}$, where ν_0 is some fixed frequency. Making the simplifying assumption that both point source fluxes and the primary beam are independent of frequency, Equation (35) can be analytically integrated. We further make coordinate transformations $x = (\nu' - \nu)/2$ and $y = (\nu' + \nu)/2$, and assume $\nu' = \nu$ in all the quantities except those in the exponent containing their difference, which allows us to use $y \simeq \nu$:

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle \simeq \sum_i\sum_j \int dx \int dy F_{i\nu} F_{j\nu}$$
$$\times \exp(i2\pi[\boldsymbol{u}_\nu.\boldsymbol{\theta}_i - \boldsymbol{u}_\nu.\boldsymbol{\theta}_j$$
$$+ 2(d\boldsymbol{u}_\nu/d\nu.\boldsymbol{\theta}_i - \tau)x])A_\nu(\boldsymbol{\theta}_i)A_\nu(\boldsymbol{\theta}_j). \tag{36}$$

Integrals over x and y are now separated which gives us

$$\langle V_\tau(\boldsymbol{u}) V_\tau^*(\boldsymbol{u}')\rangle \propto \int dx \exp[-i2\pi(d\boldsymbol{u}_\nu/d\nu.\boldsymbol{\theta}_i - \tau)x]. \tag{37}$$

As noted above, $d\boldsymbol{u}_\nu/d\nu = \boldsymbol{u}_{\nu_0}/\nu_0$, or it is independent of frequency. The integral in the equation is insignificant only when $\tau \simeq d\boldsymbol{u}/d\nu.\boldsymbol{\theta}_i$. This linear relation between τ and the baseline $\boldsymbol{u}$ gives a region bounded by a "wedge" in the τ–$d\boldsymbol{u}/d\nu.\boldsymbol{\theta}_i$ space for a spatial distribution of point sources (e.g., see Datta et al. 2010; Parsons et al. 2012b; Vedantham et al. 2012; Thyagarajan et al. 2013, 2015a; Dillon et al. 2014; Liu et al. 2014a).

Another possible way to understand the nature of spectrally smooth foregrounds is to first compute the correlation in the frequency space. Using Equation (34) gives us

$$\langle V_\nu(\boldsymbol{u}_\nu) V_{\nu'}^*(\boldsymbol{u}'_{\nu'})\rangle = \sum_i \sum_j F_{i\nu} F_{j\nu'}$$
$$\times \exp(i2\pi[\boldsymbol{u}_\nu.\boldsymbol{\theta}_i - \boldsymbol{u}'_{\nu'}.\boldsymbol{\theta}_j])$$
$$\times A_\nu(\boldsymbol{\theta}_i)A_{\nu'}(\boldsymbol{\theta}_j). \tag{38}$$

Using $\boldsymbol{u}'_{\nu'} = \boldsymbol{u}_\nu + \Delta\nu\boldsymbol{u}_{\nu_0}/\nu_0$ and substituting into Equation (38), and performing a single Fourier transform with respect to $\Delta\nu$, we recover the main expected feature of the foreground "wedge" described above. Computationally, if the variation of other quantities with frequency, primary beam, and source fluxes, is neglected, this method is completely equivalent to the one based on Equation (36).

Even though we used a set of point sources, the main inferences of the analysis also follow for diffuse sources. In particular, the frequency space integrals used to prove our case are exactly the same.

For our simulations, we assume a set of point sources isotropically distributed with fluxes above 1 Jy at 150 MHz. We construct this flux distribution from the radio source count at 1.4 GHz, which is given by (Hopkins et al. 2003)

$$\log\left[\frac{dn/ds}{S^{-2.5}}\right] = \sum_{i=0}^{6} a_i \left[\log\left(\frac{S}{\text{mJy}}\right)\right]^i \qquad (39)$$

for flux range $0.05\,\text{mJy} \leqslant S \leqslant 1000\,\text{mJy}$. The constants are $a_0 = 0.859$, $a_1 = 0.508$, $a_2 = 0.376$, $a_3 = -0.049$, $a_4 = -0.121$, $a_5 = 0.057$, and $a_6 = -0.008$. We simulate sources over the entire hemisphere (nearly 15,000 sources) to suitably take into account the contribution from MWA primary beam sidelobes. We extrapolate the distribution to the frequencies of interest to us by assuming a spectral index $\alpha = -0.7$.

A.1. Thermal Noise

Thermal noise is independent of the baseline and depends on three parameters: system temperature, integration time, and channel width. The rms of thermal noise associated with a visibility measurement for channel width $\Delta\nu$ and integration time Δt is

$$\sigma(\nu) = \frac{T_{\text{sys}}}{K\sqrt{\Delta\nu\Delta t}} \qquad (40)$$

where T_{sys} and K denote the system temperature and antenna gain respectively. For MWA, $K = A_{\text{eff}}/(2k_B)$ with $A_{\text{eff}} = 21.5\,\text{m}^2$ for MWA at $\nu = 150\,\text{MHz}$ (Tingay et al. 2013). In our analysis we choose $\Delta\nu = 40\,\text{kHz}$, $\Delta t = 10\,\text{s}$, which are very small compared to the frequency and time coherence of the signal (Paul et al. 2014). The system temperature has two components: sky temperature (dominant source of noise at low frequency) and receiver temperature. We consider $T_{\text{sys}} = 250\,\text{K}$ for a single polarization which is consistent with the reported system temperature at 154.24 MHz for the MWA pointing we consider in this paper.

It is fair to assume that the thermal noise for a radio interferometer follows Gaussian statistics with zero mean. In our simulation, we follow the same pipeline used for analyzing the real data to estimate the thermal noise power. We use the baseline distribution from the observation with $u_{\text{max}} = v_{\text{max}} = 250$. For every (u, v) point the noise is drawn from a Gaussian distribution with zero mean and the rms given by Equation (40).

APPENDIX B
POWER SPECTRUM ESTIMATION

As discussed in Section 4.1, the power spectrum from the data is computed in two stages. First the power spectrum and its rms are computed for a single cell in which the H I signal is expected to be near coherent and then an average is obtained across cells assuming the H I signal to be incoherent for different cells (Figure 7).

As noted in the text, the H I signal can be recovered from a visibility cross-correlation by inverse weighting with $\mathcal{W}$. We denote such a cross-correlation as $\hat{S} \equiv VV/\mathcal{W}$; $\hat{S}$ is generally a complex number. For optimal averaging to get the lowest noise estimator, one needs to sum over these cross-correlations by inverse weighting with the square of the rms of each cross-correlation σ_i. For pure noise, $\sigma_i \propto 1/\mathcal{W}$ and one can obtain Equation (32). Notice that this estimator is invariant under an

overall scaling of σ_i. The error in power spectrum for each cell σ_p is $\sigma_p^2 = \langle\hat{S}^2\rangle - \langle\hat{S}\rangle^2$, where the average is obtained optimally from the data for all the cross-correlations. It can be shown that if each cross-correlation is assumed to be uncorrelated, as would be the case for pure noise, $\sigma_p \simeq (1/\sum \sigma_i^2)^{-1/2}$. Notice that if the rms for all the cross-correlations is the same, as would be the case if all measurements are equally weighted, then this expression reduces to $\sigma_p = \sigma_i/\sqrt{N_c}$, where N_c is the number of all the cross-correlations within a cell.

This procedure yields an estimate of the power spectrum (Equation (32)) and its error σ_p for each cell.

For averaging over cells, we repeat the procedure described above by taking the estimated power spectrum for a cell as the signal and σ_p as the weights. This allows us to estimate the 2D and 1D power spectrum and its rms. For pure noise, the final error in the power spectrum is expected to approach $\sigma_{\text{fin}} = \sigma_p/\sqrt{N}$, where N is the number of cells used for obtaining the average.

We briefly discuss some shortfalls of such a procedure. First, we do not construct the covariance matrix of the power spectrum estimator. We only estimate its diagonal terms, σ_p^2 and σ_{fin}^2. This means that we are not able to assess the extent of cross-correlation between two neighboring bins in Figure 9. Such cross-correlation might contain important information about systematic errors, foreground leakage, and H I signal and noise cross-correlation.

Second, we do not include the H I signal in our estimation procedure. This is justified for the present work as the observed signal is clearly dominated by noise and foreground residuals (Figure 9). We briefly assess the impact of the H I signal in computing the power spectrum error.

We assume the following estimator for computing the power spectrum for a cell and consider the contribution of only the H I signal:

$$\hat{S} = \frac{1}{N_c} \sum_{ij} \frac{V_i V_j}{\mathcal{W}_{ij}}. \qquad (41)$$

As noted above, this estimator allows us to recover the H I signal. The subscripts ij correspond to a pair for visibilities and the sum is carried over all the cross-correlations. After further computation, we obtain the error on the signal as

$$\Delta\hat{S}^2 = \frac{\hat{S}^2}{N_c^2} \sum_{kl} \sum_{ij} \left[\frac{\mathcal{W}_{ik}\mathcal{W}_{jl} + \mathcal{W}_{il}\mathcal{W}_{jk}}{\mathcal{W}_{ij}\mathcal{W}_{kl}}\right]. \qquad (42)$$

If all the weights are unity this reduces to the usual cosmic variance expression $\Delta\hat{S} = \sqrt{2}\hat{S}$. Even though this term is negligible for our purposes, this would need to be included for longer integration times.

REFERENCES

Ali, S. S., Bharadwaj, S., & Chengalur, J. N. 2008, MNRAS, 385, 2166
Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, ApJ, 809, 61
Barkana, R., & Loeb, A. 2001, PhR, 349, 125
Bharadwaj, S., & Ali, S. S. 2005, MNRAS, 356, 1519
Bharadwaj, S., & Sethi, S. K. 2001, JApA, 22, 293
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183
Chapman, E., Abdalla, F., Harker, G., et al. 2012, MNRAS, 423, 2518
Chapman, E., Abdalla, F. B., Bobin, J., et al. 2013, MNRAS, 429, 165

Christiansen, W. N., & Hogbom, J. A. 1969, Radiotelescopes (Cambridge: Cambridge Univ. Press)
Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, ISTSP, 2, 647
Datta, A., Bowman, J. D., & Carilli, C. L. 2010, ApJ, 724, 526
Datta, K. K., Roy Choudhury, T., & Bharadwaj, S. 2007, MNRAS, 378, 119
Dillon, J., Liu, A., & Tegmark, M. 2013, PhRvD, 87, 43005
Dillon, J. S., Liu, A., Williams, C. L., et al. 2014, PhRvD, 89, 023002
Dillon, J. S., Tegmark, M., Liu, A., et al. 2015, PhRvD, 91, 123011
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Harker, G., Zaroubi, S., Bernardi, G., et al. 2009, MNRAS, 397, 1138
Haslam, C. G. T., Stoffel, H., Salter, C. J., & Wilson, W. E. 1982, A&AS, 47, 1
Hazelton, B. J., Morales, M. F., & Sullivan, I. S. 2013, ApJ, 770, 156
Hobson, M. P., Lasenby, A. N., & Jones, M. 1995, MNRAS, 275, 863
Hopkins, A. M., Afonso, J., Chan, B., et al. 2003, AJ, 125, 465
Jacobs, D. C., Hazelton, B. J., Trott, C. M., et al. 2016, ApJ, 825, 114
Jelic, V., Zaroubi, S., Labropoulos, P., et al. 2008, MNRAS, 389, 1319
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Liu, A., Parsons, A. R., & Trott, C. M. 2014a, PhRvD, 90, 023018
Liu, A., Parsons, A. R., & Trott, C. M. 2014b, PhRvD, 90, 023019
Liu, A., & Tegmark, M. 2011, PhRvD, 83, 103006
Lonsdale, C. J., Cappallo, R. J., Morales, M. F., et al. 2009, IEEEP, 97, 1497
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127
McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, ApJ, 653, 815
Mitchell, D. A., Greenhill, L. J., Wayth, R. B., et al. 2008, ISTSP, 2, 707
Morales, M. F. 2005, ApJ, 619, 678
Morales, M. F., Hazelton, B., Sullivan, I., & Beardsley, A. 2012, ApJ, 752, 137
Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7
Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127
Nuttall, A. H. 1981, ITASS, 29, 84
Offringa, A. R., Wayth, R. B., Hurley-Walker, N., et al. 2015, PASA, 32, 8
Ord, S. M., Mitchell, D. A., Wayth, R. B., et al. 2010, PASP, 122, 1353
Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639

Parsons, A., Pober, J., et al. 2012a, ApJ, 753, 81
Parsons, A. R., & Backer, D. C. 2009, AJ, 138, 219
Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, ApJ, 788, 106
Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012b, ApJ, 756, 165
Paul, S., Sethi, S. K., Subrahmanyan, R., et al. 2014, ApJ, 793, 28
Pen, U. L., Chang, T. C., Hirata, C. M., et al. 2009, MNRAS, 399, 181
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 751, A16
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, arXiv:1502.01589
Planck Collaboration, Ade, R., Aghanim, N., et al. 2016, arXiv:1605.03507v2
Pober, J. C., Parsons, A. R., Aguirre, J. E., et al. 2013, ApJL, 768, L36
Presley, M. E., Liu, A., & Parsons, A. R. 2015, ApJ, 809, 18
Rau, U., & Cornwell, T. J. 2011, A&A, 532, A71
Schwab, F. R. 1984, AJ, 89, 1076
Singh, S., Subrahmanyan, R., Udaya Shankar, N., & Raghunathan, A. 2015, ApJ, 815, 88
Sullivan, I. S., Morales, M. F., Hazelton, B. J., et al. 2012, ApJ, 759, 17
Taylor, G. B., Carilli, C. L., & Perley, R. A. 1999, in ASP Conf. Ser. 180, Synthesis Imaging in Radio Astronomy II (San Francisco, CA: ASP)
Thyagarajan, N., Parsons, A., DeBoer, D., et al. 2016, arXiv:1603.08958
Thyagarajan, N., Udaya Shankar, N., Subrahmanyan, R., et al. 2013, ApJ, 776, 6
Thyagarajan, N., Jacobs, D. C., Bowman, J. D., et al. 2015a, ApJ, 804, 14
Thyagarajan, N., Jacobs, D. C., Bowman, J. D., et al. 2015b, ApJL, 807, L28
Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, 7
Trott, C., Wayth, R., & Tingay, S. 2012, ApJ, 757, 101
Trott, C. M., Pindor, B., Procopio, P., et al. 2016, ApJ, 818, 139
Van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Vedantham, H., Udaya Shankar, N., & Subrahmanyan, R. 2012, ApJ, 745 176
Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, ApJ, 608, 622
Zaroubi, S. 2013, The First Galaxies (Astrophysics and Space Science Library, Vol. 396; Berlin: Springer)
Zheng, Q., Wu, X.-P., Hollitt, M. J., Gu, J.-H., & Xu, H. 2016, arXiv:1602.06624v1